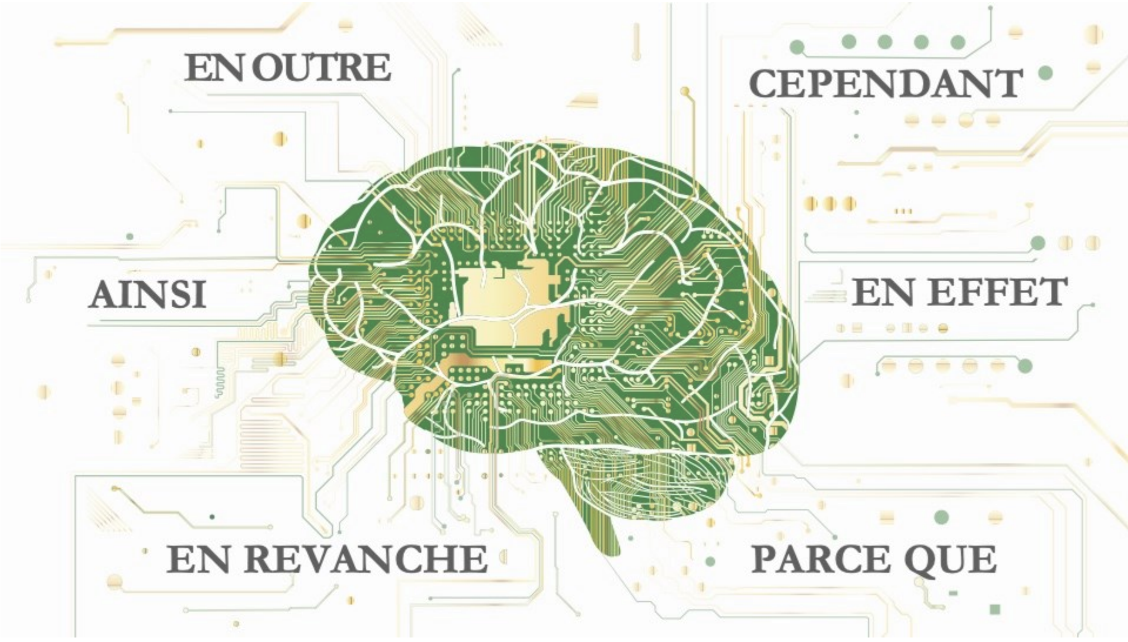
Developing the competence with discourse connectives during teenage years



by

Ekaterina Tckhovrebova

Dissertation submitted to the Faculty of Humanities at the University of Bern, in partial fulfilment of the requirements for the degree of Doctor of Philosophy

June 2023

Supervisor

Prof. Dr. Sandrine Zufferey
Institute of French Language and Literature
University of Bern

External examiner

Prof. Dr. Ted (T.J.M.) Sanders
Department of Languages, Literature and Communication
Utrecht University

To my family

Table of Contents

Acknowledgments

It is with great pleasure that I express my gratitude to all the people who have contributed to the completion of my PhD journey. It would be definitely less joyful without your unwavering support.

First of all, I would like to thank Sandrine Zufferey. Dear Sandrine, I am really grateful that you gave me the opportunity to pursue a PhD. Looking back at these four years of working together, I can most certainly say that I could not have had a better supervisor. Since my first day at the University of Bern, I felt your support and quietly assured guidance. Thank you for generously sharing with me your scientific knowledge, introducing me to the fascinating world of experimental linguistics, believing in me, and supporting me in so many ways. There is a long list of other things that I could mention here, but they will not fit this page. So, I will just say that you have been a real mentor to me in various aspects of my life, and I cannot be grateful enough for this.

Ted Sanders. Thank you for accepting to be an external examiner for this thesis. Your own work as well as the work of all the numerous linguists that you have raised were an inspiration for me.

Pascal Gygax. Thank you, Pascal, for everything you taught me about statistics, for your detailed feedback on every paper, and for being always available in spite of your intense schedule of public activities. Thank you also for your sense of humour that always makes the routine much more cheerful.

Elena Tribushinina. I am really happy that I had an opportunity to work with you and to learn from you. Thank you for your insightful comments on our papers, for your precise methodological approach, and sound advice over the rounds of reviews.

I must also mention all the school teachers in Switzerland and in Saint Petersburg who were interested in my experiments and let me conduct them in their classes. Needless to say, without your collaboration, this thesis would not have been realised.

I would also like to extend my gratitude to the linguistics team of the University of Bern, including Mathis, Jennifer, Ruedi, Sandra, and Baptiste. I am happy that I was part of such a positive and hardworking team. Thank you for all the lunches and laughs together. It was a real pleasure to have such supportive and motivating colleagues by my side.

Mathis: thank you for being my partner in crime in this SNF project, for all the discussions and jokes about connectives, linguistics and life in general. I was really lucky to go through this PhD adventure together with you.

Jennifer: thank you for being my first PhD buddy, my guide to the life of the University of Bern and to the mysteries of Bridge. It is always a pleasure to have a good laugh as well as a deep talk with you. Baptiste: thank you for helping me with the annotation of the data from the experiment on alterative list signals. I really appreciate your meticulous and diligent collaboration.

Thank you also to my colleagues from the literature department, Matthieu and Helder, with whom we have shared many pleasant conversations over the semester dinners. I want to thank especially Matthieu for all the coffee chats and day-to-day support. Your sense of humour and integrity have definitely made the routine work more enjoyable.

Many thanks also to my colleagues from the CSLS, Danielle and Matthew, for re-reading and commenting some of my chapters.

Special thanks also to the PhD colleagues from the University of Fribourg, Justine and Alizée. It was a great experience and a lot of fun to organise a workshop together.

Thank you to all my friends outside the university, and especially to Anya, on whom I can always count and who can always give a good piece of advice.

To my family, and especially mum, dad, brother, and grandmother. It might sound clichéd, but there are no words that can express how grateful I am for everything you do for me. Thank you for being there for me even at distance, for day-to-day encouragement, and a healthy dose of constructive criticism. Without your love, support, and belief in me, everything would be more difficult. Нет таких слов, которые могли бы выразить всю мою благодарность за то, что вы для меня делаете. Спасибо за то, что вы рядом даже на расстоянии, за ежедневную поддержку и необходимую долю конструктивной критики. Без вашей любви, опоры и веры в меня всё было бы намного труднее.

Finally, Alberto. Thank you for being by my side no matter what, for listening and truly understanding all my struggles and joys, for always finding the right words to cheer me up, and for encouraging me to do more than I thought I could.

Zurich, June 2023

Abstract

Built upon a combination of off-line production experiments and corpus analysis of data, this thesis provides a multifaceted investigation into the level of competence with discourse connectives during teenage years. Firstly, it discerns endogenous factors, related to the linguistic characteristics of connectives, that may render some of them more difficult than others. Secondly, it investigates exogenous factors, related to inter-individual differences in linguistic competence between young speakers, that may influence their mastery of discourse connectives. Thirdly, the current work also explores whether the mastery of connectives hinges on the context in which they are used. Furthermore, this thesis raises the question of young speakers' sensitivity to different types of coherence signals spread in discourse, and whether they are receptive to non-connective cues when inferring coherence relations. Finally, this work also focuses on the potential for improving competence with connectives through learning activities. The evidence obtained in this work serves to both deepen our understanding of the development of the competence with connectives during a transitional period of teenage years and to enrich the ongoing discussion about the type of meaning that connectives encode, namely, whether this meaning is procedural, conceptual or a combination of both.

Introduction

Results of international examinations state that the reading comprehension skills of schoolchildren are deteriorating. The Programme for International Student Assessment (PISA) reported in 2018 that more than ten million 15-year-old students from all over the world had difficulties with reading, as they were unable to complete even the most basic reading tasks (Schleicher, 2019). The Progress in International Reading Literacy Study (PIRLS) provided similar results, showing a decline in average reading achievement between 2016 and 2021 among the majority of fourth-year students from more than 50 countries (Mullis et al., 2023). Similarly, the National Assessment of Educational Progress (NAEP) in the United States revealed a significant decrease in the average reading score at both fourth and eighth grade in 2022. This trend is worrying because reading comprehension is at the heart of educational processes, as it is through reading that most of the knowledge on various subjects is transferred. As a result, poor reading comprehension skills can be detrimental for general academic success across different domains, such as natural sciences (see, e.g., Akbaşlı et al., 2016; Korpershoek, 2015; O'Reilly & McNamara, 2007), history (see, e.g., Beek, 2020), and mathematics (see, e.g., Fuentes, 1998; Salihu et al., 2018).

Among factors affecting reading comprehension is the knowledge of discourse connectives. Understanding connectives, elements that are crucial for linking parts of discourse and making explicit the underlying coherence relations, generally facilitates on-line (see, e.g., Deaton & Gernsbacher, 2000; Gaddy et al., 2001; Millis & Just, 1994; Sanders & Noordman, 2000) and off-line (see, e.g., Degand & Sanders, 2002; Lorch & Lorch, 1986; Meyer et al., 1980) reading comprehension. For instance, a study by Britton et al. (1982) demonstrated that texts including connectives are processed faster than those without connectives. Degand et al. (1999) reported that readers also perform better on off-line comprehension questions after reading expository texts with causal and consequence connectives than without them.

Considering the central place of connectives for reading comprehension, it is important to gain a better understanding of the level of competence in these linguistic items in young speakers who are exposed to connectives on daily basis. Connectives are widely spread in textbooks of various school subjects, and their mastery can therefore impact the comprehension of these subjects. Here is an extract from a history textbook used by the 10th grade students (13–14 years) in the French-speaking part of Switzerland that includes the concessive connective *bien que* 'even though' (1). This connective is crucial for the correct interpretation of the meaning of this sentence, as it is challenging to infer the relation of concession when it is not explicitly marked. As a result, not knowing the function of this connective may affect the ability to learn the content of the school subject.

(1) Bien que les cantons suisses ne combattent pas directement dans la guerre de Trente Ans, ils sont mentionnés dans les Traités de Westphalie.

'Although the Swiss cantons did not fight directly in the Thirty Years' War, they are mentioned in the Treaties of Westphalia.'

Notwithstanding their importance for the comprehension of written learning material, and hence, academic success, only few studies have examined how the ability to use and understand connectives progresses during teenage years (see Kleijn et al., 2019; McClure & Geva, 1983; Nippold et al., 1992; Van Silfhout et al., 2015; Zufferey & Gygax, 2020b, for the exceptions). However, it is during this period that readers are building a more advanced general linguistic proficiency (see, e.g., Berman, 2004a; Berman & Slobin, 1994; Berman & Verhoeven, 2002), involving a comprehensive mastery of a wide range of lexical and grammatical elements, as well as the ability to adapt their linguistic choices to various communicative contexts (Berman, 2004b).

Research questions of the thesis

The current thesis attempts to fill the gap between studies on childhood and adulthood, and to establish the level of competence with discourse connectives during teenage years (RQ1, Chapters 1–7). A better understanding of how the mastery of connectives develops during teenage years is achieved by investigating it from several perspectives. First, this thesis identifies endogenous factors, related to the characteristics of connectives, that may render some of them more difficult than others (RQ2). Factors such as their frequency in corpora (Chapters 3–7), the mode in which connectives are typically used (Chapters 4, 5 & 6), the cognitive complexity of coherence relations that they encode (Chapters 3, 4, 5 & 7), and their mono- or polyfunctional nature (Chapters 3, 5 & 7) are assessed across the experiments presented in this thesis.

A second aim is to trace exogenous factors, related to inter-individual differences in linguistic competence between young speakers, that may impact their mastery of discourse connectives (RQ3). Among the examined factors are namely cumulative linguistic competence, as revealed by speakers' academic background (Chapter 3), degree of exposure to print (Chapters 3–7), vocabulary level (Chapters 4 & 5), and grammatical competence (Chapter 3).

Third, the current work also explores whether the mastery of connectives depends on the context in which they are used. Chapter 3 examines whether the ability to use connectives hinges on the more or less close-to-real-discourse context of texts and pairs of sentences, respectively (RQ4). Chapter 6 raises the question of young speakers' sensitivity to different types of coherence signals spread in discourse, and whether they are receptive to non-connective cues when inferring coherence relations (RQ5). More precisely, it examines teenagers' sensitivity to alternative signals of the list relation, and how this sensitivity is modulated by the presence of connectives varying in frequency and expressing compatible or non-compatible coherence relations.

Finally, this thesis also focuses on examining whether the competence with connectives can be enhanced by learning activities (RQ6, Chapter 7). Revealing endogenous and exogenous factors, accounting for the difficulties in the mastery of connectives, and whether and how the latter can be improved will also allow us to shed light on the linguistic nature of connectives. The evidence obtained in this work enriches our knowledge about the type of meaning that connectives express, by contributing to the discussion about whether they encode procedural meaning, conceptual meaning or both.

All these research questions are assessed using evidence from the combination of experimental and corpus analysis. Through the use of experimental methods, I examined the performance of teenagers with specific types of connectives, while isolating and controlling for the factors that may affect it. The corpus analyses, in turn, complement the experimental methodology, by providing necessary information on the behaviour of the tested connectives in language use. More precisely, it facilitates the selection of the most suitable connectives to assess the role of the above-mentioned endogenous factors, by establishing the connectives' frequency in oral and written language, and the distribution of different functions of polyfunctional connectives. Ultimately, the current examination of teenagers' performance with connectives is mostly based on production tasks, as the correct use of connectives in discourse represents a particular challenge for young speakers (Cain & Nash, 2011). Moreover, the ability to use appropriate connectives in written essays constitutes a separate criterion for evaluating the quality of the students' writing, and therefore, was an additional reason to examine this ability more in depth.

Structure of the thesis

The present work is organised as follows. **Chapter 1** provides the theoretical basis for the following chapters of the thesis. More specifically, it presents different theoretical approaches to discourse coherence, and to the meaning that connectives convey.

Chapter 2 provides an overview of empirical research on the acquisition of discourse connectives, and presents factors that account for the difficulties in using and understanding them during childhood and that even remain for a portion of adults. After reviewing the factors affecting the performance with connectives in children and adults, this chapter explains why the teenage period occupies a particular place for linguistic development in general, and for the development of a more fine-grained proficiency with discourse connectives in particular. This chapter is concluded by the formulation of hypotheses for the main research questions of the current thesis.

Chapter 3 opens the empirical part of the thesis, and reports a series of experiments addressing research questions 1–4. This chapter assesses the use and comprehension by teenagers of four French connectives, typically used in the written mode and varying in frequency, polyfunctionality, as well as the cognitive complexity of the coherence relations they signal. The inter-individual variation in linguistic competence among participants is evaluated with exposure to print tests, grammatical test, and academic background. Finally, sentence- and text-cloze tests were conducted to examine the role of context for the use of discourse connectives, and the results of the text comprehension test provided evidence on the difference between production and comprehension of connectives.

Chapter 4 continues the examination of factors accounting for the use of connectives during the teenage period, by increasing the number of tested connectives to twelve, including connectives from both oral and written modes, and restricting their choice only to monofunctional ones (RQ 1–2). Based on the results of Chapter 3, different measures of individual variation were included in this study, namely, a vocabulary level test and an adapted version of the exposure to print test (RQ3).

In order to obtain more solid and generalisable evidence on the level of connective mastery in teenage years, **Chapter 5** replicates the study from Chapter 4 in Russian. Specifically, it assesses the use of twelve equivalent mono- and polyfunctional connectives from oral and written modes (RQ 1–2) by young Russian speakers with different degrees of exposure to print and lexicon size (RQ3).

Considering that connectives are not the only type of linguistic cues signalling coherence relations in discourse, **Chapter 6** extends the current investigation to alternative signals of coherence relations (RQ5) and examines whether the sensitivity to alternative and less prototypical signals of the list relation is impacted by the presence of more salient and prototypical coherence cues such as connectives. Furthermore, measures of exposure to print from Chapter 4 are included to assess if the sensitivity to non-connective list signals varies in young speakers according to individual differences in linguistic competence (RQ3).

Chapter 7 pursues the study of factors that influence the use of connectives during teenage years. This chapter focuses on particularly infrequent connectives used predominantly in writing and the connectives signalling less typical coherence relations (RQ 1–2). In addition, this chapter also explores whether the use of connectives can be improved with a learning activity, consisting in explaining their functions and providing exercises to train students (RQ6), or whether it is implicit exposure to the written language as measured by the adapted versions of the ART, that allows development of the mastery of connectives (RQ3). Finally, **Chapter 8** presents an overview of the major findings reported in this work, discusses their implications and limits, and proposes possible avenues for the future research.

Preliminary notes

- Several chapters were published or submitted for publication as individual papers. Although all the papers have co-authors, the development of main ideas, set-up and execution of experiments as well as analysis and description of the results were my own. Publishing information is provided at the beginning of every chapter that is concerned.
- The present study is funded by Swiss National Science Foundation Grant 100012_184882 and adheres to its ethical principles of research on human subjects.

1. The Nature of Discourse Connectives

This chapter provides a theoretical foundation for understanding the role of connectives in discourse. It starts by presenting different approaches to coherence relations and specifies which framework is adopted in the present thesis. Subsequently, this chapter explores linguistic parameters that allow to distinguish between different types of connectives. These parameters include the cognitive complexity of the coherence relations that connectives signal, the number of functions they can have, the mode in which they are typically used, and their domain of use.

1.1. The role of connectives in discourse

A difference between a set of unrelated sentences and a text is that units of text are linked in a coherent way. Coherence is an important element that assures text's intelligibility and allows to construct a comprehensible mental representation of discourse[1] (Blass, 1990; Charolles, 1983; Sanders et al., 1992). Two different approaches can be distinguished on how the coherence of a text is realised. The first approach analyses the organisation of content in discourse, by looking at whether there is sense, topic or reference continuity (see, e.g., De Beaugrande & Dressler, 1996; Blass, 1990; Giora, 1985). For the followers of this approach, it is important to distinguish the notions of cohesion and coherence as two distinct types of connectedness present in discourse. Cohesion is described as connectedness between elements of discourse at a formal, surface level. Coherence instead refers to the connectedness at a deeper level of semantics and pragmatics.

The second approach to coherence rather focuses on the relations connecting the units of discourse (Ballard et al., 1971; Halliday & Hassan, 1976; Knott & Dale, 1994; Mann & Thompson, 1988; Sanders et al., 1992; Van Dijk, 1977). According to the Rhetorical Structure Theory (RST), developed by Mann and Thompson (1988), for instance, a hierarchical structure of a text holds together because of the relations between its parts. Halliday and Hassan (1976) rather refer to a within-text cohesion that builds on semantic relations between its units. The idea of a certain type of relations that are present between segments of discourse was proposed by many linguists from different fields, such as comparative linguistics (Ballard et al., 1971; Grimes, 1975), structuralist discourse analysis (Roulet et al., 1985), systemic grammar (Martin, 1983, 1992), computational linguistics (Asher & Lascarides, 2003; Grosz & Sidner, 1986; Hobbs,

[1] I use the terms *text* and *discourse* as synonyms in the present thesis, referring to all types of oral and written communication.

1985; Knott & Dale, 1994; Lascarides & Asher, 2007), and psycholinguistics (Sanders et al., 1992).

Different theoretical frameworks developed different terms, referring to these relations, depending on the focus of their research, such as *deep grammar relations* (Ballard et al., 1971), *rhetorical relations* (see, e.g., Asher & Lascarides, 2003; Grimes, 1975; Grosz & Sidner, 1986), *relational propositions* (Mann & Thompson, 1986), *semantic relations* (Halliday & Hassan, 1976) and *coherence relations* (Hobbs, 1979, 1985; Sanders et al., 1992). The present thesis follows the second approach to coherence and uses the term *coherence relations*, when referring to the relations that exist between discourse segments of various levels.

Different approaches to coherence and to the annotation of coherence relations also propose different inventories of those relations. These inventories vary in a number of relations that are distinguished as well as in types of labels attributed to conceptually the same coherence relations. For instance, the Cognitive Coherence Relation (CCR) model, developed by Sanders and colleagues (1992), distinguishes approximately 12 relations and the RST (Mann & Thompson, 1988) includes around 20 relations. The Penn Discourse Treebank (PDTB-2.0) (Prasad et al., 2008) and the Segmented Discourse Representation Theory (SDRT) (Asher & Lascarides, 2003) count more than 30 relations each and the RST Discourse Treebank (RST-DT) (Carlson et al., 2003) includes more than 50 relations. In contrast, Hobbs' (1985) proposes a taxonomy of 8 relations and Grosz and Sidner (1986) distinguish only two types of relations. Such variation in the inventories of coherence relations is explained by differences of granularities between frameworks. For instance, in the theory of Grosz and Sidner (1986), there was made a distinction between a coordinating relation of "satisfaction-precedence" and a subordinating relation of "dominance". In comparison, Carlson and colleagues (2003) divide coherence relations into 17 different classes, such as cause, elaboration or condition, each of which includes at least two types of coherence relations.

Furthermore, even if conceptually similar relations are inventorised in different frameworks, such relations do not always share the same names. For instance, the PDTB-2.0 includes a general class of expansion relations, while the RST-DT labels a similar type of relations as elaboration relations. The result relation from the SDRT, in turn, could be mapped onto the reason relation in the PDTB-2.0. This thesis follows the CCR model (Sanders et al., 1992) when referring to different types of coherence relations. This model offers a structured approach by breaking down each coherence relation into distinct features, or primitives. This systematic decomposition facilitates the establishment of connections between different labels of coherence relations and enables the scaling of each relation type in terms of cognitive complexity (Sanders et al., 1992; Sanders et al., 2018). A more detailed account of the CCR model is provided in Section 1.3.1.

Coherence relations can be implicit, like the consequence-cause relation in (1); or they can be explicitly expressed by different types of signals.

(1) July went to bed early last night. She was really exhausted.

A corpus analysis, conducted by Das and Taboada (2018) on the RST Discourse Treebank (Carlson et al., 2002), allowed the authors to distinguish at least nine types of signalling devices, such as connectives, reference, semantic, morphological, lexical, syntactic, graphical, genre, and numerical features. This thesis mainly focuses on one specific type of signals of coherence that is connectives.

Connectives are lexical items that render explicit coherence relations between parts of discourse, such as consequence or contrast (e.g., Mann & Thomson, 1986; Roulet et al., 1985; Sanders, 1997). For example, a cause-consequence relation, linking the two segments of the sentence (2), is marked by the connective *so*; and the relation of contrast in (3) is signalled by the connective *but*.

(2) Paul ran out of milk, *so* he went to the supermarket.

(3) Lucy is a talented piano player, *but* her brother is not.

There are different terms referring to the linguistic devices that were previously named *connectives*, such as *discourse markers* (see, e.g., Fraser, 1990; Iten, 2000; Schiffrin, 1987), *discourse connectives* (see, e.g., Blakemore, 1987; Sanders et al., 1992; Sperber & Wilson, 1986, 1995), *cue phrases* (see, e.g., Knott & Dale, 1994), *conjunctions* (see, e.g., McClure & Steffensen, 1985; Prasad et al., 2008; Prasad, Joshi & Webber, 2010), *sequencers* (Hempel & Degand, 2008), *pragmatic operators* (Ariel, 1998), *indicating devices* (see, e.g., Katriel & Dascal, 1984) and many more. Each term reflects a theoretical framework, in which it is applied, as well as angle and granularity of the studied phenomenon. *Discourse markers*, for example, is a hyperonym term that can be applied not only to connectives, but to a wider range of markers like modalizers, turn-regulators, and reformulators (Pons Bordería, 2001; Schiffrin, 1987). The terms *conjunctions* and *sequencers*, in contrast, describe more specific categories inside the class of connectives. Hempel and Degand (2008), for instance, use the term *sequencers* when describing lexical elements, used in series and introducing a new sequence of text. The authors distinguish sequencers of spatial (*on the one hand ... on the other hand ...*), temporal (*first ... then ... finally*), and enumeration (*firstly ... secondly ...*) nature. As for the term *conjunctions*, it rather refers to a specific syntactic function of certain connectives like coordinate and subordinate conjunctions *or* and *because*. Throughout this thesis, I will use the term *discourse connectives* when referring to all types of lexical items that tie segments of discourse of different levels. The choice of this term over others is motivated by several reasons. First, it is in line with the cognitive account of coherence, developed by Sanders and colleagues (1992), on which the present work is based. Second, this term reflects well the nature of the linguistic items in question, as their main purpose is to connect parts of discourse, making explicit underlying coherence relations. Third, the term *connective* is the best suited because of the granularity of the linguistic phenomenon that it describes, as it is neither too broad, like the term *discourse marker*, nor too narrow, like the term *sequencer*.

1.2. The meaning conveyed by connectives

According to a number of linguistic theories, connectives are not just regular lexical elements, as they play a more particular role in a sentence than a simple content word. In the Gricean approach to implicatures, connectives are described as a tool allowing to perform higher-order speech acts over lower-order speech acts, by relating them between each other (Grice, 1989). For instance, in the example (4a), Grice distinguishes two lower-level speech acts about the brother-in-law (4b) and the aunt (4c) and the third higher-level speech act that indicates that there is a contrast between the two lower-level acts (4d). The connective *on the other hand* thus is a signal of the higher-level speech act (4d), performed over (4b) and (4c).

(4a) My brother-in-law lives on a peak in Darien; his great aunt, *on the other hand*, was a nurse in World War I (Grice, 1989: 361).

(4b) My brother-in-law lives on a peak in Darien.

(4c) His great aunt was a nurse in World War I.

(4d) (b) and (c) are in contrast between each other.

A slightly different vision of the meaning expressed by connectives, was introduced by Diane Blakemore (1987) into the framework of Relevance Theory (Sperber & Wilson, 1986). She distinguished words like *cat* and *talk* that encode a conceptual meaning from other words, like discourse connectives *therefore* and *but*, that rather give procedural instructions about how conceptual representations should be manipulated in a text. In the sentence (5), for instance, the connective *but* indicates that the upcoming proposition provides evidence against an implication that in summer the weather typically is warm.

(5) It's summer, *but* it's not warm outside.

In other words, conceptual expressions belong to the so-called *language of thought* (Fodor, 1975), i.e., a system of conceptual representations, and normally are accessible to consciousness. In contrast, procedural expressions trigger cognitive procedures that are part of the *machine language* (Wilson, 2011), which is not accessible to consciousness and is difficult to conceptualize. As a result, the meaning of procedural expressions is more difficult to grasp and explain (Wilson & Sperber, 1993). In a more recent account of Relevance Theory, Wilson (2011) however notices that conceptual and procedural meaning are not mutually exclusive and that majority of conceptual items also encode procedures. This procedural meaning "… might become more specific over time, to a point where the original conceptual content becomes entirely redundant" (Wilson, 2011, p. 18). Moreover, Wilson suggests that procedural expressions have a more general function of triggering, not necessarily related exclusively to the guiding of the comprehension process, but also to other capacities, such as "mindreading, emotion reading, social cognition, parsing and epistemic vigilance" (Wilson, 2011, pp. 26–27).

The work of Ducrot and his colleagues (Anscombre & Ducrot, 1997; Ducrot, 1972, 1980) also referred to a similar distinction between two types of meaning. Their claim however was that majority of words in a language express both a conceptual content and, what they call, an *argumentative orientation*, which is similar to the idea of procedural meaning from Relevance Theory. They illustrate the fact that conceptual and procedural meaning are not mutually exclusive with an example of French expressions *peu* 'little' and *un peu* 'some' (Ducrot, 1972). On the one hand, these expressions encode conceptual information about quantity. On the other hand, they orient speakers towards opposite interpretations for the upcoming segment of discourse. Saying that Martin drank little (*peu*) wine (6) orients the hearer towards an assumption that he shows a sober behaviour, while saying that he drank some (*un peu*) wine (7) provides an opposite orientation.

(6) Martin a bu *peu* de vin hier. Il semble devenir sobre.

'Martin drank <u>little</u> wine yesterday. He seems to be getting sober.'

(7) Martin a bu *un peu* de vin hier. Il semble devenir moins sobre.

'Martin drank <u>some</u> wine yesterday. He seems to be getting less sober.'

Further development of the conceptual-procedural account of meaning, applied to discourse connectives, took two diverging directions. One current of research (e.g., Carston, 2002; Sperber, 2001) supported the claim that linguistic items cannot encode both procedural and conceptual meaning at the same time. As a result, they distinguished connectives conveying conceptual meaning (such as *and, if*, and *or*) from the connectives having only procedural meaning (such as *although, whereas,* and *since*).

Other researchers (Fraser, 2006; Hussein, 2008; Mauri & van der Auwera, 2012; Moeschler, 2002, 2005) supported the view that connectives encode both concepts and procedures. According to Fraser (2006), for instance, connectives are not devoid of conceptual meaning, as, for instance, such connective as *so* can express at least four different concepts. This connective can encode concepts of consequence (8a), logical inference (8b), contextual inference (8c), and purpose (8d).

(8a) Jack was forced to work overtime. *So* he quit his job.

(8b) He likes sweets, *so* he has to like chocolates.

(8c) The movie was over *so* we didn't bother hurrying.

(8d) She shut the door *so* the cat couldn't get out. (Fraser, 2006, p. 27)

Moeschler (2002, 2005) also supported the view that connectives encode both types of meanings and occupy an intermediate position between conceptual and procedural information. However, he rather represented conceptual and procedural meanings as two intersecting poles. Therefore, depending on where a connective is situated with respect to these poles, it can have a stronger or a weaker conceptual meaning. For example, the French connective *et* 'and' gives procedural information about the order of events and has a weak conceptual meaning, as this connective does not trigger one specific coherence relation, but rather can be used to express multiple relations, such as addition, temporality, cause etc. As a result, it can be placed closer to the centre of procedural pole. The connective *parce que* 'because' carries procedural information on the order of events, but it also has a strong conceptual meaning, as it is used to signal only one type of coherence relation, namely causality. Therefore, this connective can be situated in the middle between conceptual and procedural poles.

The idea that connectives may occupy an intermediate position between conceptual type of information and procedural one is also endorsed by neurolinguistic theories. The declarative-procedural model of language, developed by Ullman (2001), argues, for instance, that the use of language is supported by two different brain memory systems that are responsible for distinct cognitive mechanisms and linguistic skills. Associative, or declarative, memory supposedly is rooted in temporal lobe and is accountable for learning arbitrary related information, such as associative binding of words and irregular morphology, and conscious use of mental lexicon. Procedural memory, in contrast, probably is embedded in frontal- and basal-ganglia regions and is associated with unconscious learning of mental rules governing the principles of regular morphology and hierarchical syntactical structures. Thus, it is possible that connectives belong to both types of brain memory systems. On the one hand, using connectives in discourse as signalling devices can be considered as part of procedural knowledge that is unconscious and automatized in speakers with greater experience. On the other hand, connectives are lexical items and, therefore, are part of declarative knowledge.

1.3. Classification of connectives

1.3.1. Cognitive complexity

As mentioned previously, there are frameworks that distinguish between connectives encoding different types of meaning (conceptual or/and procedural) and the degree to which this meaning is present in a connective (weak versus strong). However, there are also other ways that allow to make distinction between connectives. According to the Cognitive Coherence Relation (CCR) model (Sanders et al., 1992; see also Sanders et al., 2018 for a more recent account of the model), connectives can also be classified depending on the cognitive complexity of the coherence relation that they encode. The CCR model suggests that coherence can be analysed in terms of five primitives (Evers-Vermeul et al., 2017; Sanders et al., 2018), namely *basic operation* (additive versus causal), *source of coherence* (objective versus subjective), *implication order* (basic versus non-basic), *polarity* (positive versus negative), and *temporality* (temporal versus non-temporal).

Basic operation refers to an underlying relation between two segments of discourse that can be additive or causal. According to the CCR model, an additive relation is a relation of a logical conjunction that weakly connects two segments, while a causal relation is a relation of implication that creates a strong connection between segments of discourse. The primitive of the coherence source distinguishes objective from subjective relations, named pragmatic and semantic in the initial account of the CCR model (Sanders et al., 1992). An objective relation connects two segments at the propositional content level and does not require a speaker's participation in the creation of this relation. The speaker simply reports two real-world situations in two segments of discourse. In a subjective relation, two segments are connected at the illocutionary level, and a speaker actively participates in the construction of this relation, by performing speech acts or by reasoning.

The implication order, initially labelled as the order of segments (Sanders et al., 1992), concerns only causal relations because additive relations are symmetrical. The implication order is considered as basic if the first segment includes an antecedent and the second segment has a consequent. In a relation with a non-basic order, it is rather a consequent that precedes an antecedent. On the dimension of polarity, positive relations are opposed to negative ones. A relation is negative when negated counterparts of propositions P or Q, – i. e., not-P or not-Q, – are implicated in the relation between two segments. A positive relation is the relation, in which both P and Q are not negated. Finally, temporality is the primitive that was not distinguished in the first version of the CCR model and was added later (Evers-Vermeul et al., 2017; Sanders et al., 2018). This primitive indicates whether a relation between two segments is temporally ordered or not. The relation is non-temporal if the temporal order of segments is not essential. The relation is temporal, in contrast, when two segments are ordered in time and their order is relevant for the interpretation of the relation between the segments. Such relations can be chronological, anti-chronological, or simultaneous.

One facet of each primitive is considered to be cognitively more challenging than the other. For example, negative, causal, subjective, non-basic-order, non-temporal relations supposedly are cognitively more difficult to process and comprehend than positive, additive, objective, basic-order, temporal relations. Thus, concessive relations, which have negative polarity, are considered to be cognitively more complex than causal relations, which have positive polarity. Causal relations, in their turn, are deemed to be more challenging than additive ones, since they differ on the dimension of basic operation.

1.3.2. Number of functions

Not all connectives signal exclusively one coherence relation though. Depending on the number of coherence relations encoded by a connective, a further distinction can be made between monofunctional and polyfunctional connectives. Monofunctional connectives encode only one type of coherence relation. This is the case of the English connective *moreover* that typically signals the relation of addition (Das et al., 2018). Polyfunctional connectives, in contrast, signal multiple coherence relations, depending on the context in which they are utilized. For instance, the connective *since* can convey not only the causal relation, as in (9), but also the temporal relation, as in (10).

(9) Audience did not want to leave *since* it was the last concert of their favourite singer.

(10) They have not seen their daughter *since* she went to the university.

1.3.3. Language modality

If we look at connectives not from the perspective of coherence relations that they can express, but rather considering how they are used in a language, we can differentiate between connectives that are typically used in oral modality and those that are mostly used in written modality. Although every lexical item technically can be used in both contexts, there is a number of characteristics that function differently in oral and written language (see, e.g., Chafe, 1982; Halliday, 1987; Horowitz & Samuels, 1987). With regard to connectives, a more reduced number of them appears in the oral speech in comparison to the writing. Moreover, not only their repertoire is less diverse, but they are also used with less precise functions (Biber, 2006; Crible & Cuenca, 2017). For instance, Crible and Cuenca (2017) report that the connective *and* is used almost exclusively in one sense in the written modality, namely to mark the additive relation (90.76% of uses). In contrast, in speech, this connective is used in a greater number of functions, among which the signalling of the additive relation represents 57.11% of occurrences.

1.3.4. Domain of use

Finally, connectives can also be distinguished on the basis of the domains in which they can be used. According to Sweetser (1990), relations between parts of discourse linked by a connective correspond to relations between three domains of human cognition. There are content domain, epistemic domain, and speech act domain, respectively related to socio-physical world, mental states and communication (see Zufferey, 2010 for a detailed argumentation in favour of a tripartite distinction between the domains).

(11a) Laura a faim *parce qu*'elle n'a rien mangé depuis le matin.

‘Laura is hungry <u>because</u> she hasn't eaten anything since the morning.’

(11b) Laura quitte le bureau de plus en plus tôt *parce qu*'elle n'est plus intéressée à faire une carrière.

'Laura is leaving the office earlier and earlier because she is no longer interested in having a career.'

(11c) Tu veux entrer? *Parce qu'*il fait froid.

'Do you want to come in? Because it's cold.'

For instance, one of the connectives that can appear in all three domains of use is the French connective *parce que* 'because'. In example (11a), the fact of not eating leads to a physical reaction of being hungry, which means that this use of *parce que* matches the content, or physical world, domain. Example (11b), illustrates the use of this connective in the epistemic, or mental states, domain, as the fact that Laura does not stay long hours in the office leads a speaker to infer that she is not motivated to make a career. The sentence (11c) corresponds to the speech-act, or communication domain, as the segment following the connective *parce que* is a rationale for producing the speech act or question in the preceding segment.

1.4. Conclusion

This chapter has provided a theoretical background on different approaches to discourse coherence as well as on the linguistic nature of connectives. The next chapter makes an overview of empirical studies, focusing on the particularities of the acquisition and development of the competence with discourse connectives, and provides a rationale for the current work.

2. Empirical Evidence on the Developing Competence with Discourse Connectives

In this chapter, I review empirical studies examining how discourse connectives are acquired in the course of first language acquisition, and how this competence is further developed during the adult lifespan. I especially focus on different factors that impact the acquisition and development of the competence with connectives at three developmental stages, namely during childhood (0–11), teenage years (12–18), and adulthood (over 18). I will refer to these terms for convenience throughout the thesis, even though the distinction between these three periods is not clear-cut.

2.1. Connective acquisition during childhood

During childhood years, several factors seem to have the most impact on the acquisition of connectives. These factors are both related to the characteristics of connectives and to inter-individual differences between young speakers. I first present the linguistic predictors related to connectives themselves, such as the cognitive complexity of coherence relations encoded by connectives, connectives' frequency in the child's input, and differences between their domains of use. Afterwards, I mention factors regarding individual differences between children, namely age, memory, the socio-economic status of their family, and academic background.

2.1.1. Cognitive complexity

Already in early childhood, the cognitive complexity of coherence relations (Sanders et al., 1992), signalled by connectives is a strong predictor of the acquisition of discourse connectives in several ways. First of all, it is an important predictor for the order of acquisition between coherence relations. The analysis of spontaneous oral productions of children aged 2 to 3, made by Bloom et al. (1980), showed, for instance, that children produced on average additive relations first (1), followed by temporal (2), then causal (3), and finally adversative relations (4). All examples come from Bloom et al. (1980, pp. 7–8).

(1) Maybe you can carry that <u>and</u> I can carry this.

(2) I going this way to get the groceries <u>then</u> come back.

(3) Get them <u>cause</u> I want it.

(4) … cause I was tired, <u>but</u> now I'm not tired.

The authors explained the observed order of production of different types of relations by what they call a cumulative principle. To put it another way, it means that some relations involve less semantic components than others and are therefore acquired earlier. For example, adversative relations, including meaning components of addition, temporal order and opposition, are produced after temporal relations, which only have additive and temporal order components. However, the data of this study revealed that children were not consistent in the order of production of different types of connectives encoding these relations.

In order to explain the variation in the order of production (and acquisition) of different types of connectives in the data by Bloom et al. (1980), Evers-Vermeul and Sanders (2009) proposed a more nuanced approach, suggesting that the order of acquisition of connectives is predicted by the cumulative cognitive complexity of coherence relations, based on the CCR model (see Section 1.3.1 for a description of this model). According to this approach, complexity of each coherence relation can be estimated based on the primitives from the CCR model, where one dimension of each primitive is considered to be cognitively more complex (hereafter marked by "+") than another (hereafter marked by "–"). Thus, based on the primitive called basic operation, causal (+) connectives should not be produced before additive (–) ones, as the dimension of causality is deemed to be more complex. Further, based on the primitive of polarity, it is negative (+) connectives that should not emerge before positive (–) ones. The analysis of recordings from 12 Dutch native speakers, aged 1 to 5, indeed showed that on the primitive of basic operation, all children produced the cognitively less complex additive connective *en* 'and' (–) before the more complex causal *want* 'because' (+). As regards the primitive of polarity, the majority of children also produced the less complex positive connective *en* 'and' (–) before the more complex negative connective *maar* 'but' (+).

Referring to the variation in the order of production of connectives, the authors hypothesized that it can occur when there is an interaction between different primitives. For instance, based on the interaction between basic operation and polarity, positive additive (–, –) connectives can be followed by both positive causal (–, +) or negative additive (+, –) ones. Since the CCR model does not make assumptions about the relative complexity between the dimensions of different primitives, such as negative polarity and causal basic order, both types of connectives are deemed to have the same degree of complexity. They have therefore the potential to appear in children's speech approximately at the same time, which explains the variation in the order of production of connectives. The results supported the initial hypothesis, and indeed revealed a variation in children's productions when an interaction between the two primitives was involved. The negative additive connective *maar* 'but' (+, –) was produced before the positive causal *want* 'because' (–, +) by nine children, while the order was reverse in the productions of three other children.

The cognitive complexity of coherence relations does not only predict the order of production of connectives in early childhood, but also the difficulty in the comprehension of different types of coherence relations in older groups of children. Spooren and Sanders (2008), for instance, assessed the comprehension level of children aged 8 to 12 with a multiple-choice sentence-continuation task of two types. The first one included a discourse passage, formulating a general rule, which had to be completed with one of the two possible continuations, as in (5). The second type involved a discourse passage, presenting a statement, and participants had to choose out of two options a general rule that corresponded the best to this statement, as in example (6).

(5) People who are lok wear sweaters.

Els is lok but she is . . .

a. not wearing a sweater.

b. wearing a sweater.

(6) Manel has a lot of money, but she is not bap.

a. People who have much money are not bap.

b. People who have much money are bap.

Results revealed that children had a good comprehension of connectives with a lower degree of cognitive complexity, namely with positive causal (–, +) (*and therefore*) and negative additive (+, –) (*but*) relations[1]. However, the scores for the more complex negative causal relation (+, +) (*but*), on average, were significantly lower, and younger participants (8–9) had more difficulty with this type of relation than older ones (11–12).

The results from the study of McClure and Steffensen (1985), in which children had to write a continuation after additive (*and*, *but*) and causal (*because*, *even though*) connectives with different polarities, also suggest that negative relations are more difficult to understand for speakers aged 8 to 15, as indicated by the lower number of correct continuations that they produced after the connectives signalling negative relations. However, the developmental progress in the comprehension level up to the age of 15 was observed only for the negative causal connective *even though*, representing the highest degree of complexity. The acquisition of less complex coherence relations, encoded by all other connectives (*and, but, because*), in contrast, seems to be achieved by the age of 12.

The study of Blything, Davies and Cain (2015) suggests that certain connectives, signalling the relation of temporality, are fully mastered even earlier, namely by the age of 7. The authors asked children aged 3 to 7 to watch a short cartoon, representing two actions, and then to select an action that they think happened first, by touching a corresponding image on a screen. The main goal of this task was to assess the role of the connective knowledge (*before* and *after*), the role of the order of actions (chronological or reverse-chronological) and the world knowledge on the comprehension of the

[1] Note that Spooren and Sanders (2008) did not include an inference task (second type) for the negative additive relation in their study.

temporal relation. Their results revealed that world knowledge did not predict the accuracy in the task across all the participants. As for the other two factors, the youngest group of children (3 to 4) did not understand very well the meanings of the two temporal connectives, and mostly relied on the order of appearance of the two actions on the screen to infer temporality. In contrast, the oldest group of children aged 6 to 7 had a high level of accuracy in the task, suggesting that they had a full understanding of the meanings of connectives. Similar results were obtained in the study by Blything and Cain (2016) who measured not only the accuracy, but also the reading times in an analogous touch-screen paradigm task.

In a different type of experimental design, however, it was shown that the temporal connective *after* can cause difficulties even for 12-year-olds (Pyykkönen & Järvikivi, 2012). The authors of this experiment proposed Finnish native-speaking children aged 8–12 to read pairs of sentences, linked by the connectives *ennen kuin* 'before' and *sen jälkeen kun* 'after', and then to select in a multiple-choice paradigm which of the two actions happened earlier, or if they took place at the same time. Their findings indicate that, on average, children performed lower for the items including the connective *sen jälkeen kun* 'after', than for the items with the connective *ennen kuin* 'before'. Moreover, the performance of all the children decreased even more when the connective *sen jälkeen kun* 'after' was placed in sentence-medial position (i.e., the main clause preceded the subordinate one), making the order of events presented in discourse dissimilar to the order of events in the world. In other words, children aged 7 may have a high level of comprehension of chronological and anti-chronological temporal relations, signalled by the connectives *after* and *before*, when this comprehension is measured with the tasks, not involving reading (Blything et al., 2015; Blything & Cain, 2016). However, once reading skills are necessary to successfully accomplish the task

(Pyykkönen & Järvikivi, 2012), the anti-chronological order of events as well as the position of a connective in a sentence seem to still add a layer of complexity and prevent even 12-year-olds to perform at a high level with temporal connectives.

In addition, when contrasted to the performance of adults, the performance with connectives in children is found to be poorer, even for connectives that encode simple relations from a cognitive perspective. Cain and Nash (2011) revealed, for example, that children benefit from the presence of connectives during reading at least as early as aged 7. This finding was obtained in a series of self-paced reading experiments in which the authors found that children aged 7–10 years read the sentences linked by an appropriate connective faster in comparison to sentences that did not include any connective or included an inappropriate or underspecified one (*and*). The results from the off-line tasks suggest, however, that implicit and explicit knowledge of connectives do not follow the same developmental pace. Thus, children had generally higher scores for causal (*because, so*), temporal (*after, before*) and adversative (*although, but*) connectives in a sentence cloze task, and could discriminate well between appropriate and inappropriate uses of these connectives. Moreover, there was an increase in performance between the younger (7–8 years) and the older (9–10) groups of children in both tasks. The older children scored significantly higher for temporal and adversative connectives in a cloze test and were more accurate at judging appropriate and inappropriate uses of causal and adversative connectives than the younger group of children. Finally, the performance of the older children did not significantly differ from that of adults in the judgement task. Yet, it was significantly lower than that of adults in the sentence cloze task, suggesting that the explicit knowledge of connectives continues to develop beyond the age of 10 even for less cognitively complex connectives.

When dealing with more complex coherence relations, in contrast, even at the age of 11, children were found to face difficulties with judging their appropriate and inappropriate uses or with choosing a continuation after them. Knoepke et al. (2017) revealed, for instance, that German-speaking children aged 7–11 cannot distinguish between coherent and incoherent uses of the concessive connective *trotzdem*, signalling a negative causal (+, +) relation. In contrast, their performance with the positive causal connectives *darum, daher, deshalb* and *denn* was comparable to that of adults. The authors proposed that children treated the negative causal connective as a positive causal one, because they evaluated as incoherent its coherent uses and vice versa. A similar result was obtained in the experiment of Spenader (2018) who asked Dutch-speaking children aged 7–10 to select a conclusion for sentences containing either the positive causal connective *want* or the negative causal connective *maar*. The results indicated that children had a high level of competence with the sentences containing the positive causal connective, whereas the performance with the sentences including the negative causal connective was below chance level.

2.1.2. Frequency

Besides the complexity of coherence relations, the frequency of connectives in a child's input was also found to play a role for their acquisition. Frequency of linguistic elements in the input in general was found to be important by usage-based accounts of language acquisition (e.g., Tomasello, 2000, 2003; Kidd et al., 2006). According to this theoretical framework, the input that children receive from their environment is considered to be crucial for explaining the course of acquisition between various linguistic elements.

Diessel (2004) found, for instance, that there was a correlation between an overall frequency of a clause introduced by a specific connective in the mother's speech and the order of appearance of this clause in the speech of five English-speaking children aged 1;8–5;1[2]. A more fine-grained analysis of parental input was made by Van Veen and colleagues (2009). The authors studied the production of the German connectives *aber* 'but,' *damit* 'so that,' *und* 'and,' *weil* 'because,' and *wenn* 'when' in a longitudinal corpus data from a German-speaking child aged 1;11 to 2;11. The authors analysed the effects of age, as well as of short-term and long-term connective frequency in the input on the production of connectives by the child. By a short-term frequency, the authors defined the frequency of a certain connective in the child's input in the space of the same recording as the child's own production. A long-term frequency was calculated, in contrast, over all the recordings preceding the moment of the child's production.

Results showed that all three factors predicted the production of connectives in the child's speech. There were significantly more productions of connectives as the child grew older. However, the age growth curve varied a lot between different connectives and not all the connectives occurred in the child's speech at the same time. The order of emergence of connectives in speech followed the logic of the cognitive complexity of coherence relations (Sanders et al., 1992), as the positive connective *und* occurred before the negative *aber*, and the additive connectives *und* and *aber* were produced before the causal connectives *weil* and *damit*.

Once a connective occurred in the child's production, it underwent three periods, characterised by different degrees of influence from the parental input. During the first

---

[2] In this coding of age, the numeral before the semicolon stands for the number of full years and the numeral after the semicolon represents the number of full months.

period, the child's speech was almost not affected at all by the parental input. The second period, in contrast, was characterised by a very strong and important influence of the input. Finally, during the third period, the child's production again did not seem to be influenced much by parental input. The absence of a parental input effect in the first stage was attributed to the lack of cognitive ability to deal with more complex linguistic elements, such as discourse connectives. The absence of the effect at the final stage was explained by the fact that the child had acquired a sufficient mastery of a connective, whose usage became independent from parental input.

In a different study, Van Veen and colleagues (2013) examined not only the effect of the parental input, but also whether productions of connectives were influenced by "audience design" (Clark & Murphy, 1983) – in other words whether parents adapted the complexity of their speech to the age of their child. The authors analysed longitudinal corpora of spontaneous productions by five native German- and English-speaking children between the ages of 0;10 and 4;3. Not only did the children themselves produce many causal connectives, namely *because* and *weil,* when replying to *why*-questions, but they also received a lot of parental input on how to use these connectives in response to their own *why*-questions. In other words, the children used more causal connectives in speech following parental input, and also actively elicited the type of input they received, by asking causal questions themselves.

2.1.3. Domains of use

Among other sources of difficulties accounting for the order of production and comprehension of connectives, there is also the factor of the domain of use. As presented in Section 1.3.4, three main domains of use can be distinguished for coherence relations, namely content, speech act and epistemic domains (Sweetser, 1990). The distinction between the domains can be considered as a more refined extension of the CCR model, as these domains can be mapped onto the primitive of the source of coherence (Sanders et al., 1992). Thus, the content domain expresses an objective coherence relation, and speech act and epistemic domains are considered to be more subjective, as they actively involve a speaker in the construction of their meaning (Pander Maat & Degand, 2001; Pit, 2003; Stukker & Sanders, 2012).

Results from the experimental studies show that different domains of use are not acquired at the same time. Spooren and Sanders (2008), for instance, studied the production of causal relations by Dutch-speaking children aged 6 to 7 and 11 to 12 in a description and conversation tasks. In the description task, children had to describe a series of pictures, representing various causally related events, and thus this task was biased towards the production of content relations. In the conversation task, the participants had to express an opinion on a number of topical issues, which made it biased for speech-act and epistemic relations. The authors found that the younger children on average produced more content relations than the older ones. However, they did not observe any difference in the number of epistemic and speech-act relations between the two age groups. In a different series of elicited production tasks, targeting content, speech-act, and epistemic relations, Evers-Vermeul and Sanders (2011) showed that Dutch-speaking children could produce causal connectives in all three domains as early as at the age of 3. However, their longitudinal corpus study among children aged 1;6–5;6 demonstrated that children always produced connectives in the content and speech-act domains earlier than the ones in the epistemic domain, even though the age of the first production in each domain varied between children.

A similar order of acquisition between the domains was reported in the study by Zufferey (2010), examining the acquisition of the French causal connective *parce que* 'because' by four French-speaking children aged 1;10–4;3. This work showed that, in order to signal content and speech-act relations, this connective was used already starting from the age of 2;6; whereas to indicate epistemic relations, it was used only starting from the age of 3;0. These findings suggest that the epistemic domain probably represents a higher level of difficulty for children and therefore appears later in the children's speech. A possible explanation for this difficulty is that epistemic relations require metacognitive, or theory of mind abilities (Zufferey, 2010), meaning that a listener has to take another person's perspective to be able to interpret these relations; while content and speech-act domains do not involve such ability.

The comprehension study of Zufferey and colleagues (2015) also showed converging findings, supporting the claim about a greater level of difficulty of the epistemic domain. Native speakers of French and Dutch aged 5 to 8 participated in an off-line comprehension task, in which they had to read short stories, including objective (content domain) and subjective (epistemic domain) causal relations, and then to answer *why*-questions about the content of these stories. Results revealed that the comprehension of epistemic relations was significantly lower than that of content relations across all age groups in both languages, and that this difference in comprehension did not decrease with age. Since at the age of 8 the theory of mind ability had normally been already acquired, the authors argued that a greater difficulty of the epistemic relations may be related to the ability of abstract reasoning, as described by Spooren and Sanders (2008).

Finally, the fact that epistemic relations are more difficult is also supported by an on-line comprehension study of van Veen (2011). The author measured the comprehension of objective (content domain) and subjective (epistemic domain) relations by children aged 2;0 and 3;4 via an eye-tracking visual world experiment. The results demonstrated that the two age groups did not differ in their performance with content relations. However, the older children had a better and faster performance with the epistemic relations in comparison to the younger group of children. In other words, even if all the children managed to understand both types of relations and to fulfil the task, the epistemic domain seemed to cause a greater processing cost for the younger speakers.

2.1.4. Individual differences

Individual differences in children's linguistic competence are observed in all the linguistic domains, such as phonetics, vocabulary, and morpho-syntax, and stem from a number of factors, related to both social environment and cognitive skills (see Kidd & Donnelly, 2020 for a detailed overview of these factors). Among one of the most prominent predictors of linguistic competence in normally developing children is their chronological age. The studies that were presented before demonstrate that with age children tend to master the connectives encoding relations with a greater degree of cognitive complexity better (see, e.g., Spooren & Sanders, 2008), as well as those having lower frequency in the input (see, e.g., Van Veen et al., 2009), and belonging to the more complex epistemic domain (see, e.g., Evers-Vermeul & Sanders, 2011).

It is important to mention, however, that chronological age can be considered as an abstraction that reflects the cognitive maturation of speakers. Even if a number of studies report notable differences in the ability to use and to understand connectives related to age, variation in this ability is found even between children of the same age group. Thus, Blything, Davies and Cain (2015) and Blything and Cain (2016) found that not only children aged 3–7 performed better with the increase of age in the task examining temporal connectives, as mentioned in Section 2.1.1., but their performance was also predicted by their working memory (WM) capacity. Importantly, the factor of

WM was a much stronger predictor than the factor of age across all the children. Karlsson et al. (2019) made a more nuanced distinction between WM capacity and WM updating, where the former referred to the ability to store information in WM and the latter to the ability to update and track the information stored in WM. Their results also revealed that both types of WM were important predictors of comprehension and processing of sentences, including the Dutch temporal connectives *nadat* ‘after’ and *voordat* ‘before’, by Dutch-speaking children aged 9–12.

As regards exogenous factors, affecting children’s ability to use and comprehend connectives, Volodina and Weinert (2020) tested the ability of German-speaking primary school children to understand 39 connectives representing temporal, concessive, modal, causal, and conditional relations in a cloze-sentence task. They included a variety of predictors, namely receptive grammar, language background, and the socio-economic status of parents, as reflected by parental occupation. The authors found that parental socio-economic status (SES) together with receptive grammar were the strongest predictors for the initial level as well as for the development of the performance in the connective task, as revealed by a growth curve analysis. This finding is revealing of the multifaceted and complex nature of the SES measures. SES may for instance be related to the quantity and quality of the input that a child receives, as, at least in western cultures, a child from a lower-SES families is exposed on average to 3.5 times less words than a child from a higher-SES environment (Hart & Risley, 1995). Moreover, environments with a higher level of education often provide children with a more diverse and rich linguistic input (Huttenlocher et al., 2007) that is also linked to the development of the lexicon and grammar (Huttenlocher et al., 2010).

Finally, there is also evidence that general performance in a certain language at school is related to the mastery of connectives in late childhood. As reported in the study by Oğuz and Özge (2020), high-achieving students of Turkish aged 8–10 performed significantly better than low-achieving ones in a cloze task across causal, adversative, and temporal connectives in Turkish. In other words, it means that cumulative linguistic proficiency in a certain language, as reflected by a grade for this subject at school, is an important predictor of the mastery of connectives. However, it seems that also a reverse direction between the mastery of connectives and the performance in school subjects is true. Volodina et al. (2021) found, for instance, that German-speaking children aged 7–8 with a better comprehension level of connectives had not only better reading comprehension skills in German, but also a greater performance on measures of mathematical skills and overall higher grades in mathematics.

2.1.5. Conclusion

In this section, I have outlined the most prominent factors that account for the order of acquisition of discourse connectives as well as for the difficulties in their use and comprehension during childhood. In spite of a general consistency in the acquisition process, important inter-individual differences related to cognitive factors, social environment, and general academic level are observed at least from the age of three. In the next section, I present the main findings for adult speakers in order to understand how the competence with connectives evolves over the years and what an adult level of mastery with connectives looks like.

2.2. Mastering connectives in adulthood

2.2.1. Connectives and the comprehension of discourse

A vast body of research has explored the impact of connectives on the comprehension and processing of discourse by adults. There is, however, no consensus either from on-line or off-line experimental methods on whether the presence of connectives always enhances comprehension. An important number of studies have demonstrated that when connectives are present in discourse, they accelerate its processing (see, e.g., Bestgen & Vonk, 1995; Britton et al., 1982; Deaton & Gernsbacher, 2000; Gaddy et al., 2001; Haberlandt, 1982; Millis & Just, 1994; Sanders & Noordman, 2000). For instance, Haberlandt (1982) found that adult speakers of English read sentences without connectives more slowly than those that contained adversative connectives such as *but, yet, instead, however*, and *nevertheless* as well as consequence connectives such as *so, consequently*, and *therefore*. A similar result also was reported by Millis and Just (1994). The authors demonstrated that speakers were faster on a verb recognition task as well as faster and more accurate in answering comprehension questions when sentences were linked by the causal connective *because* or the concessive connective *although*.

Moreover, there is also evidence that connectives facilitate the off-line comprehension of written texts (Degand et al., 1999; Degand & Sanders, 2002; Loman & Mayer, 1983; Lorch & Lorch 1986; Meyer et al., 1980; Sanders & Noordman, 2000). Degand & Sanders (2002), for instance, studied the effect of causal and consequence connectives and signalling phrases on the comprehension of expository texts in French and Dutch, which functioned as L1 and L2. The authors found that the scores on comprehension questions were higher for the texts that contained causal signalling devices than for texts without such signals.

Another bulk of research suggests, however, that connectives might hinder (Millis et al., 1993) or not affect at all the comprehension of discourse (Meyer, 1975; Sanders & Noordman, 2000; Spyridakis & Standal, 1987). Sanders and Noordman (2000), for instance, did not report any difference in the performance with the recall task between the texts that included connective expressions of causal and additive relations and those without such markers. Millis et al. (1993) found that the recall of expository texts with temporal (*before, and then*), causal (*which caused, which enabled*), and intentional (*in order that, so that*) connectives and connective phrases was inferior to those without connectives. However, this study was later criticised on methodological grounds. For instance, Degand and Sanders (2002) suggested that connectives had a negative effect on the recall task because experimental texts were not well designed and the content of the segments was difficult to match with the coherence relation, signalled by connectives. The authors also argued that free recall tasks, in general, may be not very sensitive to the effect of connectives, as these tasks assess a more global level of comprehension, going beyond the comprehension of more local relations between the segments of a text.

Furthermore, it seems that not all types of coherence relations equally benefit from the presence of connectives. On the contrary, it might be easier to convey certain relations implicitly, i.e., without connectives. Zufferey and Gygax (2016), for instance, showed in a self-paced reading experiment that processing of an explicit confirmation relation, signalled by the French connective *en effet,* was significantly faster than processing of an implicit confirmation relation. When the authors compared the reading of implicit and explicit causal relations, they did not find such a notable facilitating effect from the explicit marking by the connective *en effet*, used in its causal function. This finding is particularly interesting, as it broadens the continuity hypothesis (Murray, 1995; 1997; Segal et al., 1991). The latter states that readers expect the next discourse segment to form a congruent meaning with the preceding segment in two ways, namely that it should be causally continuous and should present events in a temporally linear manner. Therefore, relations that are discontinuous from the point of view of causality and temporal order require a greater processing effort than continuous relations if the former are not marked by a connective. The study of Zufferey and Gygax (2016) extends this notion of continuity, showing that it can be disrupted not only on the grounds of temporal order and causality, but also because of the perspective shift. According to the authors, to understand the conditional relation, speakers should make a shift of perspective "from the reality space to the hypothetical space" (Zufferey & Gygax, 2016, p. 549), and this shift hinders the processing of the implicit conditional relation.

As suggested by Degand and Sanders (2002), the contradictory findings on the role of connectives in discourse may stem from the lack of control over factors such as type of coherence marker, type of text, type of task, and background knowledge of the reader. In addition, these differences can also derive from the type of coherence relation encoded by a connective, as not all relations profit in the same manner from the explicit marking by connectives. However, once these factors are taken into account, it seems that the presence of connectives does stimulate a better representation of a text not only during, but also after reading.

2.2.2. Factors accounting for the difficulty of connectives in adulthood

Besides the role of connectives in the comprehension of discourse, it is not less important to examine why certain connectives are more difficult to use and understand during adult years and how these difficulties differ from those that arise during childhood years. Different connectives are not processed and used in the same way by adult speakers. A greater cognitive complexity of certain coherence relations (Sanders et al., 1992) can be one of the important factors of the variation in the mastery of connectives also during adulthood. Thus, subjective coherence relations are processed more slowly that objective ones (Canestrelli et al., 2013; Traxler, Sanford, et al., 1997). In an eye-tracking reading task, for instance, Canestrelli et al. (2013) observed a processing delay for pairs of sentences including the subjective causal connective *want*, such as (7), in comparison to the sentences with the objective causal connective *omdat*, such as (8), in Dutch.

(7) Hanneke had haast, want ze was vier trappen afgerend om de post te halen.
'Hanneke was in a hurry, because she ran down four stairs to get the mail.'

(8) Hanneke was buiten adem, omdat ze vier trappen was afgerend om de post te halen.
'Hanneke was out of breath, because she four stairs was ran down to the mail get.'

There is also evidence that connectives encoding negative coherence relations involve an additional processing load than connectives signalling positive relations (see, e.g., Drenhaus et al., 2014; Koehne & Demberg, 2013; Morera et al., 2017; Xu et al., 2015). The visual-world-paradigm study and a reading experiment of Koehne and Demberg (2013) showed, for instance, that concessive, or negative causal (+, +), connectives in German were processed with a delay compared to positive causal (–, +) connectives. The subsequent ERP study, conducted by Drenhaus et al. (2014), supported these results. The authors revealed that, in contrast positive causal connectives, the processing of the concessive connectives in German and English also was associated with a higher late positive component (P600), reflecting the process of updating or reorganisation of the mental representation of the ongoing discourse (Brouwer et al., 2012). A similar result also was obtained by Xu et al. (2015) in Chinese. The data from self-paced reading and eye-movement experiments demonstrated that concessive relations were processed slower than causal ones, independently of the explicit marking of these relations by connectives.

A faster processing of positive causal relations may, however, be explained by the *causality-by-default* hypothesis (Sanders, 2005), proposing that speakers have a preference for establishing a causal relationship between sentences, unless there are signals suggesting otherwise. In consequence, such cognitive preference may account for a faster processing of this particular relation (see, e.g., Black & Bern, 1981; Haberlandt & Bingham, 1978; Keenan et al., 1984; Kuperberg et al., 2011; Mak & Sanders, 2013; Sanders & Noordman, 2000). Sanders and Noordman (2000) reported, for instance, that Dutch speakers read texts with causal problem–solution relations faster than those with additive list relations, and also had a better performance with the causal relations on verification and recall questions. Kuperberg et al. (2011), in turn, observed in an ERP study that causal relations were associated with a lower processing effort than non-causal ones, as suggested by a larger N400 amplitude.

The factor of the relation type and a potential difference in cognitive complexity appears less prominent in studies using off-line reading comprehension measures. Zufferey and Gygax (2020a) revealed that it was frequency in corpora that predicted the performance in a correctness-evaluation task for the French connectives *en effet* 'for/because', *en outre* 'moreover', *aussi* 'therefore', and *toutefois* 'however', mostly used in a written language; whereas the factors of cognitive complexity and polyfunctionality did not play any role on the performance with this task.

In addition, different functions of polyfunctional connectives can be associated with different levels of difficulty, depending on their frequencies (Asr & Demberg, 2020; Zufferey et al., 2015). Asr and Demberg (2020) showed in a series of experiments that the inference bias as well as the processing difficulty associated with different functions of the polyfunctional connectives *but* and *although* were consistent with their distributions in the Penn Discourse Treebank (PDTB; Prasad et al., 2008). In other words, the connective *but* most of the times generated the inference of a contrastive relation that this connective encodes more frequently in corpus data. Moreover, the processing of the more frequent contrastive function was also faster than that of the less frequent concessive one. In contrast, the connective *although* generated an equal number of inferences of the contrastive and concessive relations, and both relations were processed with a similar speed, which reflects the balanced distribution of both functions in corpus data.

2.2.3. Individual differences

Although adults have on average a high level of mastery of connectives, especially of those that are frequently used in oral speech (see, e.g., Cain & Nash, 2011), recent research reveals that there is an important inter-speaker variation in the performance with discourse connectives among this cohort. The study of Zufferey and Gygax (2020a), which was presented in the Section 2.2.2, revealed not only that the performance with less frequent written connectives was lower than that of the more frequent ones, but also that this performance varied according to the speaker's level of exposure to print, as measured by the author recognition test (ART; Stanovich & West, 1989). Moreover, the scores on the ART correlated with the grammatical competence of participants. These results suggest that the ability to distinguish between the correct and the incorrect uses of connectives is probably linked to a greater exposure to written language and to a greater grammatical awareness. The fact that exposure to written language is important for enriching the competence with connectives is in line with the findings on the specificities of the written modality. In fact, it is in writing that a greater variety of connectives with more precise functions are used (Biber, 2006; Castellà, 2004; Crible & Cuenca, 2017; Crible, 2020). It should therefore be expected that speakers that are more exposed to written language also have a higher proficiency with discourse connectives.

Besides the competence with discourse connectives, individual differences in the degree of exposure to print were also shown to predict speakers' sensitivity to alternative lexical signals, which are less prototypical and salient markers of coherence relations (Scholman et al., 2020). The work of Scholman and colleagues showed that speakers with higher scores on the ART were also more sensitive to the alternative signals for the relation of list. In other words, in a sentence continuation task, these speakers produced more list continuations after the sentences containing the expressions of quantity *a couple, a few, multiple*, and *several*, than after the sentences without these expressions. In contrast to the research on children (see Section 2.1.4), this study did not find an effect of verbal and nonverbal working memory. The authors argued that, probably, this effect was absent because they used an off-line test to measure participants' sensitivity to alternative signalling devices, whereas the WM effects may have been more apparent if they assessed this sensitivity via an on-line measure.

2.2.4. Conclusion

In this section, I reviewed the main findings on the use and comprehension of discourse connectives in adult years. It was outlined that the presence of connectives has a variable effect on the comprehension of texts, especially if such factors as type of text and type of coherence relation were not controlled for. Furthermore, it appears that different methods (on-line versus off-line) reveal that different factors predicting the degree of connectives' difficulty. Coherence relation type as well as the complexity of coherence relations signalled by connectives affect on-line processing, whereas frequency in corpus data is a strong predictor of the competence with connectives in studies measuring their use and comprehension off-line. Finally, the most prominent factor of inter-speaker variation in the mastery of connectives during adulthood seems to be the degree of exposure to print.

In the next section, I will explore how the competence with connectives develops during the transitional period between childhood and adult years, which seems to be less examined so far. I will therefore outline why it is essential to study this period more in-depth and will trace the directions of investigation that will be addressed in my thesis.

2.3. Mastery of connectives during teenage years

2.3.1. The importance of teenage years to study linguistic development

Teenage years are an important period of linguistic development between the emergence and mastery of language (Berman, 2004b). Language development in teenagers continues on lexical, semantic, syntactic, and pragmatic levels and has a slower speed and less salient character in comparison to early years of acquisition (see, e.g., Nippold, 2004, 2006, 2008; Berman, 2004b). The linguistic growth is greatly supported by the development of cognitive skills, metalinguistic competence, abstract thinking and social interactions, as well as by the schooling process. Through schooling, for instance, older children and teenagers start to be extensively exposed to a new type of linguistic input, that is written language, and are required to use more complex linguistic elements and engage in metalinguistic tasks (Nippold, 2004).

As a result, at the lexical level, teenagers learn a larger variety of words that rarely appear in oral speech (Ravid & Tolchinsky, 2002), such as terms related to specific academic subjects (e.g., *hypothesis*, *friction*). Moreover, they gradually improve their mastery of abstract nouns (e.g., *courage*, *kindness*), verbs related to metalinguistic tasks (e.g., *imply*, *predict*), and secondary meanings of polysemous words (e.g., *strike*, *sharp*) (Nippold, 2008). At the level of semantics, teenagers also increase their understanding of figurative expressions, such as metaphors, idioms, proverbs and slang. The learning of the new vocabulary and figurative expressions is accompanied and facilitated by growing metalinguistic skills. These include the ability to infer the meaning of a word or an expression from a larger context as well as the strategies of morphological (in case of new words) and lexical (in case of new expressions) analysis.

The syntactic development is characterised by the growth of sentence length, cohesion and low-frequency grammatical structures, such as subordinate clauses, passive voice, and non-finite verbal phrases (e.g., Loban, 1976; Nippold, 2006). However, sentence length and the complexity of syntax differs between discourse genres. Longer and more complex syntactic structures tend to appear more often in persuasive and expository discourse compared to narrative discourse (e.g., Crowhurst, 1980; Leadholm & Miller, 1992; Berman & Verhoeven, 2002).

The increasing number of social interactions also promotes the development of pragmatic skills, such as making appropriate comments and remaining on topic during conversation (Nippold, 2006). One of the most important pragmatic competences that is gradually acquired during teenage years is the ability to take the perspective of another person. Thus, teenagers improve their ability to tailor the style and the content of their discourse to the interlocuter and the context of the conversation, to make smooth transitions between topics and to present opposite sides of an argument (Nippold, 2008).

The mastery of connectives is at the interface between the lexical, syntactic and pragmatic skills, which are actively developing during teenage years. On the one hand, connectives are part of a more complex lexical knowledge and many of them are mostly bound to written contexts. On the other hand, the increasing production of longer phrases with subordinate clauses as well as texts with greater cohesion also relies a lot on the mastery of connectives (Nippold, 2008). Yet, even the most common connectives such as *before* and *because* are not completely mastered by the age of 12 (see, e.g., Flores d'Arcais, 1978; Cain & Nash, 2011; Irwin & Pulver, 1984) and some low-frequency connectives are challenging even for adults (Zufferey & Gygax, 2020a).

These findings suggest that the competence with connectives develops throughout the life span, with the major growth most probably occurring during teenage years (Nippold, 2008). This may be related to the fact that, during teenage years, the cognitive potential for a complex (hypothetical) thinking develops (Cook-Gumperz & Gumperz, 1992; Santrock, 1996). Furthermore, connectives are likely to appear in argumentative and expository texts, introducing complex reasoning and ideas, to which speakers start to be exposed mostly from middle school, when they become more autonomous readers (Nippold, 2004). Taylor et al. (2019) showed indeed that the use of connectives is linked to a more complex reasoning ability. The authors analysed the use of argumentative strategies and four types of connectives, namely temporal, causal, additive, and adversative, in the argumentative essays of middle-school students from grades 6–8 (approximately aged 11–14). The results revealed that connectives, especially the adversative ones, were used to express the most complex types of arguments.

Besides their central role in the linguistic system and in the ability to express complex reasoning, connectives also play a significant part in the academic success. Discourse connectives improve overall writing quality in argumentative essays (Andreev & Uccelli, 2023; Duggleby et al., 2016; Uccelli et al., 2013) and are considered to be an integral part of fundamental academic language (see, e.g., RAND Reading Study Group & Snow, 2002). Academic language is mostly used in written formal educational contexts and has a certain number of characteristic linguistic features (Barr et al, 2019; Snow & Uccelli, 2009). For instance, it employs a more diverse and precise vocabulary, has a highly organised structure, and a complex syntax involving the use of discourse connectives (Snow & Uccelli, 2009). As its name suggests, this type of language is used not only in language and literature classes, but in all other subjects, as a means to transfer academic knowledge. Therefore, the success in these subjects depends on the ability to use this language with all its typical features as well as to read and understand texts written in it. Indeed, ample evidence suggests that performance in various academic domains such as mathematics (see, e.g., Fuentes, 1998; Jordan et al, 2006; Salihu, Aro & Räsänen, 2018), history (see, e.g., Beek, 2020), and natural sciences (see, e.g., Akbaşlı et al., 2016; Imam et al., 2014; Korpershoek, 2015; O'Reilly & McNamara, 2007), is linked to reading comprehension skills.

2.3.2. Factors accounting for the difficulty of connectives in teenage years

By the age of 11–12 years old, teenagers are able to use and understand frequent connectives, signalling main types of coherence relations such as addition, causality, contrast, and temporality (see, e.g., Blything et al., 2015; Blything & Cain, 2016; McClure & Steffensen, 1985; Spooren & Sanders, 2008). However, when compared to the performance of adults, teenagers are reported to have lower scores on connective tasks even during late teenage years (Nippold et al., 1992; Zufferey & Gygax, 2020b).

Zufferey and Gygax (2020b) assessed, for instance, the performance with four French connectives, namely *en outre* 'moreover', *aussi* 'therefore', *en effet* 'for/because', and *toutefois* 'however', in a cloze-sentence task by pre-university and pre-vocational students (M_{age}=16) and a control group of university students (M_{age}=22). The authors chose these connectives because they allowed to assess the effect of cognitive complexity, frequency, and polyfunctionality on the performance in the cloze task. The selected connectives are mostly used in the written language and signal four coherence relations, varying in cognitive complexity from the less complex relation of addition to the most complex relation of concession. Moreover, the chosen connectives differ in frequency and the number of functions that they signal. The connectives *en outre* and *aussi* are less frequent in written corpora than the connectives *en effet* and *toutefois*; and the connectives *aussi* and *en effet* are polyfunctional, while *en outre* and *toutefois* are monofunctional. Results revealed that teenagers generally had significantly lower accuracy scores than the control group of adults across all connectives. Furthermore, the results differed also between the two groups of teenagers. Pre-university students outperformed pre-vocational students for the more frequent connectives *en effet* and *toutefois*. In contrast, their scores did not differ much for the less frequent connectives *aussi* and *en outre*. In line with the findings by Zufferey and Gygax (2020a) on adults, it was the factor of frequency that best predicted accuracy scores in late teenage years, while cognitive complexity and polyfunctionality did not seem to influence the outcome of the experiment.

The idea that the frequency of connectives in corpora affects the use and comprehension of connectives in teenage years had already been suggested earlier by Nippold et al. (1992). The authors studied the performance with connectives signalling consequence (*therefore*, *consequently*), additive (*similarly*, *moreover*), contrastive (*rather*, *contrastively*), and concessive (*nevertheless*, *however*) relations in sentence continuation and sentence cloze tasks by teenagers and young adults aged 12–23. Results demonstrated that, first of all, there was a gradual increase in performance from the youngest to the oldest participants in both tasks. Next, the continuation task overall was more challenging than the cloze task for all the participants. Lastly, the authors did not find important differences between relation types. To be more precise, they did not observe a discrepancy between concordant (*consequently, moreover, similarly, furthermore,* and *therefore*) and discordant (*nevertheless, however, contrastively, conversely,* and *rather*) connectives, which correspond to positive or negative connectives according to the CCR framework (Sanders et al., 1992). In a post hoc discussion of the results, Nippold et al. (1992) suggested that the performance with connectives was probably related to their frequency in school reading materials, reported in Carroll et al. (1971), as the connectives with the lowest frequency scores tended to also have the lowest scores in both experimental tasks.

In a similar vein, Crosson and Lesaux (2013) proposed that it was the degree of familiarity with connectives that could be a predictor of their appropriate use already in early teenage years. To verify this hypothesis, the authors assessed the use of English connectives signalling additive, adversative, causal, and temporal relations, and having different levels of familiarity, in a cloze-sentence task by fifth-graders (approximately aged 10–11), having English as L1 and as L2. Different degrees of familiarity were established based on the number of teenagers from a certain age group who actually knew these connectives. Thus, for example, among the temporal connectives *after*, *meanwhile*, *eventually*, and *subsequently*, *after* was categorised as *easy*, or the most familiar, and *subsequently* as *very difficult*, or the less familiar. The results revealed that, indeed, participants systematically gave more accurate responses for the connectives that were more familiar, regardless of the coherence relation type.

In a text comprehension study by Kleijn et al. (2019), however, the type of coherence relation encoded by connectives did affect the level of comprehension. The authors found that the presence of connectives contributed to a better comprehension of more difficult expository texts on both global and local levels by Dutch-speaking teenagers aged 13–16. Moreover, the level of comprehension was not the same for different coherence relations. For instance, additive connectives such as *and* decreased the level of comprehension; whereas contrastive and causal connectives such as *but* and *as a result* had a facilitating effect on comprehension. The authors suggested that more complex contrastive and causal relations are also probably more informative, as by rendering a coherence relation salient, they help readers make the necessary interpretation. In contrast, additive connectives may have put an excessive focus on the expected relation, triggering a search for a more elaborate interpretation. Finally, the results of this study also showed that there was no important individual variation between participants, related to their academic level and reading proficiency.

The role of connectives and inter-speaker differences were also examined by Van Silfhout et al. (2015) in text processing and comprehension tasks. First of all, the authors evaluated with an eye-tracking reading task whether the presence of connectives facilitated the processing of narrative and expository texts. In addition, to measure the off-line comprehension of these texts, the authors also included inference bridging questions, targeting local comprehension within a paragraph, and a sorting task, targeting the global level of comprehension. In total, 141 Dutch-speaking teenagers aged 12 to 15, coming from pre-university and prevocational classes and differing in a reading proficiency level, participated in the study. The findings revealed that the presence of connectives accelerated processing and enhanced local comprehension of texts, irrespective of their genre. However, it seems that connectives did not affect a global comprehension, as teenagers did not benefit from them, when doing a sorting task. Moreover, all the participants, independently of their academic background and reading proficiency, equally benefitted from the presence of connectives. Nevertheless, in general, the reading proficiency score was reported to be a more reliable measure of reading competence. A higher reading proficiency level overall predicted shorter rereading times and better scores on both comprehension tasks.

2.4. The motivation for the present thesis

2.4.1. Examining the period of teenage years

The ability to use and comprehend connectives in discourse is central for the development of a full-fledged linguistic competence, allowing not only to successfully communicate, but also to access knowledge from written texts. It seems, however, that the majority of research on this topic has mostly focused on childhood and adulthood, leaving out the period of teenage years (see for the few exceptions Section 2.3.1). That is the reason why in my thesis I examine how the use and comprehension of connectives develops during this transitional period between the acquisition of basic skills and a higher adult-level proficiency.

> H1: The mastery of connectives is not completed by the end of childhood (see, e.g., Geva, 2006; Knoepke et al., 2017; McClure & Steffensen, 1985; Pyykkönen & Järvikivi, 2012) and competence keeps on developing over the period of teenage years and beyond it, as the teenage period is a transitional stage between the acquisition of basic linguistic skills and the development of an advanced language proficiency (Berman, 2004a, 2004b).

2.4.2. Exploring factors accounting for the difficulties with connectives

An overview of the literature revealed that not all factors equally affect the performance with discourse connectives in children and adults. While during childhood years, the cognitive complexity of coherence relations is a prominent predictor of their correct use, processing and comprehension; during adulthood, this complexity accounts mostly for the differences in on processing between different connectives (see Section 2.2.2). Furthermore, frequency of connectives in a child's input and frequency of connectives in corpora predict the performance with connectives, respectively, in childhood and adulthood. However, during childhood years, it seems that connectives' frequency starts to affect the competence with connectives only after the complexity level corresponding to a certain relation has been acquired (Van Veen et al., 2009). It is therefore important to unveil how the effects of cognitive complexity of coherence relations and frequency are transformed over the years between childhood and adulthood in order to better understand the developmental path during teenage years.

H2: Considering the evidence from previous research (see, e.g., Evers-Vermeul & Sanders, 2009; McClure & Steffensen, 1985; Spooren & Sanders, 2008 for the childhood period; Zufferey & Gygax, 2020b for teenage period, and Zufferey & Gygax, 2020a for the adulthood period), the cognitive complexity of coherence relations should not be a prominent factor affecting the ability to use connectives in off-line tasks.

H3: In line with the findings of Van Veen et al. (2009) and Zufferey & Gygax (2020a, 2020b), corpora frequency should in contrast be a more important factor explaining differences of mastery between connectives during teenage years.

Similar to the effects of frequency (H3) and based on the evidence on general linguistic development during teenage period (Berman, 2004), the mode, in which connectives are typically used (oral versus written), may be another predictor of an accurate use of connectives.

H4: Since a speaker starts to be exposed to written language later than to oral speech and predominantly at school (Nippold, 2004, 2008), teenagers should demonstrate less accurate usage of written connectives compared to oral ones.

The last predictor that is important to consider is that of polyfunctionality. It seems that the only evidence on the teenagers' competence with polyfunctional connectives is provided by Zufferey and Gygax (2020b). However, the authors examined only two polyfunctional connectives, and a limited cohort of older teenagers (18 high-school students with $M_{Age} = 16.2$ and 22 professional school students with $M_{Age} = 15.6$) participated in the study. Therefore, further examination of the role of polyfunctionality is necessary to build a more solid knowledge on the development of the mastery with connectives during teenage years.

H5: Based on the findings of Zufferey and Gygax (2020a, 2020b), different levels of mastery for different functions of polyfunctional connectives may disappear with time and should not predict the teenagers' use of connectives in off-line production tasks.

2.4.3. Evaluating the interaction between connectives and alternative signals of coherence relations

In addition to the use and comprehension of connectives themselves, I also study in this thesis how they interact in discourse with other, alternative signals of coherence relations. Considering the less salient nature of alternative signals, it is important to explore whether the sensitivity to such signals is affected by the presence of more salient and prototypical signalling devices such as connectives. Moreover, it is also important to determine whether younger speakers are sensitive at all to these alternative signals of coherence relations. The answer to this question will shed light on the acquisition of coherence more generally, and on how it develops. For instance, whether teenagers already rely on more fine-grained cues such as alternative signals or whether they are mostly guided by connectives in their inference of coherence relations.

H6: Teenagers may already be sensitive to alternative signals of coherence. However, the presence of connectives combined with alternative signals might affect the inference generation in teenagers more than in adults.

2.4.4. Identifying predictors of individual variation

Interindividual variation in the ability to use and understand connectives in discourse is apparent throughout the whole lifespan of speakers, from childhood to adult years. Hence, learning about the sources of this variation at different developmental stages may inform us about the underlying mechanisms of functioning and the linguistic nature of discourse connectives. For instance, identifying the link between the competence with connectives and grammatical awareness and exposure to print would provide evidence for the theories, suggesting that connectives are part of procedural knowledge that is unconscious and is accessed through a frequent exposure to the contexts where connectives are used. In contrast, the relation between this competence and lexicon size would plead for the view that connectives are not deprived of conceptual meaning. Based on previous research, the following hypotheses about the inter-individual variation in the mastery of connectives can be formulated:

H7: Considering that connectives are used in a greater variety and with more precise functions in writing than in oral speech (see, e.g., Biber, 2006; Crible & Cuenca, 2017), speakers with a greater degree of exposure to print should also have a more advanced overall mastery of connectives.

H8: Given that connectives might encode both conceptual and procedural meanings (see, e.g., Ducrot, 1972; Wilson, 2011) and that previous research on adult native and non-native speakers of French (Wetzel et al., 2020) confirmed the importance of vocabulary knowledge for the performance with connectives in a cloze test, a broader lexicon should predict a better use of connectives in native French-speaking teenagers.

H9: Since there is evidence from the research on adults (Zufferey & Gygax, 2020a) and primary school children (Volodina & Weinert, 2020) that the use and comprehension of connectives are predicted by grammatical competence, teenagers with a greater knowledge of grammar should be more accurate in using connectives.

H10: Academic background is an indicator of a cumulative level of linguistic proficiency that may reflect a complex set of skills, such as explicit awareness of discourse structure and a more diverse vocabulary (Snow & Uccelli, 2009; Welie et al., 2017). Therefore, in line with previous studies (Oğuz & Özge, 2020; Zufferey & Gygax, 2020b), teenagers with a more advanced academic background may also be more proficient when using connectives.

Moreover, I also examine the interaction between cognitive maturation (see, e.g., Mills et al., 2016) related to chronological age, and measures of various linguistic competences, as age is one of the primary factors affecting connective, and more generally, language acquisition in childhood, whereas it is a much less central factor for linguistic processes occurring in adulthood. Uncovering that the competence with connectives is better predicted by one type of factor than by another would shed light on the particularities of teenage years as a transitional phase between childhood and adult years.

H11: Based on the evidence from the studies on childhood (see, e.g., Evers-Vermeul & Sanders, 2011; Spooren & Sanders, 2008; Van Veen et al., 2009) and the lack of such evidence from the studies on adulthood (see Section 2.2.2), the role of chronological age should be less important during teenage years, especially compared to other measures of linguistic proficiency.

2.4.5. Studying the role of instruction as a way to improve competence with connectives

Another central issue is whether and how the competence with connectives can be improved via training at school, since connectives are an integral part of academic language skills (see, e.g., Snow & Uccelli, 2009). Evaluating the role of instruction for a more efficient development of the ability to use and understand connectives would allow to rein the difficulties related to the linguistic characteristics of connectives, and to potentially overcome important inter-speaker variations. The question of instruction is especially crucial owing to the double-edged nature of connectives between concepts and procedures (see Section 1.2), as procedural meaning is particularly challenging to explain and seize (Wilson & Sperber, 1993).

H12: Providing readers with an explicit knowledge of the connectives' functions in the form of learning activities may help them to build the link between their form and function and, eventually, with practice, to internalise their functioning in discourse (Ellis, 1994). In consequence, training activities about the functioning of connectives should enable speakers to apply the rule in the context of an off-line cloze test, and as a result, have a higher accuracy score.

2.4.6. Assessing the role of context for the competence with connectives

Previous research, examining how teenagers use connectives, was based on the evidence from sentence-cloze tasks (e.g., Nippold et al., 1992; Zufferey & Gygax, 2020b). In real-life discourse, however, connectives do not usually link isolated pairs of sentences. Sentences exist in a broader discourse, and the interpretation of each new sentence is established from the context of preceding sentences (see, e.g., Kamp, 1981). Therefore, assessing the use of connectives in a more ecological context of broader texts will show whether the results of previous studies on the same phenomenon were partially biased by the non-realistic nature of experimental context. Based on the contrastive results from the research on the role of context for the interpretation of a coherence relation by adults (Scholman & Demberg, 2017; Yung et al., 2019), the following hypothesis can be formulated:

> H13: If a greater context facilitates the insertion of an appropriate connective, it would mean that the level of connective mastery was underestimated by previous studies, and the accuracy scores on the text-cloze test should be higher than those on the sentence-cloze test. In contrast, if the context does not affect the performance on the insertion task, it would imply that the reasons underlying the mastery of connectives extend beyond the context in which they are employed.

2.5. Conclusion

In this chapter, I outlined how the ability to comprehend and use connectives develops from early childhood to adulthood. I also provided rationale on why it is essential to study how the competence with connectives evolves over teenage years and which factors are mostly involved in the development of this competence. In the next chapters, I present experimental evidence, allowing to better understand how teenagers improve their mastery of connectives and which exogenous and endogenous factors predict the most the accuracy of the connectives' usage and comprehension. Chapter 3 starts with the presentation of a series of experiments, assessing the role of cognitive complexity and frequency as well as inter-individual differences in academic background, exposure to print, and grammatical competence for the use of four French connectives in a more or less ecological experimental context.

3. The Mastery of Connectives from the Written Mode across Different Experimental Contexts[1]

I examine in this chapter several research directions that were defined in the previous chapter. As mentioned in Section 2.3.1, teenage years are an important period for building a higher level of linguistic proficiency that involves the use of connectives typically found in written, formal academic contexts. Despite the significance of teenage years for the development of the linguistic proficiency in general and for the mastery of connectives in particular, only few studies have investigated how teenagers use and understand discourse connectives (see Crosson & Lesaux, 2013; Kleijn et al., 2019; Nippold et al., 1992; Van Silfhout et al., 2015; Zufferey & Gygax, 2020b for exceptions). However, there is no ample evidence about how the competence with connectives develops between childhood and adulthood in a continuous manner (see Nippold et al.,

[1] This chapter is partially published in Tskhovrebova, Zufferey and Gygax (2022).

1992 for the only example of such a study). Therefore, I aim to trace in the present chapter how this competence develops over a longer time period between early and late teenage years (from 12 to 18 years old) and compare it to the performance of adult speakers (over 18 years old). I expect that, with age, speakers should on average display a better use of connectives, being older may be associated with a potentially longer time of exposure to these linguistic elements, as well as to greater cognitive maturity.

The next important dimension is to explore how the development of the ability to use and understand discourse connectives varies between different types of connectives and speakers. In other words, how the linguistic characteristics of connectives themselves as well as various external factors, such as general linguistic proficiency, affect the development of their mastery. Studies focusing on childhood years show that the cognitive complexity of coherence relations is one of the main factors predicting the order of acquisition of connectives both from the perspective of their production and comprehension, as revealed by on-line processing and off-line comprehension experimental measures (see Section 2.1.1). In contrast, it seems that for older speakers, it is the frequency of connectives rather than the cognitive complexity of the relations they encode that predicts variation in their mastery, especially when the latter is measured via off-line production and comprehension methods (Nippold et al., 1992; Zufferey & Gygax, 2020a, 2020b). However, it is not clear from what point onwards frequency becomes a more prominent predictor of connective mastery than the type of coherence relation that they signal. To address this issue, I replicate the connective cloze task from Zufferey and Gygax (2020b), assessing the use of four connectives typical from the written mode, namely *aussi* 'therefore', *en outre* 'moreover', *en effet* 'for', and *toutefois* 'however', and expand it to a broader group of speakers, including teenagers not only from one high school level, but also from secondary school as well as adults from a wider age range.

The fact that all four connectives are typically used in written language was important for testing the effect of frequency, by disentangling it from the effect of mode. As a matter of fact, connectives mostly used in speech have on average higher frequency than those mostly used in writing. For instance, the four connectives chosen for this study are easily comparable in terms of their frequencies in written corpora, with two of them having lower frequency (107 and 73 occurrences per million words, respectively for *aussi* 'therefore' and *en outre* 'moreover') and other two being more frequent (211 and 185 occurrences per million words, respectively for *en effet* 'for' and *toutefois* 'however') (Zufferey & Gygax, 2020a). Including connectives typically used in speech would render such comparison quite difficult, as the average frequency of these connectives is significantly higher. Compare, for example, the frequencies of the four chosen connectives with the frequency of the oral connective *mais* 'but' (3'924 occurrences per million words) (Zufferey & Gygax, 2020a).

It is important to mention that the connectives *aussi* and *en effet* are polyfunctional. The connective *aussi* can also signal a relation of addition (equivalent to the English *also*) and the connective *en effet* a relation of confirmation (equivalent to the English *indeed*). Considering that in this study, the production of connectives is assessed in sentence-initial position only, the two polyfunctional connectives are tested in their dominant function for this position. As regards the connective *en effet*, its dominant sentence-initial function is causality, since it accounts for 80% of uses in written corpora (Zufferey & Gygax, 2020a, p. 4). The only possible relation that can be signalled by the connective *aussi* in a sentence-initial position is that of consequence (Roze et al., 2012, p. 11).

Finally, the connectives chosen for this study also signal four types of coherence relations with different degrees of cognitive complexity, as theorised in the CCR model of Sanders et al. (1992) (see Section 1.3.1 for a more detailed description of the model). The additive relation, encoded by the connective *en outre*, is the one with the lowest degree of complexity. This relation is followed by the relation of consequence (*aussi*), of causality (*en effet*), and finally of concession (*toutefois*). To summarise, the connectives *en outre* and *aussi* signal simpler relations in terms of cognitive complexity and are less frequent in written corpora. In contrast to *en effet* and *toutefois* that encode more complex relations while being more frequent in written corpora. Hence, these four connectives are well fitted to assess the effects of cognitive complexity and frequency on teenagers' developing competence.

As for the external predictors of the competence with written connectives, I argue that this competence may be related to a general proficiency in written language. This modality requires different sets of skills from the spoken one, including connective uses. For instance, connectives appear in a greater variety in written language and take more precise functions (Biber, 2006; Castellà, 2004; Crible & Cuenca, 2017). Therefore, I hypothesize that speakers who are more often exposed to print and have a better awareness of written grammar should also have a better knowledge of connectives from the written mode.

To assess speakers' degree of exposure to print, I use a French version of the Author Recognition Test (ART-F) (Zufferey & Gygax, 2020a) and a newly developed French version of the Title Recognition Test (TRT-F). This family of measures used to assess the level of exposure to print are particularly interesting for this study, as they were found to predict various written linguistic competences, such as orthographic processing (Stanovich & West, 1989) and general reading ability (Spear-Swerling et al., 2010; Mano & Guerin, 2018). Moreover, I included two versions of this test in order to increase the chances of capturing variation in the degree of exposure to print within speakers from different age groups. Zufferey and Gygax (2020a) demonstrated, for instance, that the ability of adult French-speakers to distinguish between correct and incorrect uses of connectives bound to the written mode was predicted by their performance on the ART-F. However, there is evidence that younger speakers may be more receptive to a recognition test based on book titles rather than on author names (Cunningham & Stanovich, 1990; Echols et al., 1996). Hence, the TRT-F, adapted for the Swiss French-speaking teenagers, was included as a second measure of exposure to print. Based on these previous results, I expect the TRT-F to be a more accurate predictor of performance in the sentence cloze test for younger participants aged 12-18 and the ART-F to better predict this performance within older speakers who have finished high school and are over 18 years old.

To measure participants' grammatical awareness, I use a test of written grammatical competence, developed by Zufferey and Gygax (2020a). The authors showed that adults' performance in the connective task was related not only to the ART-F, but also to the ability to identify correct and incorrect written grammatical forms in French. Furthermore, a study by Volodina and Weinert (2020) also suggests that comprehension of connectives is predicted by receptive grammatical competence already in primary school children. I therefore expect that a better mastery of written connectives may be predicted by their level of written grammatical competence also in the case of teenagers.

As an indicator of a cumulative level of linguistic proficiency, I also take into consideration participants' academic background in French, as this indicator may reveal a different set of language abilities. For instance, a more proficient user of academic language makes more use of complex grammar, has a better explicit awareness of discourse structure, and a lexicon that is more diverse and precise (Snow & Uccelli, 2009). In addition, Welie et al. (2017) demonstrated that connective processing and use are facilitated by a greater metacognitive awareness of text structure, reading, and writing strategies. In consequence, and similarly to the findings of Oğuz and Özge (2020) on children and those of Zufferey and Gygax (2020b) on late teenage years, I suggest that the higher the academic level of a speaker, the more advanced must be their academic language proficiency, linguistic metacognitive awareness, and, therefore, their competence with connectives.

Besides the exogenous and endogenous factors discussed above, the experimental context itself may affect the use of connectives. In real-life written communication, connectives rarely link isolated pairs of sentences, as it was for instance the case in the studies of Nippold et al. (1992) and Zufferey and Gygax (2020b). Sentences exist in the context of a larger discourse and each new sentence is interpreted based on the context provided by preceding sentences (Heim, 1982; Kamp, 1981). Previous studies on the role of context for the interpretation of a coherence relation report contrastive findings for different coherence relations. Scholman and Demberg (2017) showed, for instance, that a greater context facilitated the insertion of an appropriate connective, especially in the case of additive relations. In contrast, the study by Yung et al. (2019) revealed that the context almost did not impact the performance on the insertion task for most of the relations and was beneficial only for causal, consequence, and level-of-detail (as labelled in PDTB 3.0, Prasad et al., 2019) relations. Comparing the teenagers' performance on sentence- and text-cloze tests would complement this line of research and verify whether the results of previous studies on the teenagers' competence with connectives, based on sentence-cloze tasks (Nippold et al., 1992; Zufferey & Gygax, 2020b), were partially biased by the non-realistic experimental environment and may have reported an artificially low level of performance with connectives, especially those signalling additive, causal, and consequence relations. Alternatively, if the results of the two types of cloze tests are similar, it would indicate that the factors that account for the difficulty or ease with connectives are more profound than the nature of the context in which they are employed.

3.1. Experiment 1. Sentence-level cloze test (Sentence-CT)

The materials from this study are available on the OSF repository[2].

3.1.1. Participants

A group of 191 teenagers aged 12 to 22 years old (M_{age}=15.87; SD=1.48) participated in Experiment 1. All the participants were native French speakers according to their French teachers. The experiment was carried out among teenagers of two school levels: secondary school (M_{age}=14.48; SD=0.78) and high school (M_{age}=16.80; SD=1.4). Each level included different types of classes. The secondary school group contained classes of Level 1 and Level 2, where the former is followed by students with a lower performance in French, while the latter is for the students with higher grades. According to the Swiss education system, the separation between two levels is made on the basis of the annual mean grade that a student gets in the discipline. The high school group was

[2] https://osf.io/gqpy7/?view_only=1f466fe2e9cc42aba5b53b721d9475ba

made of 3 types of classes that I will call classes of Levels A, B, and C. Students from Level A are those who chose a specialisation in business and who can directly go to work in the commercial sector after their studies or, if they want, can continue their studies in a professional school of management. Participants from Level B obtain the *General Culture Certificate* at the end of their studies which gives them access to professional schools in the sectors of health, social work, sports etc. Finally, studies in Level C allow students to enter university or polytechnical schools. All the mentioned levels have different entry requirements increasing from Level A to Level C. The youngest participants are from Levels 1 and 2. The eldest are the pupils from Level A and the ones in the middle are the participants from Levels B and C (see Table 1 for age distribution among all the mentioned groups).

Two groups of adults also performed the test. First, native French-speaking adults recruited on the Prolific© platform (Prolific, Oxford, UK). This group was included to assess the extent of individual variations among adults. Second, a group of native-speaking university students studying French. This second group was included to measure the highest possible degree of proficiency with these connectives.

Table 1. Distribution of participants across all groups in Experiment 1

Group	Group specifications	N	Mean	SD	Range
All Teenagers		191	15.70	1.8	12 – 22
Secondary School		68	14.48	0.78	12 – 16
Level 1	Less advanced curriculum in French	31	14.46	0.88	13 – 16
Level 2	More advanced curriculum in French	37	14.49	0.7	12 – 16
High School		123	16.80	1.4	15 –22
Level A	Curriculum preparing for a direct professional path	37	17.70	1.18	16 – 22
Level B	Curriculum preparing for a professional High School	22	15.91	0.81	15 – 17
Level C	Curriculum preparing for a university path	64	16.16	0.95	15 – 19
All Adults		116	25.43	8.3	18 – 64
University Students	Specialisation in French Studies	63	22.16	3.93	18 – 38
Other Adults	Heterogeneous backgrounds	53	29.26	10.25	18 – 64
Total		307	19.51	7.02	12 – 64

3.1.2. Materials and procedures

3.1.2.1. Sentence-level cloze test

The material for this task was the same as in Zufferey and Gygax (2020b) and included 40 items. Each item represented a pair of two sentences separated with a blank. The participants had to fill it in with an appropriate connective making a choice between *aussi* 'therefore', *en effet* 'for', *en outre* 'moreover', and *toutefois* 'however'. Ten items were included for each coherence relation, namely consequence, cause, addition, and concession. All the items were conceived in such a way that there was only one possible correct interpretation in this particular discourse context. All the sequences of sentences were checked for ambiguity by the authors. An example of item for each type of relation is illustrated in (3) to (6). The full materials for this task can be accessed at https://osf.io/yxj8q/.

(3) Consequence (correct answer *aussi* ‘therefore’)

Pascal n’avait pas pris ses clés. ________ il dû attendre son collègue pour rentrer dans son bureau.

‘Pascal didn’t take his keys. _______ he had to wait for his colleague to enter the office.’

(4) Cause (correct answer *en effet* ‘for’)

Marc ne fut pas content de ses résultats. ________ il n’avait eu que des mauvaises notes aux examens.

‘Marc wasn’t happy about his results. ________ he received only bad marks for the exams.’

(5) Addition (correct answer *en outre* ‘moreover’)

Georges avait une grande culture musicale. ________ il était passionné de photographie.

‘Georges had a great knowledge of music. _______ he was passionate about photography.’

(6) Concession (correct answer *toutefois* ‘however’)

Le maître n’avait pas terminé de présenter son cours. ________ il libéra ses élèves à l’heure prévue.

‘The teacher didn’t finish his class. _______ he let the students go at the scheduled hour.’

Part of the experiment was carried out in class with pen and paper (n=131) before the Covid-19 pandemic. Due to health security measures, the participants who took the test after March 2019 had to perform it online via a weblink.

3.1.2.2. Author recognition test for French-speaking adults (ART-F)

The material for this test was also developed by Zufferey and Gygax (2020a), and was designed to represent a French equivalent of the English ART (Stanovich & West, 1989). The task contained a list of 40 real names of authors and 40 fake ones (the full test is available at: https://osf.io/yxj8q/). The participants had to tick a box for all the names that they recognised as authors. They were told that some names were fake and that they should select only the ones that they knew, as one point would be removed per incorrect answer. Thus, for each correct author, participants obtained 1 point and for each false one -1. The general score was calculated with the formula *correct items score + incorrect items score*. The maximum score is 40 and the minimum score is -40. In addition to the task itself, participants performed a subjective evaluation of their degree of exposure to print (i.e., *How regularly do you read?*), on a 11-point scale ranging from 0=*never* to 10=*every day*.

3.1.2.3. Title recognition test for French-speaking teenagers (TRT-F)

The TRT-F is another measure of exposure to print, similar to the ART, that aims to assess the level of exposure to print of French-speaking teenagers. This version was developed for this study because previous research showed that, to capture the variation in exposure to print in an efficient manner, the ART should be adapted to the language and cultural context (see, e.g., Stainthorp, 1997) as well as to the age of participants (see, e.g., Cunningham & Stanovich, 1990). Moreover, there is evidence that a test based on children's book titles instead of author names is better suited to assess print exposure of younger cohorts (Echols et al., 1996).

The development of the TRT-F followed the procedure from Cunningham and Stanovich (1990). The TRT-F included titles of the books that were not part of school curriculum and that were popular among French-speaking Swiss teenagers aged 12–18, according to three libraries for children and one major national chain of bookstores in Switzerland. The test consisted of 22 real titles and 22 false ones. The false titles were created by changing an original title of a book. For instance, the original title "La maison Ipatiev" 'Ipatiev house' was transformed into "La maison Tourgueniev" 'Turgenev house'. In order to assure greater reliability and grasp of this measure, I designed two versions of the TRT-F. The correct items from the Version 1 ("Le Quidditch à travers les âges") were transformed into the incorrect ones in Version 2 ("Le Quidditch à travers les états") and vice versa. Both versions were randomly distributed among an equal number of teenagers. The attribution of scores followed the same procedure as in the ART-F. The maximum possible score was 22 and the minimum -22.

3.1.2.4. Written grammatical competence

As in Zufferey and Gygax (2020a), the materials included 40 sentences, 20 of which were grammatically correct and 20 were incorrect. The participants had to judge the correctness of the sentences on a 11-point scale ranging from 0=*I am sure that it is incorrect* to 10= *I am sure that it is correct*. The incorrect sentences contained an error typical of the written mode, such as silent agreements (7) or misuse of diacritical marks (8). The full materials can be accessed at https://osf.io/yxj8q/.

(7) Après avoir <u>manger/mangé</u>, ils sont partis.

'After they had eat/eaten, they left.'

(8) <u>Vôtre/votre</u> fille est adorable.

'Yours/your daughter is adorable.'

To calculate a general grammatical competence score per participant, I subtracted the mean of all the incorrect answers on the incorrect sentences from the mean of all the correct answers on the correct sentences. The maximum score is 20 and the minimum score is -20.

3.1.3. Results

3.1.3.1. Connectives in the Sentence-level cloze test

I analysed participants' responses in terms of correctness (i.e., right or wrong) by fitting a mixed-effects logistic regression model on the binary variable. The analysis was conducted using the R software (R Core Team, 2020), and models were tested with the *glmer* function of the *lme4* package (Bates et al., 2015). Models were compared with the *anova()* function which calculates the Chi-square value of the log-likelihood in order to evaluate the difference between models, following Baayen's (2008) procedure. Similar to the previous studies on the same issue (e.g., Zufferey and Gygax, 2020b), models were compared using a forward-testing approach. Fixed effects were included one at a time (main and interaction effects), and each resulting model was compared to a model that did not include the added factor. In order not to violate the assumption of an absence of multicollinearity in logistic regressions (Schreiber-Gregory, 2018), I did not include all the measures of written language competence in the models, but rather added only one measure that provided the best improvement to the model fit. When comparing models, I also evaluated the contribution of random slopes to the models by using log likelihood tests, when the random slopes were justified by the design (as suggested by Barr et al., 2013). To obtain p-values for the final model, I used the *summary()* function from the *lmerTest* package (Kuznetsova et al., 2014). Finally, all the continuous variables were centred (Schreiber-Gregory, 2018); and the statistical significance level was set at 5%.

Global analysis

First, I built a model assessing the results of all participants together. I did not include the measures of written language competence in this model, as not all of them were completed by all the participants. For instance, the TRT-F was done by secondary and high school students, but not by the group of adults. Therefore, I assessed the contribution of different measures of linguistic proficiency to the model fit separately for each age group of participants (see Analyses per group). Our model fit kept improving after adding Group (secondary school, high school, and adults) and Connective (*en effet* 'for', *toutefois* 'however', *en outre* 'moreover', and *aussi* 'therefore') (both main and interaction effects) as fixed factors. Finally, our model further improved after including Connective as a random slope by participant (see Table A1 in Supplementary materials for the comparison of the estimates of the model fit across all participants). Since adding other random slopes did not have further positive effect on the model, our final model included Group and Connective as fixed factors, Item and Participant as random intercepts and Connective as random slope by Participant. I did not add the task mode (online versus offline) into the general model because it coincided with the factor of group in the case of adults. Treatment contrasts were applied for the unordered factors of Connective and Group. The "Cause" (i.e., *en effet*) was set as reference level for comparing the scores associated to the different connectives, since *en effet* constitutes the most frequently used connective in written French. For the factor of Group, I chose "Adults" as reference, as this group was assumed to include speakers with the highest level of competence (see Table 2).

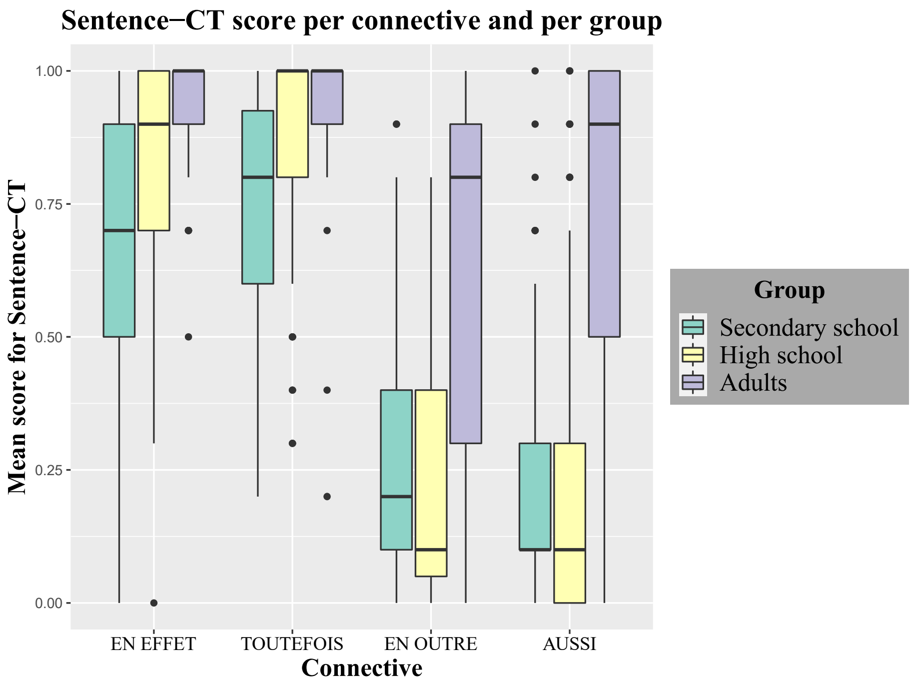
Table 2. Model's estimates for the best fitting model within all the tested groups of participants in Experiment 1

	Estimate	SE	z-value	Pr(>|z|)		Estimate	SE	z-value	Pr(>|z|)
All Participants					**High School**				
(Intercept)	3.83	0.26	14.77	**<0.001**	(Intercept)	1.08	0.40	2.70	**0.007**
GROUP					TYPE OF CLASS				
High School	-1.86	0.26	-7.10	**<0.001**	Level B	-0.53	0.34	-1.56	0.120
Secondary School	-2.89	0.29	-9.93	**<0.001**	Level A	0.01	0.28	0.05	0.958
CONNECTIVE					CONNECTIVE				
Toutefois	0.12	0.34	0.36	0.722	Toutefois	0.49	0.22	2.26	**0.024**
En outre	-3.00	0.31	-9.68	**<0.001**	En outre	-3.03	0.21	-14.73	**<0.001**
Aussi	-1.75	0.36	-4.88	**<0.001**	Aussi	-2.86	0.20	-14.01	**<0.001**
					SUBJECTIVE EXPOSURE TO PRINT	0.48	0.20	2.42	**0.016**
GROUP * CONNECTIVE					TYPE OF CLASS * CONNECTIVE				
High School by Toutefois	0.56	0.31	1.80	0.071	Level B by Toutefois	0.58	0.36	1.60	0.109
Secondary School by Toutefois	0.36	0.33	1.07	0.284	Level A by Toutefois	-0.12	0.29	-0.41	0.683
High School by En outre	-0.68	0.32	-2.12	**0.034**	Level B by En outre	-0.06	0.34	-0.18	0.860
Secondary School by En outre	0.73	0.36	2.05	**0.040**	Level A by En outre	-0.58	0.27	-2.16	**0.031**
High School by Aussi	-2.43	0.41	-5.91	**<0.001**	Level B by Aussi	-0.49	0.35	-1.41	0.158

	Estimate	SE	z-value	Pr(>|z|)		Estimate	SE	z-value	Pr(>|z|)
Secondary School by Aussi	-1.02	0.45	-2.27	**0.023**	Level A by Aussi	-1.59	0.30	-5.37	**<0.001**
Secondary School					**All Adults**				
(Intercept)	-0.08	0.48	-0.17	0.865	(Intercept)	-1.43	1.20	-1.20	0.231
TYPE OF CLASS					GROUP				
Level 1	-0.85	0.31	-2.71	**0.007**	Other Adults	-0.02	0.46	-0.05	0.964
CONNECTIVE					CONNECTIVE				
Toutefois	0.48	0.32	1.51	0.132	Toutefois	0.39	0.61	0.63	0.528
En outre	-2.25	0.36	-6.32	**<0.001**	En outre	-2.33	0.54	-4.32	**<0.001**
Aussi	-2.37	0.34	-6.95	**<0.001**	Aussi	-0.75	0.63	-1.19	0.236
TRT-F	0.50	0.23	2.18	**0.029**					
SUBJECTIVE EXPOSURE TO PRINT	0.46	0.17	2.78	**0.005**	ART-F	2.00	0.41	4.91	**<0.001**
TYPE OF CLASS * CONNECTIVE									
Level 1 by Toutefois	-0.02	0.38	-0.04	0.968	Other Adults by Toutefois	-0.68	0.55	-1.24	0.214
Level 1 by En outre	0.06	0.48	0.12	0.905	Other Adults by En outre	-1.72	0.52	-3.29	**0.001**
Level 1 by Aussi	-0.16	0.46	-0.36	0.722	Other Adults by Aussi	-2.33	0.65	-3.61	**<0.001**

As it is apparent in Figure 1, across all groups, participants performed significantly better with the more frequent connectives *en effet* ‘for’ (M= .83, 95% CI [.78, .89]) and *toutefois* ‘however’ (M= .87, 95% CI [.83, .92]) compared to the less frequent connectives *en outre* ‘moreover’ (M= .39, 95% CI [.30, .47]) and *aussi* ‘therefore’ (M= .41, 95% CI [.31, .51]), with the estimates' difference between Cause (set as reference level) and Addition being 6.83, and between Cause and Consequence 5.58. These findings replicate the result from Zufferey and Gygax (2020 a,b) on adults and high school participants. Moreover, participants from secondary school and high school on average scored lower than adults, which was reflected by the difference in estimates with the Intercept of 6.72 for Secondary School and 5.69 for High School. These estimates also show that there was a developmental trend between the two school levels, with 1.03 increase in estimates from Secondary to High School.

Figure 1. Mean proportions of correct connective production by three groups of participants across connectives in Experiment 1, with boxplots representing 50% interquartile range (IQR)

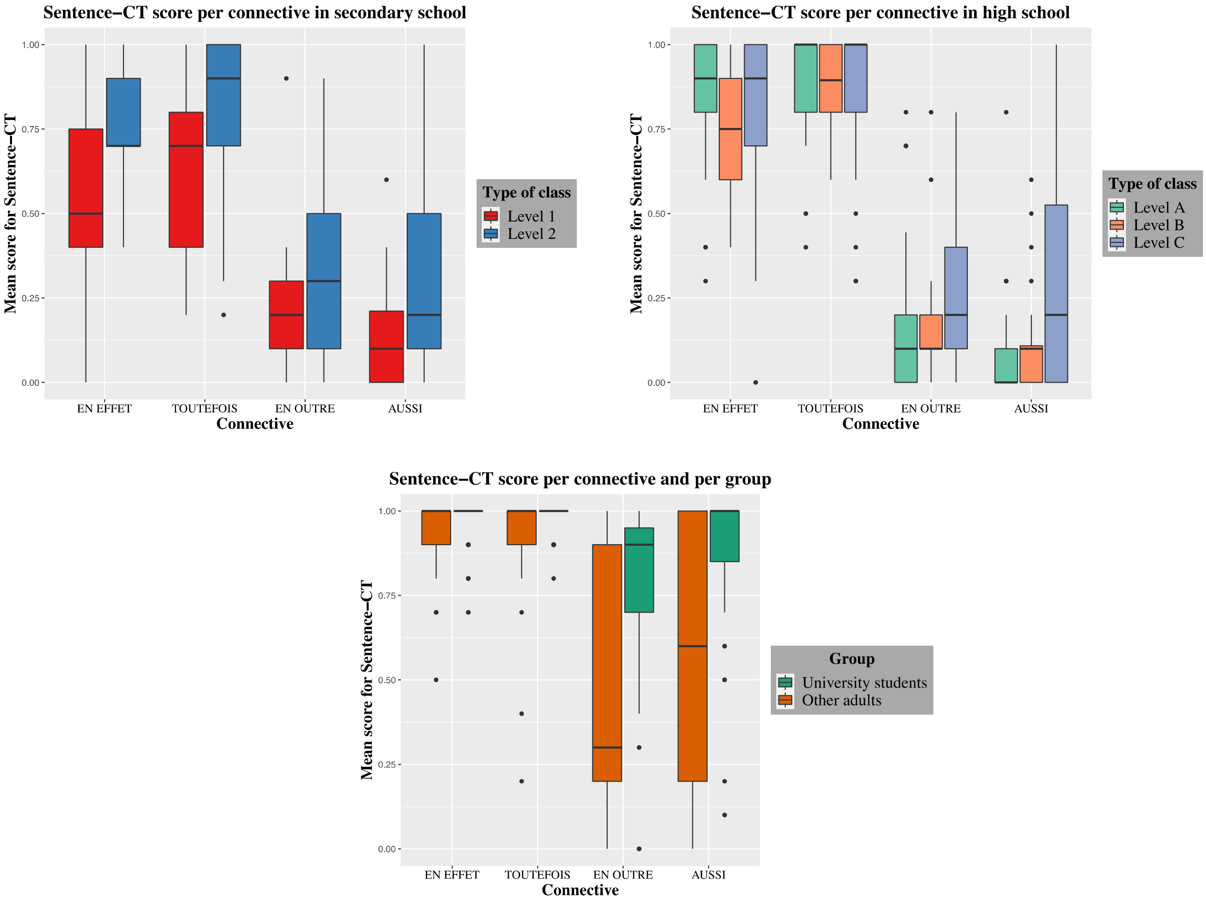


In addition, I performed a pairwise comparison between groups and connectives, using the *lsmeans()* function of the *emmeans* package in R (Lenth, 2020). First, this comparison confirmed that adults significantly outperformed teenagers across all connectives (see Table A2 in Supplementary materials for the estimates). Second, it also revealed that there was an important developmental effect for the connectives *en effet* and *toutefois* between secondary school ($M_{\text{en effet}}$=.67, 95% CI [.60, .73], and $M_{\text{toutefois}}$=.74, 95% CI [.67, .80]) and high school ($M_{\text{en effet}}$=0.81, 95% CI [.76, .87], and $M_{\text{toutefois}}$=.88, 95% CI [.84, .92]), with an increase by 1.03 and 1.23 estimates, respectively. Finally, I did not observe a significant effect for the connectives *en outre* and *aussi* between secondary school ($M_{\textit{en outre}}$ =.27, 95% CI [.21, .33], $M_{\textit{aussi}}$=.23, 95% CI [.17, .29]) and high school ($M_{\textit{en outre}}$ =.23, 95% CI [.17, .28], $M_{\textit{aussi}}$ = .21, 95% CI [.14, .28]), although participants from secondary school scored slightly higher for these two connectives.

Analyses per group

After performing a global analysis, I decided to explore the results within each age group and created three separate statistical models following the same procedure. Since there were unordered factors, I also used treatment contrasts. As described above, "Cause" (i.e., *en effet*) was set as the reference level for comparing connectives across all models. Models for secondary and high schools included the factor Type of Class, with Levels 1 and 2 for the former and Levels A, B and C for the latter. "Level 2" was chosen as the reference level for comparing classes within secondary school, and "Level C" was set as a reference for high school. Finally, the model for adults had a factor Group (University Students and Other Adults), where "University Students" were a reference for comparing the two groups of adults.

Figure 2. Mean proportions of correct connective production in Experiment 1 across the participants from secondary school (up left), high school (up right) and across adults (low), with boxplots representing 50% IQR



Secondary school. The final model, designed for the secondary school, included Type of Class, Connective, the TRT-F, and Subjective Exposure to Print as fixed factors, Item and Participant as random intercepts, and Connective as random slope by Participant (see Table A1). The ART-F and Grammatical Task did not improve the model's fit. After I tried to add other random slopes, the model did not converge. Among the participants from secondary school, students from Level 1 completed the task in class, while those from Level 2 completed it online. Since the factors of academic level and the modality of the task coincide, I did not include the task mode in the final model. Figure 2 demonstrates that teenagers from secondary school on average have higher scores for the two more frequent connectives *en effet* 'for'

(M=.67, 95% CI [.60, .73]) and *toutefois* ‘however’ (M=.74, 95% CI [.67, .80]) than for the less frequent *en outre* ‘moreover’ (M=.27, 95% CI [.21, .33]) and *aussi* ‘therefore’ (M=.23, 95% CI [.17, .29]), as was also found in the previous analysis for the whole population. According to the final model estimates (Table 2), the results for Addition (i.e., *en outre*) and Consequence (i.e., *aussi*) are indeed significantly lower than those for Cause (i.e., *en effet*). In addition, the final model estimates as well as the output of the post hoc analysis (see Table A2) indicate that students from Level 1 obtained significantly lower scores across all connectives. Out of all the measures of linguistic competence, it was the TRT-F and Subjective Exposure to Print that predicted the best the performance in the Sentence-CT by secondary school students.

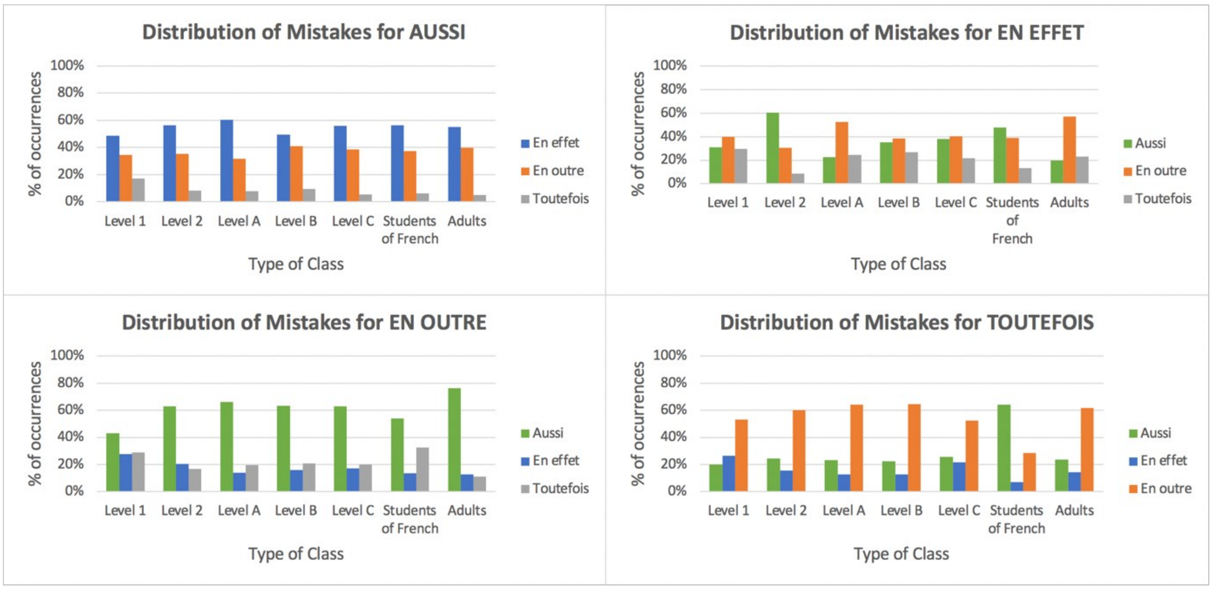
High School. The final model for high school participants included Type of Class, Connective, and Subjective Exposure to Print as fixed factors, Item and Participant as random intercepts (see Table A1). Out of all the measures of language competence, Subjective Exposure to Print was the only one that improved the model’s fit. After adding Connective as random slope by Participant and by Item, the model stopped to converge. Furthermore, the model did not improve when I added the task mode into it. Figure 2 shows that, on average, high school students also scored better with the two more frequent connectives *en effet* (M=.81, 95% CI [.76, .87]) and *toutefois* (M=.88, 95% CI [.84, .92]) than with less frequent *en outre* (M=.23, 95% CI [.17, .28]) and *aussi* (M=.21, 95% CI [.14, .28]). Similar to the findings for secondary school, the estimates of our final model (Table 2) demonstrated that the results for Addition and Consequence were indeed significantly lower than those for Cause. In addition, according to the pairwise analysis among three classes, Level C on average had higher scores than the other two levels for the connective *aussi* (Table A2). Finally, it seems that the class of Level C also scored significantly higher than the class of Level A for the connective *en outre*.

Adults. The final model for the adult participants included Group, Connective, and ART-F as fixed factors, Item and Participant as random intercept and Connective as random slope by Participant (see Table A1). The ART-F was selected for the final model, as this measure improved the model's fit the most. When adding other random slopes to the model, it stopped converging. As shown in Figure 2 and Table 2, adults also performed better with *en effet* (M=.95, 95% CI [.93, .97]) and *toutefois* (M=.95, 95% CI [.92, .98]) than with *en outre* (M=.63, 95% CI [.54, .72]) and *aussi* (M=.73, 95% CI [.64, .81]). Finally, I found a significant difference between the results of the group of adults with various backgrounds and university students, the former scoring significantly lower for the less frequent connectives *en outre* (M_{Uni}= .79, 95% CI [.72, .85], versus $M_{\text{Non-Uni}}$= .44, 95% CI [.35, .53]) and *aussi* (M_{Uni}=.88, 95% CI [.82, .93], versus $M_{\text{Non-Uni}}$= .55, 95% CI [.45, .65]) (see Table A2).

Distribution of errors

In order to ensure that the results were not due to some systematic errors because one or the other connective was consistently being chosen by mistake, I also explored the distributions of errors per connective and per type of group (Figure 3). I noticed that, both for teenagers and for adults, the most frequent error for *en outre* 'moreover' was *aussi* 'therefore', and for *aussi* it was *en effet* 'for'. There was no systematic pattern of error for *toutefois* 'however' and *en effet*.

Figure 3. The percentage of errors across all tested groups and connectives in Experiment 1



3.1.3.2. Measures of written language proficiency

Generally – and as expected –, across proficiency measures, participants from high school performed better than those from secondary school (Table A3 in Supplementary materials). Interestingly, Level 2 students outperformed Level 1 students throughout all the language measures despite the fact that both classes are from the same age group (Table 1). Among the classes from high school, it was Level C students who obtained the highest scores for all the tasks, except for the grammatical competence measure, and Level B which had the lowest means in all tasks. Neither group of teenagers reached the level of adults in any of those tasks, even though adults with various backgrounds scored lower than university students across all measures. The reliability scores of ART-F, TRT-F, and Written grammatical competence tasks are reported in the Supplementary materials.

3.1.4. Discussion

Results from Experiment 1 showed that, across all groups, teenagers performed better with more frequent connectives (*en effet* 'for' and *toutefois* 'however') than with less frequent ones (*en outre* 'moreover' and *aussi* 'therefore'). These findings are in line with the performance of the adult groups as well as with the results of previous studies on children, adults, and high school teenagers (e.g., Nippold et. al., 1992; Zufferey & Gygax, 2020 a, b). Therefore, our result supports the frequency hypothesis for connectives used in the written mode.

Another important finding concerns the difference in performance between secondary and high school students. This developmental difference was especially pronounced for more frequent connectives in general and for the connective *en effet* in particular. The analysis of results within each school level also suggested that students with a more advanced curriculum (Level 2 for Secondary School and Level C for High School) scored better across all connectives, independently of their age. For instance, there was no age difference between the classes of secondary school, while, among the classes of high school, the students from Level C had a similar age as those from Level B and were even younger than those from Level A (see Table 1). Interestingly, among high school students, Level A reached the lowest score for the connectives *en outre* and *aussi*, in spite of their being older in chronological age. In addition, Level 2 and Level C outperformed all other classes from the same school level also across all tasks of linguistic competence.

These findings demonstrate that academic background is indeed a strong predictor for students' ability to use connectives. Besides, this predictor stays valid not only during teenage years, but also for adults, as evidenced by the results from the two groups of adults. University students studying French scored higher than adults from more diverse backgrounds, especially for the less frequent connectives *aussi* and *en outre*. Furthermore, besides academic background, different measures of exposure to print also predicted the performance of different groups of participants in the Sentence-CT. The Subjective exposure to print and the TRT-F were the best predictors for the group of secondary school students, and the Subjective exposure to print was the only measure that predicted the use of connectives in high school. The ART-F was the best predictor of the competence with connectives in adults.

The distribution of errors seems to suggest some interesting mechanisms. First, regarding the connective *aussi*, its frequent replacement by *en effet* in consequence relations seems to indicate that participants are not aware of its causal meaning. This interpretation is reinforced by the fact that they tend to erroneously use it instead of *en outre* in additive relations. Taken together, these mistakes indicate that participants only associate the more frequent additive meaning to *aussi*, even though this meaning cannot be obtained in sentence initial position. This result therefore confirms that polyfunctional connectives represent an area of difficulty, even for native speakers. Interestingly, when participants use an erroneous connective in concessive relations it is the additive connective *en outre*. This misuse seems to suggest that our participants have not yet encoded a specific value to this connective, as negative relations are rarely mixed up with positive ones, for example in discourse annotation experiments. This problem is a further indication that infrequent connectives are not mastered by a good portion of native speakers.

As mentioned earlier, our results may result from the use of an artificially low context, limited to two sentences. In Experiment 2, I address this issue by conducting a similar experiment, yet in a more ecological experimental design with texts instead of pairs of sentences.

3.2. Experiment 2. Text-level cloze test (Text-CT)

3.2.1. Participants

A group of 85 teenagers from 13 to 19 years old (M_{age}=15.51; SD=1.29) took part in Experiment 2. Some participants came from secondary school (Level 1 and 2) and others from high school (Level C). A more detailed view of the distribution of participants between levels and classes is reported in Table 3. As for the adult groups, they included 40 university students specialised in French and 47 adults recruited on the Prolific platform. All the participants were native French speakers.

Table 3. Distribution of participants across all groups in Experiment 2

		N	M_{age}	SD	Range
All Teenagers		**85**	**15.51**	**1.29**	**13 - 19**
Secondary School		43	14.5	0.80	13 - 16
	Level 1	29	14.46	0.90	13 - 16
	Level 2	14	14.58	0.51	14 - 15
High School (Level C)		42	16.43	0.91	15 - 19
	Year 1	19	15.68	0.67	15 - 17
	Year 2	23	17.04	0.56	16 - 19
All Adults		**87**	**27.06**	**8.55**	**18 - 71**
	University Students	40	22.73	3.73	18 - 33
	Other Adults	47	30.74	9.73	18 - 71
Total		**172**	**17.92**	**4.25**	**13 - 71**

3.2.2. Materials and procedures

3.2.2.1. Text-level cloze test

The materials for this task consisted in 8 short texts (approximately 250 words each) with blanks that participants had to fill in with an appropriate connective. All the omitted connectives were in sentence-initial position, similarly to Experiment 1. They had to choose between the same four connectives as in the Sentence-CT (i.e., *aussi* 'therefore', *en effet* 'for', *en outre* 'moreover', and *toutefois* 'however'). The texts were real texts taken from websites presenting news and various cultural and historical phenomena for teenagers (https://dimoitou.ouest-france.fr; https://www.1jour1actu.com) and adapted to the purposes of the study. The texts covered a wide range of topics: new technologies in archaeology, Black Friday and the environment, urbanisation, fake news, housework and gender equality, recurrent flooding in Venice, the dangers of smoking and the 70th anniversary of China. In their final version, each text included four paragraphs with one missing connective per paragraph. In total, each connective had to be used once in every text. All the paragraphs were controlled for alternative interpretations so that only the intended connective was possible to build a coherent text. The participants first saw the whole text with blanks and had to read it. Afterwards, they saw one by one separate paragraphs of the same text and had to fill in the blanks with an appropriate connective. All the participants did this task online via a weblink. To access the full materials for this task, check the supplementary materials.

3.2.2.2. ART-F, TRT-F, and Written grammatical competence

The three measures of language competence were the same as in Experiment 1.

3.2.3. Results

3.2.3.1. Connectives in the Text-level cloze test

I analysed the results using the same procedure as in Experiment 1, fitting mixed-effect logistic regression models on the binary variable. I started by performing a global analysis across all groups, and applied the same treatment contrasts for the factors Connective and Group as in Experiment 1.

Global analysis

The model kept improving after I added Group (Secondary School, High School and Adults) as fixed factor, then main and interaction effects of Connective (Cause versus Concession versus Addition versus Consequence), and finally Connective as a random slope by Participant. Since adding other random slopes did not improve the model fit, the final model included Group and Connective as fixed effects, Item and Participant as random intercept and Connective as random slope by Participant (see Table A4 for the estimates of the model fit for all participants).

As in the *Sentence-CT*, across all groups, participants reached higher scores for the more frequent connectives *en effet* 'for' (M=.68, 95% CI [.62, .74]) and *toutefois* 'however' (M=.69, 95% CI [.62, .76]), and significantly lower scores for the less frequent connectives *aussi* 'therefore' (M=.41, 95% CI [.33, .49]) and *en outre* 'moreover' (M=.37, 95% CI [.30, .44]), with an estimated difference of 3.64 between Cause and Addition, and 3.11 between Cause and Consequence (see Figure 4 and Table 4). In addition, a general developmental effect can be observed between the three groups. Secondary School students have an estimate of -0.79 ± 0.23 SE, High School students of -1.14 ± 0.23 SE and Adults of 1.75 ± 0.27 SE. Finally, a post hoc pairwise comparison (see Table A5) demonstrated that there was a significant difference in scores between teenagers and adults across all connectives. This means that even the most advanced students still did not reach the performance of adults even for frequent connectives. In addition, a significant developmental effect can also be observed between the two school levels for the connectives *en effet* and *toutefois*, with an estimated increase by 0.65 and 0.72, respectively.

Figure 4. Mean proportions of correct connective production across all participants in Experiment 2, with boxplots representing 50% IQR

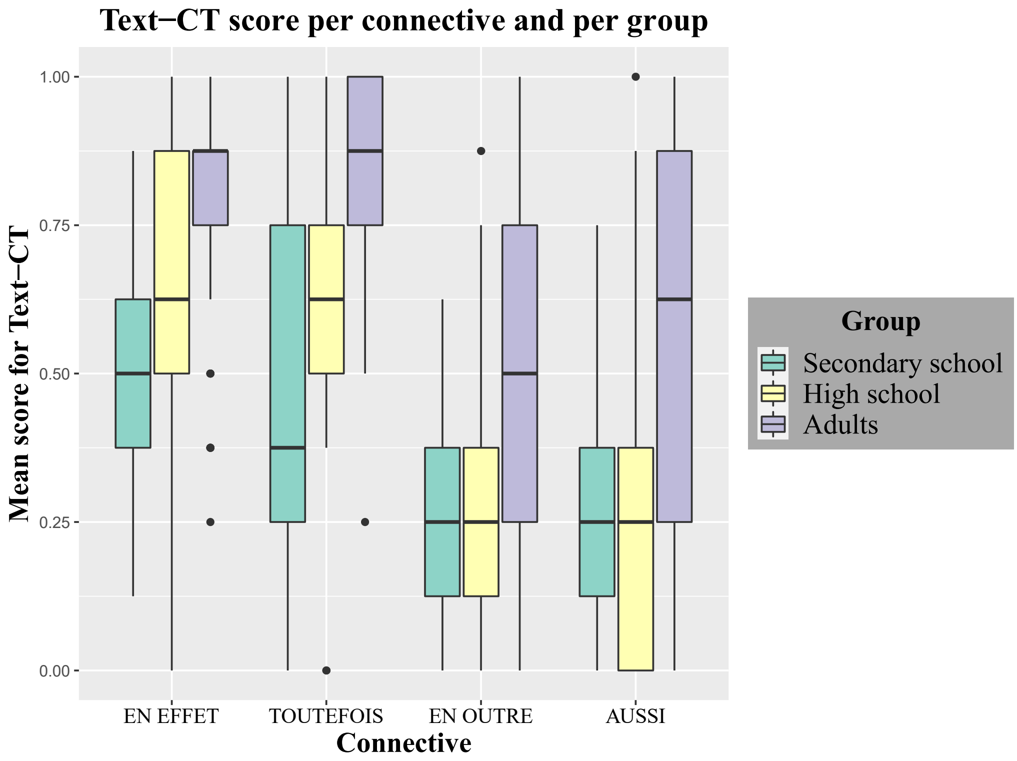


Table 4. Model's estimates for the best fitting model within all the tested groups of participants in Experiment 2

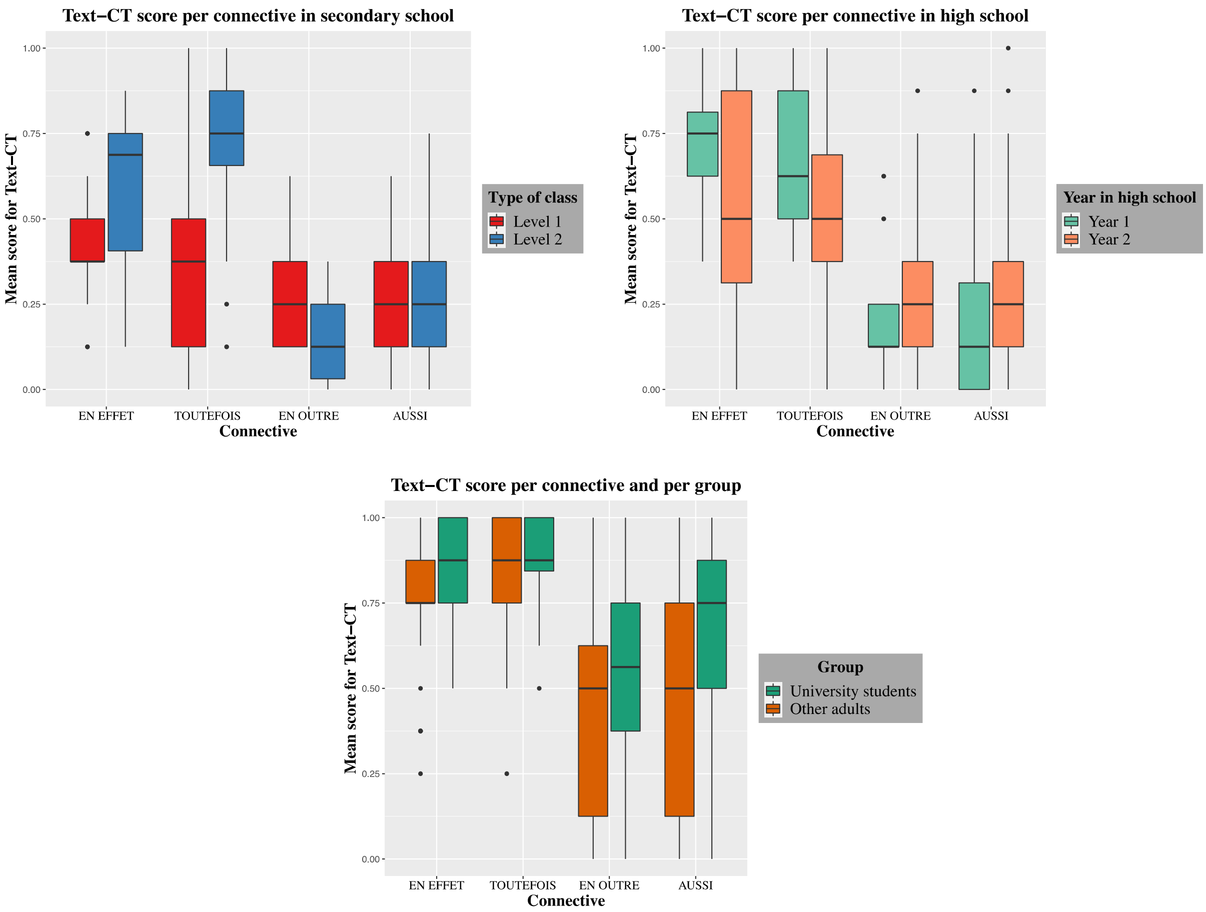
Model	Estimate	SE	z-value	Pr(>|z|)	Model	Estimate	SE	z-value	Pr(>|z|)
All Participants					**High School**				
(Intercept)	1.75	0.27	6.45	<0.001	(Intercept)	0.67	0.30	2.19	0.028
GROUP					CONNECTIVE				
High School	-1.14	0.23	-4.95	<0.001	*Toutefois*	-0.13	0.38	-0.33	0.739
Secondary School	-1.79	0.23	-7.79	<0.001	*En outre*	-1.85	0.44	-4.23	<0.001
CONNECTIVE					*Aussi*	-2.17	0.46	-4.72	<0.001
Toutefois	0.66	0.38	1.73	0.084					
En outre	-1.89	0.37	-5.11	<0.001					
Aussi	-1.36	0.39	-3.51	<0.001					
GROUP * CONNECTIVE									
High School by Toutefois	-0.73	0.29	-2.54	0.011					
Secondary School by Toutefois	-0.80	0.29	-2.74	0.006					
High School by En outre	-0.02	0.30	-0.07	0.943					
Secondary School by En outre	0.75	0.29	2.57	0.010					
High School by Aussi	-0.66	0.36	-1.82	0.068					
Secondary School by Aussi	0.04	0.35	0.11	0.911					
Secondary School					**All Adults**				
(Intercept)	0.15	0.33	0.45	0.650	(Intercept)	-2.20	0.68	-3.22	**0.001**
TYPE OF CLASS					GROUP				
Level 1	-0.84	0.30	-2.83	0.005	Other Adults	-0.02	0.30	-0.08	0.934
CONNECTIVE					CONNECTIVE				
Toutefois	0.36	0.44	0.83	0.407	*Toutefois*	0.43	0.55	0.78	0.439
En outre	-2.38	0.45	-5.33	<0.001	*En outre*	-2.10	0.51	-4.15	**<0.001**

Model	**Estimate**	**SE**	**z-value**	**Pr(>|z|)**	**Model**	**Estimate**	**SE**	**z-value**	**Pr(>|z|)**
Aussi	-1.49	0.42	-3.54	<0.001	*Aussi*	-1.27	0.54	-2.35	**0.019**
SUBJECTIVE EXPOSURE TO PRINT	0.33	0.13	2.59	0.010	ART-F	1.59	0.21	7.67	**<0.001**
TYPE OF CLASS * CONNECTIVE					GROUP * CONNECTIVE				
Level 1 by Toutefois	-0.84	0.47	-1.81	0.070	Other Adults by Toutefois	0.17	0.40	0.43	0.671
Level 1 by En outre	1.80	0.46	3.93	<0.001	Other Adults by En outre	-0.03	0.32	-0.10	0.920
Level 1 by Aussi	0.55	0.44	1.24	0.216	Other Adults by Aussi	-0.50	0.41	-1.22	0.221

Analyses per group

Secondary school. After a global analysis, I analysed the results within each group of participants. I began with the classes of secondary school. The final model included Type of Class (Level 1 and Level 2), Connective, and Subjective Exposure to Print as fixed factors, Item and Participant as random intercept and Connective as random slope by Participant (see the final model for secondary school in Table A4). I included Subjective Exposure to Print to the final model, as it was the only measure of individual variation that improved the model's fit. As can be seen in Figure 5, classes of Level 1 scored on average significantly lower than those of Level 2 (an estimate - 0.84 ± 0.30 SE), and all classes performed worse with *en outre* 'moreover' (M=.27, 95% CI [.23, .31]) and *aussi* 'therefore' (M=.25, 95% CI [.21, .30]) than with *en effet* 'for' (M=.49, 95% CI [.44, .54]) and *toutefois* 'however' (M=.47, 95% CI [.39, .54]). Moreover, the pairwise comparison revealed a significant difference between the two classes, showing that Level 2 scored significantly higher for the connectives *en effet* (M_{Level2}=.61, 95% CI [.55, .67], versus M_{Level1}=.44, 95% CI [.39, .48]) and *toutefois* (M_{Level2}=.69, 95% CI [.62, .75], versus M_{Level1}=.39, 95% CI [.30, .43]), and slightly lower for *en outre* (M_{Level2}=.17, 95% CI [.13, .21], versus M_{Level1}=.32, 95% CI [.28, .36]) (see Table A5).

Figure 5. Mean proportions of correct connective production across the participants from secondary school (up left), high school (up left) and across adults (low) in Experiment 2, with boxplots representing 50% IQR



High School. In order to analyse the results from high school students, I used a model including Connective as fixed factor, Item and Participant as random intercepts and Connective as random slope by Participant (see the final model for high school in Table A4). None of the measures of written language competence improved the model's fit. Even though this group included classes from Year 1 and 2, adding this factor to the model did not improve its fit (see Table A4), suggesting that the performance with connectives within the same academic level does not necessarily depend on the age of the participants (see Table 3 for the distribution of age among different groups of participants). Furthermore, as it is apparent in Figure 5 and Tables 9 and 10, high school students also scored significantly higher with the more frequent connectives *en effet* (M=.62, 95% CI [.55, .68]) and *toutefois* (M=.60, 95% CI [.53, .66]) than with *en outre* (M=.26, 95% CI [.20, .31]) and *aussi* (M=.26, 95% CI [.19, .33]).

Adults. The final model for adults included Group, Connective, and ART-F as fixed factors, Item and Participant as random intercepts and Connective as random slope by Participant (see the final model for adults in Table A4). The ART-F was included to the final model, as it provided to the best improvement to the model's fit. As can be observed in Figure 5 and Table 4, both groups of adults also performed better with *en effet* (M =.80, 95% CI [.76, .85]) and *toutefois* (M = .85, 95% CI [.81, .89]) than with *aussi* (M =.56, 95% CI [.48, .65]) and *en outre* (M = .47, 95% CI [.41, .54]); although their scores were higher than those of teenagers. Furthermore, the performance of the group of adults with various backgrounds was on average lower than that of the university students across all the connectives (*en effet:* M_{Uni}= .85, 95% CI [.82, .89], versus $M_{Non\text{-}Uni}$= .76, 95% CI [.71, .81]; *toutefois*: M_{Uni}= .88, 95% CI [.85, .92], versus $M_{Non\text{-}Uni}$= .82, 95% CI [.77, .87]; *aussi*: M_{Uni}=.68, 95% CI [.60, .75], versus $M_{Non\text{-}Uni}$=.47, 95% CI [.39, .55]; *en outre*: M_{Uni}=.55, 95% CI [.49, .62], versus $M_{Non\text{-}Uni}$=.41, 95% CI [.34, .47]) (see Table A5).

Distribution of errors

I also examined the distributions of errors per connective and per group (Figure 6). In this task, no specific pattern was revealed with the connectives *en effet* 'for', *en outre* 'moreover' and *toutefois* 'however' in teenagers. In contrast, the most frequent mistake for *aussi* 'therefore' was *en effet* (similar to *Sentence-CT*). Among the adult groups, there was no observable trend for mistakes with the connectives *aussi*, *en effet*, and *toutefois*, while for *en outre*, the most common mistake was *aussi* (as in the *Sentence-CT*).

Figure 6. The percentage of errors across all tested groups and connectives in Experiment 2

3.2.3.2. Measures of written language proficiency

The scores across all language measures increased between secondary and high school (Table A6). As in Experiment 1, participants from Level 2 performed better than those from Level 1 throughout all tasks, except for the TRT-F where the trend was opposite. However, even the group of the most advanced teenagers (Level C) did not manage to reach the same scores as adults in any of the tasks.

3.2.4. Discussion

As in Experiment 1, across all groups of participants, the more frequent connectives *en effet* 'for' and *toutefois* 'however' lead to higher scores than the less frequent *en outre* ‘moreover’ and *aussi* ‘therefore’. Moreover, there was a developmental trend for frequent connectives from secondary school students to adults. Yet, there was no visible progress with infrequent connectives between secondary and high school.

A more detailed analysis within each age group also revealed interesting findings. When comparing the students from secondary school, who were of the same age, but differed in their academic background, important differences of competence were observable. The students from the more advanced Level 2 scored significantly higher for frequent connectives than those from Level 1. In contrast, when comparing students from high school of different age groups but with the same academic background (Level C), no difference in performance with the Text-CT were found. This may suggest that academic level is the strongest predictor of competence with connectives.

Interestingly, the difference of academic backgrounds persisted also among adults. While the results for frequent connectives were very similar among the two groups of adults, those for infrequent connectives differed between students of French and adults from more diverse backgrounds. This result suggests again that age is not the most important factor for the competence with connectives. Instead, the mastery of connectives probably depends on the level of written language competence. Indeed, after the academic background, different measures of exposure to print were found to be important predictors of the performance in the Text-CT for the two age groups, which included participants from different backgrounds, namely secondary school students and adults. For secondary school students, it was their subjective evaluation of exposure to print and, for adults, it was the ART-F.

It was more difficult to see a trend in the production of errors by teenagers in this task, as the range of incorrect answers greatly varied. The fact that teenagers chose *en effet* instead of *aussi* may reflect that they correctly identified a causal link between the segments, but incorrectly used a causal connective instead of a consequence connective because they failed to attribute the consequence meaning to *aussi*. Generally, the fact of having more context seems to have increased uncertainty about the intended coherence relation compared to Experiment 1, as it is not clear from the pattern of errors which relations participants inferred when they misused a connective. This result brings me to the next experiment, in which I further explore how a greater context affects the competence with the four tested connectives. In other words, the next experiment aims to unveil whether a broader context hinders only the production of connectives or whether it also challenges the comprehension of the coherence relations signalled by those connectives.

3.3. Experiment 3. Comprehension test

3.3.1. Participants

A group of 114 teenagers aged 12–22 (M_{age}=16.32; SD=1.50) and 70 adults aged 18–57 (M_{age}=26.20; SD=8.99) took part in Experiment 3. All students from secondary school belonged to Level 2; whereas high school students were represented by all three academic levels (A, B, and C). As in previous experiments, the group of adults also included university students and adults from the Prolific platform. Table 5 reports a detailed distribution of participants between groups and levels. All the participants were native French speakers.

Table 5. Distribution of participants across all groups in Experiment 3

	Group specifications	N	M_{age}	SD	Range
All Teenagers		**114**	**16.32**	**1.50**	**12 - 22**
Secondary School		21	14.43	0.73	12 - 15
Level 2	More advanced curriculum in French	21	14.43	0.73	12 - 15
High School		93	16.75	1.28	15 - 22
Level A	Curriculum preparing for a direct professional path	35	17.71	1.19	16 - 22
Level B	Curriculum preparing for a professional High School	17	15.82	0.79	15 - 17
Level C	Curriculum preparing for a university path	41	16.32	0.95	15 - 19
All Adults		**70**	**26.20**	**8.99**	**18 - 57**
University Students	Specialisation in French Studies	30	23.43	4.92	19 - 40
Other Adults	Heterogeneous backgrounds	40	28.28	10.65	18 - 57
Total		**184**	**20.08**	**7.43**	**12 - 57**

3.3.2. Materials and procedures

3.3.2.1. Comprehension test

The comprehension task consisted of the same eight texts as in Experiment 2, with the only difference that connectives were not omitted this time, and each text included four written connectives, namely *aussi* 'therefore', *en effet* 'for', *en outre* 'moreover', and *toutefois* 'however'. Participants had to read a text and evaluate the truthfulness of the statements about the content of the text on a 6-point scale, ranging from from 0=*completely false* to 5=*completely true*. Four statements were true and targeted the relations between the sentences, connected by connectives; while four other statements were false and targeted other parts of the text, not signalled by connectives. All the statements about the content of the text were formulated with connectives typically used in speech and expressing equivalent coherence relations, namely *donc* 'so', *parce que* 'because', *en plus* 'and', and *mais* 'but'.

Here is an extract of a text about new technologies in archaeology (9). Statement (9a) targets the relation, signalled by the causal connective *en effet* and is formulated with the equivalent causal connective *parce que*. Statement (9b) is an example of a filler, using the same causal connective.

(9) … Les drones sont notamment très utiles. <u>En effet</u>, ils permettent de faire rapidement des milliers de photos aériennes des sites de fouilles ou de monuments. On utilise aussi des outils empruntés à la médecine, comme le scanner. Les rayons X permettent de voir ce qu'il y a à l'intérieur d'un objet sans l'abîmer ! On peut ensuite se servir d'une imprimante 3D pour créer une copie des objets que l'on a vus. …

'Drones are particularly useful. Because[1] they can quickly take thousands of aerial photographs of excavation sites or monuments. Tools borrowed from medicine, such as the scanner, are also used. X-rays allow us to see what is inside an object without damaging it! A 3D printer can then be used to create a copy of the objects seen.'

(9a) Les archéologues utilisent les drones, parce qu'ils fournissent beaucoup d'images.

'Archaeologists use drones, because they provide a lot of images.'

(9b) On utilise très peu les nouvelles technologies pour le moment, parce qu'elles ne nous apprennent pas beaucoup de choses.

'I don't use new technologies a lot at the moment, because they don't teach us much.'

3.3.2.2. ART-F, TRT-F, and Written grammatical competence

The measures of language competence were the same as in Experiments 1 and 2.

3.3.3. Results

3.3.3.1. Connectives in the Comprehension test

The data analysis was conducted using the same procedure as in Experiments 1 and 2, with the only difference that I fitted linear regression mixed-effects models on the continuous variable, which was a judgement score in the comprehension test. The models

[1] Note that this is an approximate translation in English.

were tested with the *lmer* function of the *lme4* package (Bates et al., 2015). I started by performing a global analysis across all groups, and applied the same treatment contrasts for the factors Connective and Group as in Experiments 1 and 2. Since there were less participants from secondary school than from high school and all the secondary-school students belonged to the same type of class (Level 2), I did not analyse this group of participants separately. Therefore, the global analysis is followed by per-group analyses made separately for teenagers (including both school levels) and adults.

Global analysis

The final model for all participants included Group (Adults versus Teenagers) as fixed factor and Item and Participant as random intercepts. Adding Connective as fixed factor as well as random slopes did not improve the model's fit (see Table A7 for model comparisons). Results revealed that participants had a good level of comprehension of true statements about the text, independently from the type of connective that was used and from the coherence relation that these connectives signalled in the texts, as participants gave an average judgement score of 4.21 (SD=1.25, Range: 0–5). None of the measures of written competence predicted the variation in the comprehension task. Finally, adults (*M*=4.32, SD=1.21) on average gave slightly higher judgement scores than teenagers (*M*=4.15, SD=1.27) for the true statements about the texts (see Figure 7).

Figure 7. Mean judgement score by teenagers and adults across connectives in Experiment 3, with boxplots representing 50% IQR

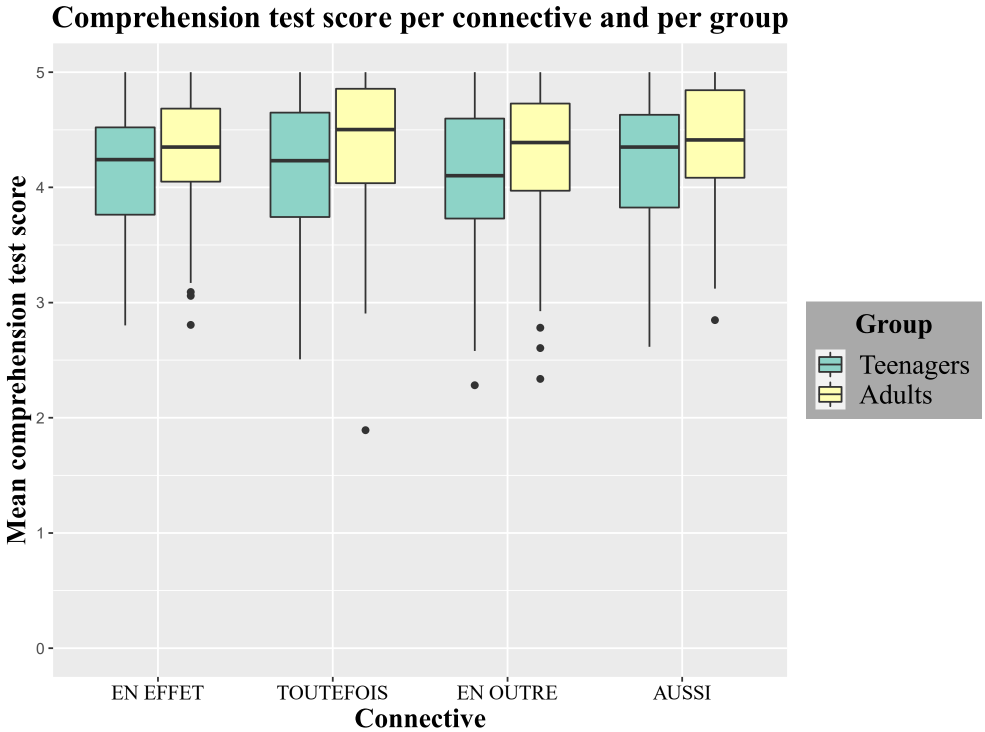


Table 6. Model's estimates for the best fitting model within all the tested groups of participants in Experiment 3

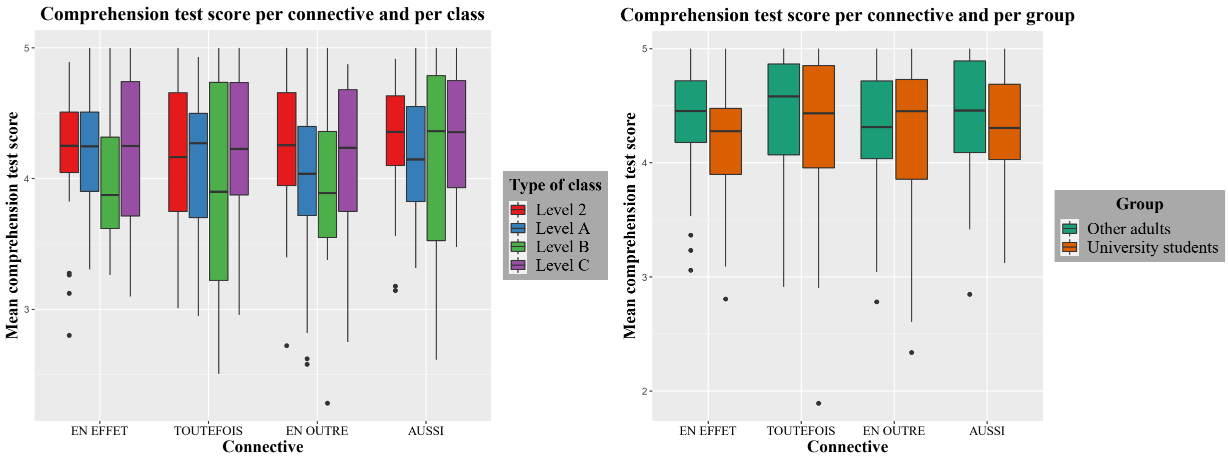
Model	**Estimate**	**SE**	**df**	**t-value**	**Pr(>|t|)**
All Participants					
(Intercept)	1.62	0.02	81.49	79.55	**<0.001**
GROUP					
Teenagers	-0.04	0.02	182.00	-2.08	**0.039**
All Teenagers					
(Intercept)	1.51	0.04	135.22	41.39	**<0.001**
Subjective Exposure to Print	0.04	0.02	112.22	2.22	**0.028**
All Adults					
(Intercept)	1.59	0.03	85.68	57.26	**<0.001**
GROUP					
Other Adults	0.06	0.03	68.00	1.98	**0.051**

Analyses per group

Teenagers. The final model for the group of teenagers consisted of Subjective Exposure to Print as fixed factor and Item and Participant as random intercepts, as all other predictors did not improve the model's fit (see Table A7 for the model comparisons). The analysis of performance within the group of teenagers showed that there was no significant difference in the judgement scores between different types of classes, connectives, and teenagers of different ages (see Figure 8 and Table 6). As for the measures of language competence, the subjective evaluation of exposure to print was the only significant predictor of the variation in the comprehension test.

Adults. The final model for the group of adults included Group (University Students versus Other Adults) as fixed factor and Item and Participant as random intercepts, as adding all other variables did not improve the model's fit (see Table A7 for the model comparisons). Analysis of results by adults showed that they also gave quite high judgement scores for the true statements, targeting different types of connectives. Variation in their responses was not predicted by any of the measures of written competence. Surprisingly, the group of adults with various backgrounds (M= 4.38, SD= 0.52) on average gave slightly higher judgement scores than university students (M= 4.24, SD= 0.63) for the true statements about the texts (see Figure 8).

Figure 8. Mean judgement score by teenagers (left) and adults (right) across connectives in Experiment 3, with boxplots representing 50% IQR



3.3.3.2. Measures of written language competence

Except for the TRT-F, the scores across all other measures increased between teenagers and adults (Table A8). However, similarly to both previous experiments, the group of teenagers with the most advanced academic background (Level C) obtained lower scores than adults in all the tasks. As for the performance in TRT-F, the younger group of secondary school students (Level 2) scored better than all three groups of high school students, which may suggest that this test is better tailored for measuring the degree of exposure to print of younger speakers.

3.3.4. Discussion

The connective comprehension test revealed that the comprehension of all four types of coherence relations was high across all the participants, as reflected by high judgement scores of the statements about those relations. In the global analysis, adults were found to score higher than teenagers; and in the within-group analysis, adults with heterogeneous backgrounds had higher scores than university students. However, it must be mentioned that the observed differences, in spite of being statistically significant, do not seem to be really big. This seems to be particularly plausible since almost none of the measures of written grammatical competence predicted the variation in the judgement scores across the participants. Only the Subjective exposure to print predicted the variation in scores among teenagers. This finding may therefore indicate that the judgement scores were not really related to the development of general linguistic proficiency.

3.4. General discussion

Across a series of production and comprehension tasks, I assessed how the use and comprehension of four different French connectives bound to the written mode develops from teenage years to adulthood. This study goes beyond previous work assessing the mastery of connectives in several ways. First, to the best of our knowledge, it is the first study to assess the ability to use and understand connectives that includes such a large age range of speakers, extending from early teenage years to adulthood. Second, within each age group, proficiency with connectives was compared between groups with various levels of academic competence, enabling us to examine the roles of age and academic competence together. Third, the use of connectives was compared across two types of cloze tests: one with a reduced two-sentence context and the other with a larger and more ecological textual context. Finally, the production of connectives in texts was compared to the comprehension of the coherence relations, encoded by these connectives, in the same texts.

3.4.1. Frequency versus cognitive complexity

The two production tasks provide converging evidence regarding the role of academic level and connectives' frequency. Indeed, significant differences were found between groups with different levels of academic achievement in both tasks and in all age groups. Similarly, in both tasks, connectives with a lower frequency in corpus data (*en outre* 'moreover' and *aussi* 'therefore') were mastered less well than connectives with a higher frequency (*en effet* 'for' and *toutefois* 'however'). This result indicates that, once speakers have acquired the basic meaning and function of connectives, their frequency becomes a better predictor of their difficulty compared to the degree of cognitive complexity of the coherence relation they encode. This is true at least in off-line tasks such as those presented in this chapter. However, it is possible that off-line measures, targeting older cohorts of speakers, cannot really capture the effect of coherence complexity on the competence with connectives. Indeed, previous studies on adults, which reported a greater difficulty in processing of subjective in comparison to objective causal relations (Canestrelli et al., 2013; Traxler, Bybee, et al., 1997; Traxler, Sanford, et al., 1997) and of negative relations as opposed to positive ones (Morera et al., 2017; Xu et al., 2015, 2018), used on-line processing measures. Future studies will need to assess whether teenagers and adults process the tested coherence relations differently when they are conveyed with connectives typical for the written mode. This would contribute to the research on the on-line processing of discourse that mostly focus on the connectives typically used in the spoken mode.

3.4.2. The role of context

The comparison between the sentence-level (Experiment 1) and the text-level (Experiment 2) cloze test also revealed important differences in the use of connectives in different experimental contexts. First, adults had lower scores across all connectives in the Text-CT compared to the Sentence-CT. A similar trend was found for teenagers who, for more frequent connectives, also scored lower in the Text-CT. This finding suggests that the process of interpretation of coherence relations in a text might require an additional cognitive load in comparison to a more reduced context of pairs of sentences. Different patterns of error distribution between the two tasks might be a reflection of this additional difficulty. The Text-CT resulted in a more heterogeneous pattern of errors, as a bigger text is probably cognitively more demanding and may therefore make it more difficult to match a connective with an appropriate coherence relation. This effect was not observed among teenagers for the less frequent connectives, probably because their scores on these connectives were already quite low in the reduced context of sentence pairs. As a result, the performance with these connectives could not really further deteriorate because of a more naturalistic context.

The fact that a more ecological experimental environment hinders the use of written connectives is still very revealing, since one might have expected that a richer context could increase the ability in infer the intended coherence relation. Even though the correct relation may indeed have been inferred, as suggested by the outcome of the connective comprehension task, it did not always lead to the use of an appropriate connective. Hence, this brings us to the conclusion that participants did not know the actual function of some connectives from the written mode.

This finding is not in line with previous studies where context was reported to facilitate the performance with connectives or not to impact it at all (Scholman & Demberg, 2017; Yung et al., 2019). The discrepancy with my findings can stem from the differences between the experimental designs. In the studies of Scholman and Demberg (2017) and Yung et al. (2019), arguments that had to be linked by a connective were marked by a black colour and the sentences giving additional context were grey. Moreover, the order of arguments and contextual elements was always the same. As a result, participants knew which sentences were supposed to be linked by a connective and which ones were there just to provide context. This was not the case in the second experiment, reported in this chapter, as participants had to establish by themselves which elements of the text were important to consider, hence making the task more challenging but also more ecologically valid.

3.4.3. Production versus comprehension

The results of the comprehension task demonstrated that the level of understanding of the four coherence relations was high among all the tested groups of speakers. In other words, all the participants successfully identified the correct statements about the parts of texts, linked by the four tested connectives, independently of the type of coherence relation that they signal and of their frequency. This finding may be revealing of a general difference between the mechanisms of production and comprehension, as speakers may be able to understand certain connectives and coherence relations that they express, but not yet be capable to use these connectives accurately in their own production (see, e.g., Cain & Nash, 2011). Moreover, the absence of difference in scores between connectives with low and high frequencies, which were observed in previous experiments, may be partially explained by the fact that both infrequent connectives signalled continuous coherence relations, namely consequence (*aussi*) and addition (*en outre*). This means that these relations can be also easily conveyed without explicit marking by connectives (see, e.g., Asr & Demberg, 2012). Therefore, even if the meaning of the less frequent connectives *aussi* and *en outre* was not clear, the meaning of the segments that they linked could have been reconstructed without the knowledge of these specific connectives.

However, the fact that the judgement scores on this task were high and that there was almost no variation between participants may also have a different interpretation, as the design of this task had several biases. It is possible that including equivalent oral connectives in the statements about the content of the texts may have impacted the results, by facilitating the comprehension of the targeted coherence relation. In addition, the outcome of this test may have been in part biased by the fact that all the non-filler items were correct and all the filler items were false. As a result, the findings from this task should be interpreted with caution and further verified.

3.4.4. Individual variation

The variation in performance with connectives was observed not only between connectives having different frequencies in corpus data and between tasks, but also between participants. In line with the initial hypothesis and with the studies by Nippold et al. (1992) and Zufferey and Gygax (2020b), there was a development in the performance with connectives between teenagers and adults across all connectives. The comparison of the results within different age groups however revealed that age was not the most relevant predictor of the competence with connectives. Even though the high school students had higher scores than the students from secondary school for frequent connectives in the Sentence-CT and Text-CT, the results for infrequent connectives did not differ between the two age groups in both tasks. Furthermore, among the secondary school students, teenagers with a lower academic level scored significantly lower compared to the students with a higher academic level on the Sentence-CT. In the Text-CT, in contrast, they differed only on the frequent connectives. It is possible that the difference between the groups of teenagers was found mostly for frequent connectives is compatible with an explanation in terms of exposure to print. Indeed, students with a higher academic level who read more as part of their curricula, have had enough exposure to frequent connectives to master them better, but not yet for infrequent ones. A different pattern is found within the adult groups. While all of them have been exposed to connectives long enough to master the frequent ones, differences between them are still found for the infrequent connectives. In this case, only students of French language and literature, who are expert of written French, master them fully.

The importance of exposure to print was also revealed in within-group analyses of the two production tasks, showing that, after the factor of academic background, different measures of exposure to print were the next most relevant predictors of connective mastery. Among adults, it was the ART-F that best predicted variation in the connective use across both tasks. Among teenagers, the predictors were more variable. The accuracy of the responses of the secondary school students was predicted by the Subjective exposure to print and the TRT-F in the sentence-level cloze test and by the Subjective exposure to print in the text-level cloze test. As for the high school students, the Subjective exposure to print predicted the variation in the Sentence-CT. However, none of the measures predicted the variation in the Text-CT, which may stem from the fact that a very homogeneous group of Level C students took part in this task. In line with the initial hypothesis, the fact that the ART-F was not an important predictor of variation among teenagers may suggest that this measure is more suitable for older cohorts of speakers, which was also reflected by higher scores on this measure among adult speakers than among teenagers. In contrast, the Subjective exposure to print tend to better reveal the variation among younger speakers. The marginal effect of the TRT-F probably suggests that this measure was not sensitive enough to reveal differences especially within the older group of teenagers and should be further improved in future research.

Finally, a better mastery of written connectives was not predicted by the teenagers' level of written grammatical competence, as measured by the Grammatical test. There might be two potential explanations to this finding. The first reason may be related to the nature of the grammatical task. The ability to identify correct and incorrect grammatical forms involves a grammatical analysis on a word level, while the ability to match a connective with an appropriate coherence relation involves a more global analysis on the level of the whole sentence or text. Therefore, the Grammatical test and the Cloze tests might simply target different, non-connected competences. Alternatively, it is also possible that the Grammatical test, used in the current study, was too difficult for the young speakers, and there was a floor effect. Hence, to avoid such effect, future research should consider adapting this test to younger and less proficient populations.

3.5. Conclusion

The study described in this chapter supports previous findings of Zufferey and Gygax (2020b) on the role of frequency for the mastery of connectives and extends it on a broader cohort of speakers from early teenage years to adulthood. Academic background was one of the strongest predictors of the variation in the competence with connectives typically used in writing. After academic background, different measures of exposure to print were important predictors of the performance with connectives within different age groups of speakers. This finding is in line with the idea that different types of measures are more suitable for the populations of different age (Cunningham & Stanovich, 1990). Moreover, this result also corroborates previous research, revealing the connection between inter-individual differences in the degree of exposure to print and the mastery of connectives (Zufferey & Gygax, 2020a) and coherence relations inferences (Scholman et al., 2020). Finally, it is important to mention that teenagers from the oldest age group (aged 16–17) did not reach an adultlike performance, even when compared to the group of adults from a general population. This suggests that the mastery of connectives continues to develop way beyond teenage years and can evolve throughout the whole lifespan.

The results of the experiments presented in this chapter have also unveiled several directions that will be examined in the next chapter. First of all, the present study has assessed the use and comprehension of a limited number of connectives, and therefore it should be extended to a broader group of connectives, varying in frequencies both in written and oral modes, in order to provide a more solid confirmation for the effects found on this rather limited set.

Moreover, it would be useful to better disentangle the effects of frequency and polyfunctionality. The error analysis, conducted in Experiments 1 and 2, showed that, oftentimes, teenagers erroneously used the connective *aussi* ‘therefore’ instead of the connective *en outre* ‘moreover’ probably because *aussi* can also be used to signal an additive relation. The additive meaning of *aussi* is also by far its most frequent meaning in language use. Hence, teenagers may have followed the probabilistic approach to connective interpretation (Asr & Demberg, 2020) and inferred the more frequent additive function of *aussi*, despite the syntactic constraints of *aussi* that can be used with an additive meaning only in sentence-medial or final position (Roze et al., 2012). It may therefore be important to rule out the effect of polyfunctionality in the next experiment, by testing strictly monofunctional connectives. In fact, previous research shows that it may be challenging even for adults to judge the appropriate uses of polyfunctional connectives, especially when these connectives are used in their infrequent functions (e.g., Zufferey et al., 2015).

Finally, it could be also noteworthy to explore other factors of inter-individual variation in the competence with connectives among teenagers, since out of several measures of exposure to print, none consistently predicted the performance in the connective task across all groups of young speakers. In contrast to the group of adults, whose performance with connectives was predicted by the ART-F, at least in the production tasks.

4. The Mastery of Monofunctional Connectives in French: The Role of Vocabulary Size and Exposure to Print[1]

This chapter continues to explore factors affecting the development of the mastery of discourse connectives during teenage years. As shown in Chapter 3, the frequency of written connectives in corpora is one of the major predictors of the correct use of connectives in a cloze task. However, it is possible that this finding was partially biased by the polyfunctional character of the connective *aussi*. This connective can signal both a consequence relation, which was targeted in previous experiments, and an additive relation, depending on its position in the sentence. Perhaps, being familiar with the more frequent additive function of *aussi*, participants might have mistakenly chosen it instead of the targeted additive connective *en outre*. As a result, the average score for *en outre*

[1] Results from this study are partly published in Tskhovrebova, Zufferey and Tribushinina (2022).

probably was lower than what it could have been without this bias. In this chapter, I aim to overcome the eventual issue of polyfunctionality by testing exclusively monofunctional French connectives.

Moreover, in order to obtain more robust results, a greater number of connectives is included in the present study. The monofunctional connectives, which are assessed in this chapter, signal six different coherence relations: addition, cause, concession, consequence, contrast, and temporality. Each coherence relation is tested using two different connectives that differ in terms of the mode in which they are used. One connective is mostly used in writing and another one is more commonly used in oral speech. The goal of testing connectives from different modes is to unveil whether, besides connective frequency, the linguistic mode in which they are typically used is an important predictor of their mastery. In contrast to oral speech, to which children are exposed since their birth, exposure to written language happens later. It is mostly through the process of schooling that a child starts to be exposed to writing on a regular basis. What is more, a really extensive exposure begins in secondary school when students become fully independent readers of various text genres (Nippold, 2004, 2008). Therefore, I hypothesise that young speakers should on average use written connectives less accurately than oral ones due to differences in the amount of exposure they have had to each mode.

Another line of investigation that I pursue in this chapter concerns the predictors of inter-individual variation in the use of connectives during teenage years. As shown in Chapter 3, the competence with connectives varies according to teenagers' academic background, as reflected by different tracks of school curriculum, and improves between the younger group of teenagers from secondary school (M_{age}=14) and those from high school (M_{age}=16). However, the role of exposure to print for the development of connectives' mastery cannot be established based on this data. The degree of exposure to print, as measured by the ART-F, was the best predictor of an appropriate use of connectives in the cloze test by French-speaking adults. In contrast, it was the subjective evaluation of exposure to print that partially predicted the accuracy of responses on the cloze test by teenagers. The TRT-F, the version of the ART adapted for teenagers, did not predict the use of connectives at all.

These findings may suggest that the tested measure of exposure to written language was simply not well suited for younger speakers. Perhaps, teenagers do not read much outside of their school curriculum or read books and texts of other genres than those included in the test. As a result, this measure has failed to capture their reading habits and hence, to predict the variation in their performance on the connective task. In order to overcome the potential issues with the TRT-F, in this study, I develop a new version of the ART, following a different logic. Instead of including contemporary books, which are popular among teenagers living in the French-speaking part of Switzerland, like it was the case for the TRT-F, I base the new ART on the names of classical authors. This way, the test items should be more familiar to young speakers and better capture their print-related activities, as the names of this type of authors are probably more wide-spread across in- and out-of-school activities in which teenagers engage. I suggest that speakers with a greater exposure to print, as measured by the new version of the ART, should also be better at using connectives, as it is mostly through the exposure to the written language and literacy activities that the biggest variety of connectives can be accessed.

In addition to exposure to print, there may also be another potentially important source of individual differences in the mastery of connectives. Vocabulary is something that actively develops over teenage years and that is variable between speakers with different personal experiences and social backgrounds (Nation & Coxhead, 2021). Moreover, a number of theories about the linguistic nature of connectives concur on the intermediate position of connectives between conceptual and procedural knowledge (see, e.g., Ducrot, 1972; Moeschler, 2002; Wilson, 2011; see Chapter 1 for a more detailed account of these theories). Indeed, on the one hand, connectives represent a specific domain of the lexicon (Crosson & Lesaux, 2013) and can be considered as part of what Ullman (2001) call declarative knowledge. On the other hand, they provide speakers with processing instructions on how they should relate segments of discourse and which coherence relation to expect in the following clause (see, e.g., Britton, 1994; Gernsbacher, 1997; Koehne & Demberg, 2013; Rohde & Horton, 2014; Sanders & Spooren, 2007; van Silfhout et al, 2015; Xiang & Kuperberg, 2015).

Therefore, assessing whether the width of general vocabulary in young speakers predicts the appropriate usage of connectives could contribute to the theories exploring the nature of connectives. In line with research on adult native and non-native speakers of French (Wetzel et al., 2020), showing that lexicon size predicts the performance with connectives in a sentence cloze test, I hypothesise that a broader vocabulary should contribute to a better knowledge of connectives also in native French-speaking teenagers. However, since connectives have a particular status between processing instructions and conceptual items, vocabulary size should not be the only predictor of their appropriate use.

The final goal of this chapter is to assess the role of age compared to vocabulary level and exposure to print in the development of the competence with discourse connectives. The results of the previous chapter hinted that, at least for the older group of teenagers aged 16, the factor of age is less important than speakers' academic background. This finding stands in opposition to what was observed for primary school children, namely that age was reported to be one of the main predictors of connective usage and comprehension (see, e.g., Spooren & Sanders, 2008; Van Veen et al., 2009; Evers-Vermeul & Sanders, 2011). The reason for such difference between childhood and teenage years may be related to cognitive processes, explained in Ullman's (2001) declarative-procedural model of language. According to this model, declarative memory becomes better with age, and as a result, older children are also better at learning new vocabulary. However, the ability for procedural learning decreases with age, and since connectives are involved in procedural knowledge, age may play a less prominent role in teenage years, when procedural learning becomes more challenging. Hence, I suggest that, in this study, the role of biological age should be less important that vocabulary size and exposure to print for the appropriate use of connectives.

4.1. Method

All materials, data, and code are available on the OSF repository[2].

4.1.1. Participants

The participants of this study were 154 French-speaking teenagers aged 12 to 19[3] (*M*=14.43, SD=1.8). All the participants were typically-developing native speakers, as

[2] https://osf.io/gqpy7/?view_only=1f466fe2e9cc42aba5b53b721d9475ba

[3] The results of two participants who were 19 years old are presented together with the group categorized as teenagers because they were recruited together with other students of the first year of high school and followed the same curriculum in French as their classmates. I did not consider it appropriate to present the results of these two participants, who were not more advanced than their younger classmates, together with the group categorized as adults, who had already finished their school studies and were recruited in a different context, namely via a crowdsourcing platform.

confirmed by their language teachers. The experiment was held in nine classes of four schools in the French-speaking part of Switzerland. Participants came from the 9th (n=53, M_{age}= 12.57, SD=0.54), 10th (n=26, M_{age}=13.73, SD=0.78), and 11th (n=14, M_{age}=14.79, SD=0.80) years of secondary school, and the first year of high school (n=61, M_{age}=16.26, SD=0.95). All schools gave their informed consent for participation in the study. A group of adults was also tested to determine the baseline of competence with connectives. For this purpose, I recruited 52 French speakers (M_{age}=30.75, SD=11.07, Range 19-58) via the crowdsourcing platform Prolific© (Prolific, Oxford, UK, www.prolific.co). All participants showed at least 90% of good ratings in previous studies on the platform and gave their informed consent for participation in the study.

4.1.2. Materials

4.1.2.1. Sentence cloze test

Selection of connectives. Six types of coherence relations (addition, contrast, temporality, consequence, cause, and concession) common in corpus data (see, e.g., Prasad et al., 2008) were tested in this study. Each coherence relation was represented by one connective more typical for written language and another one more typical for oral speech (see Table 1 for the distribution of different types of connectives).

The distinction between oral and written connectives was based on a corpus analysis of connective frequencies in corpora of oral and written language, and then was also confirmed in a judgement task performed by native speakers. The connectives' frequency in oral speech was calculated in the oral sub-corpus of French *Orféo* (Benzitoun et al., 2016), as it includes 4 million words and contains speech from a wide variety of genres, such as everyday conversation and public speech. The frequency of connectives in writing was calculated based on three corpora from different genres, including journalistic (Le Monde corpus), argumentative (the French part of the Europarl corpus, Koehn, 2005), and literary texts (the Frantext corpus, ATILF, 1998-2022). I first calculated the connective frequencies per million words separately for each corpus, and then calculated the mean frequency for each connective. Connectives that were more frequently used in oral than in written corpora were categorized as *oral*; and connectives that were more frequent in written than in oral corpora were categorized as *written*.

To verify the outcomes of the corpus analysis, I asked a group of adults to judge to what extent each of the connectives chosen for the task was common in oral conversation in informal contexts (such as family dinner or a conversation with friends), on a scale from 0 to 20. The answer 0 meant that a connective is never used in informal oral speech and 20 that it is used very frequently in this context. For every coherence relation, the connective with a higher score was labelled as *oral* and that with a lower score as *written*. The judgement task was performed online by native French speakers (*N*=102). None of them participated in the main experiment. The distinction between the connectives typically used in oral and written modes, as determined by native speakers' judgements, matched the outcome of the corpus analysis of connectives' frequencies.

Table 1. Distribution of connectives per type of coherence relation and mode with their mean subjective orality rate (M_{OR}) and frequency (per million words) in oral (Freq OR) and written (Freq WR) corpora

Relation	Mode	Connective	Translation in English	M_{OR} (SD)	Freq OR	Freq WR
Addition	oral	*en plus*	in addition (less formal)	17.57 (2.97)	501.5	57.09
	written	*en outre*	in addition (more formal)	3.63 (3.80)	1.75	91.29
Cause	oral	*parce que*	because (less formal)	18.04 (2.71)	4086.75	262.42
	written	*car*	because (more formal)	11.75 (5.44)	68	552.85
Concession	oral	*même si*	even if	16.08 (3.06)	165.5	101.76
	written	*néanmoins*	nevertheless	5.43 (4.08)	7.25	86.73
Consequence	oral	*donc*	so	17.27 (3.04)	5913	723.83
	written	*ainsi*	therefore	8.14 (5.15)	64.5	292.72
Contrast	oral	*par contre*	instead	16.25 (3.29)	251.75	15.47
	written	*en revanche*	conversely	9.10 (4.31)	6.5	46.92
Temporality	oral	*dès que*	as soon as (less formal)	14.61 (3.97)	87	52.23
	written	*aussitôt que*	as soon as (more formal)	7.96 (5.03)	0.5	7.44

Structure of the test. I asked participants to fill in gaps between two sentences with an appropriate connective. The gap was always in the initial position of the second sentence. The test included 60 items in total, 10 items per coherence relation, five of which targeted a written connective and other five oral ones. In the task, participants always had a choice between four options randomly selected out of six connectives tested in each mode. Consequently, if the expected answer was a written connective, the proposed options also belonged to this mode, and vice versa for the oral connectives (see examples for the relation of consequence (1)–(2)). This allowed to test the two modes separately and prevented participants from always choosing the oral connectives which are more common in everyday speech. For each experimental item, there was only one possible answer. The final score was calculated as the proportion of correct answers per connective.

(1) Correct answer: written connective *ainsi* ‘therefore’

Jeanne a oublié de mettre son réveil //________// elle est arrivée en retard aux cours et s'est fait punir.

‘Jeanne forgot to set her alarm clock //________// she was late for class and got punished.’

Answer options: (a) *néanmoins* ‘however’; (b) *en revanche* ‘in contast’; (c) *aussitôt que* ‘as soon as’; (d) *ainsi* ‘therefore’

(2) Correct answer: oral connective *donc* ‘so’

Le politicien a été accusé de corruption durant sa campagne //________// il n'a pas été élu.

‘The politician was accused of corruption during his campaign //________// he was not elected.’

Answer options: (a) *même si* ‘even though’; (b) *mais* ‘but’; (c) *dès que* ‘as soon as’; (d) *donc* ‘so’

4.1.2.2. Vocabulary level test

To assess the vocabulary level of participants, I created a French version of a vocabulary size test based on Nation and Beglar (2007). The participants were asked to read a definition of a word and choose one of the six words that was the best match for the definition. The test included four categories of words, based on frequency lists from the French corpus *Lexique 3.83* (New et al., 2001). Each category consisted of 30 items, which were selected from the first, second, third and fourth 5000-word families. Moreover, each word category included different parts of speech, namely, 18 nouns, 6 verbs, and 6 adjectives. Importantly, the foils also belonged to the same frequency level as the target words. The word frequencies therefore decreased from the first to the fourth category, and the participants completed the task in the order of increasing frequencies. Vocabulary scores used the proportion of correct answers per participant. The reliability of the vocabulary test, as measured by Cronbach's alpha, was high both for teenagers and adults. For teenagers, it was of .96 (95% CI [.93–.97]), and for adults, .91 (95% CI [.85–.93]).

4.1.2.3. Author recognition tests

I developed a new version of the author recognition test (ART) to assess teenagers' degree of exposure to print, as this test is not only sensitive to cultural differences (e.g., Stainthorp, 1997) but also to the age of participants (e.g., Cunningham & Stanovich, 1990). The ART (*ART-F-CL*), used in the present study, was based on the names of authors who are considered to be classics according to the listings of three big national chains of bookstores in Switzerland. The list included 40 author names and 40 names of unknown people, which were randomly mixed. The participants had to select only those names that they knew to be authors. The instruction mentioned that some of the names were not authors, and that one point would be removed if the participants checked the wrong name. For each correct answer, participants were given *1* point, and for each wrong one *-1*. I computed the general score summing up the points for correct and incorrect answers. The maximum possible score was *40* and the minimum *-40*.

For the group of adults, I used a different version of the ART (*ART-F*), developed for French by Zufferey and Gygax (2020a). It replicated the design of the original English ART (Stanovich & West, 1989) and was based on the names of best-selling and prize-winning authors (see https://osf.io/yxj8q/ for the full task). The number of items and the calculation of the final score was the same as for the teenage version of the task described before. The reliability of the two ART tests was quite high, as indicated by their

Cronbach's alphas (ART-F-CL: .88, 95% CI [.85–.91]; ART-F: .92, 95% CI [.86–.94]). In addition to the ART, all the participants were asked to give a subjective evaluation of their exposure to print. In a separate question, they were asked to estimate how regularly they read on a scale ranging from 0=*never* to 10=*every day*.

4.1.3. Procedure

All the tasks were administered online via a weblink. The link was distributed directly among the teachers of the participating classes in the case of teenagers, and via the Prolific platform (https://www.prolific.co) in the case of adults. The order of the tasks was always the same. The participants started with the connective choice task, then proceeded to the ART and finished with the vocabulary test. Once the participants gave an answer and proceeded to the next question, they could not go back and correct their initial response. There was no time limit for the task, but the participants had to finish it in one session. The teenagers spent on average one hour to complete all the tasks, and about 40 minutes for adults.

4.1.4. Analysis

I analysed the accuracy of responses (*1*=right, *0*=wrong) in the cloze test using a generalised mixed-effects logistic regression model in the statistical software R (R Core Team, 2020). I first analysed the results of all participants together in order to verify whether teenagers overall had a different performance in the connective task compared to adults. Afterwards, I made two separate analyses for teenagers and adults to assess the role of inter-individual variation within each group. I did not include the measures of individual difference in the global analysis for all participants, as different versions of the ART were fulfilled by teenagers and adults. The significance level was set at 5%.

I performed an automated backward selection of variables, as this way I could include all the tested predictors in the initial model and then automatically eliminate the nonsignificant ones. The initial full models were built with the *glmer* function of the *lme4* package (Bates et al., 2015). The full model for all participants included Group (Adults versus Teenagers), Coherence Relation, connective Mode (Oral versus Written).

The initial full model for teenagers included Vocabulary Size, the ART-F-CL, Subjective Evaluation of Exposure to Print, Age, and connective Mode as predictors of performance on the connective task. All the variables of individual difference were centred. Since ART-F-CL was highly correlated with the vocabulary score (*rho*=.50 [.37, .61], p<.001) and age (*rho*=.49 [.36, .61], p<.001), age and vocabulary score were residualized by the ART-F-CL score by means of the *umx_residualize* function of the *umx* package (Bates, 2021) to avoid multicollinearity in the statistical model. The initial full model for adults included the same predictors as the model for teenagers, except for the ART-F-CL that was substituted by the ART-F. The same centring procedure was applied to the measures of inter-individual variation; while only vocabulary score was residualized by the ART-F because of an important correlation (*rho*=.57 [.34, .73], p<.001). In all the three analyses treatment contrasts were applied for the unordered factors. The Cause was set as reference level for comparing the scores associated to different coherence relations, as speakers have a cognitive bias towards causality in the absence of other explicit cues (Sanders, 2005) and should not have a particular difficulty with this relation (see, e.g., Cain & Nash, 2011). For the factor of Group, Adults were chosen as reference, as this group was assumed to include speakers with the highest level of competence. Oral Mode was set as reference to compare the two modalities, since speakers start to be exposed to oral speech much earlier than to written language and should master it better.

Next, I conducted an automated selection of relevant predictors with *drop1* function of the *stats* package (R Core Team, 2020), deleting the fixed effects with the *p* values higher than 0.05. When several factors were selected among the relevant predictors for the reduced model, I checked whether adding an interaction between these factors improved the final model's fit. Comparison between the models without and with an interaction was done with the *anova* function of the *stats* package. The outcome of the final model was then returned with the *summary* function of the *lmerTest* package (Kuznetsova et al., 2017). Following the procedure by Schreiber-Gregory (2018), I controlled that the assumptions of logistic regressions were met (i.e., appropriate outcome structure, absence of multicollinearity, linearity of independent variables and log odds, and an appropriate sample size). Since the experiment had a repeated measures design (in that the same participants completed multiple test items, and the same test items were taken by multiple participants), the assumption of observation independence was not met. I however accounted for it by adding the random effects as intercepts for items and participants in the mixed-effects models.

Finally, in the separate models for teenagers and adults, I performed a random forest analysis (Strobl et al., 2009) based on the predictors included in the final reduced model in order to compare the impact of each relevant predictor variable on the dependent one (i.e., correctness of responses in the cloze task). The advantage of this method is that it does not have assumptions about the distribution of data and can make predictions even about highly correlated variables. Moreover, it is highly reliable, as variable importance is calculated based on a multitude of classification, or regression, trees (Strobl et al., 2009).

4.2. Results

4.2.1. Descriptive statistics for the background measures

As is evident from the descriptive statistics in Table 2, across all three measures, teenagers had on average lower scores than adults. The vocabulary level of teenagers was about 22% lower than that of adults. The ART scores were 2.2 points lower in teenagers than in adults. Finally, the subjective evaluation of exposure to print was about 1 point lower for teenagers than for adults.

Table 2. Descriptive statistics for the background measures by group

	M (SD)	**Observed range**	**Possible range**
		Vocabulary size	
Teenagers	.71 (.15)	.15-.95	0–1
Adults	.91 (.07)	.72-.99	
		Author recognition test*	
Teenagers	6.55 (5.93)	-11-28	-40–40
Adults	8.75 (7.63)	-1-33	
		Subjective exposure to print	
Teenagers	4.91 (3.10)	1-10	0–10
Adults	6.30 (2.28)	1-10	

*A different version of the ART was used for teenagers and adults

4.2.2. Performance in the connective test

4.2.2.1. All participants

Based on the step-down selection of predictors, the final model for the global analysis included fixed effects of Group (Adults versus Teenagers), Coherence Relation, and Mode (Oral versus Written), as well as an interaction between Group and Mode, and Item and Participant as random intercepts (see Table 3 for the model's estimates). Adding

Connective as random slope by Participant did not allow the model to converge, and including an interaction between Coherence Relation and Mode did not improve the model's fit (Δx^2=5.44, Δdf=5, p=0.364). Finally, I included in the analysis the categorical variable Group instead of the continuous Age because I wanted to contrast an overall performance of teenagers to that of the control group of adults.

The results revealed that teenagers performed on average quite well in the connective insertion task (see Figure 1). However, their scores across all connectives were significantly lower than those of adults, especially for the written connectives, as suggested by a significant interaction effect between Group and Mode (see Table 3). This finding indicates that teenagers have not yet reached the adult level of competence with monofunctional connectives in general and with those used in the written mode in particular.

Teenagers received the lowest scores for the connective *en outre* 'in addition' (*M*=.58, 95% CI [.45, .70]), followed by *en plus* 'in addition' (*M*=.75, 95% CI [.64, .86]) and *ainsi* 'therefore' (*M*=.75, 95% CI [.64, .86]), then concessive *néanmoins* 'nevertheless' (*M*=.79, 95% CI [.68, .89]) and *même si* 'even if' (*M*=.80, 95% CI [.70, .90]), and finally, all the remaining connectives with an accuracy superior to .85. The group of adults also obtained the lowest score for the connective *en outre* (*M*=.87, 95% CI [.78, .95]), followed by *même si* (*M*=.89, 95% CI [.81, .97]). However, accuracy of responses for all the other connectives was above .90. Given these scores, I believe that the effect of coherence relation might have been slightly overestimated for the consequence relation, as only in the group of teenagers and only one out of two connectives received a lower score.

Figure 1. Distribution of mean scores per connective in sentence cloze task among teenagers and adults

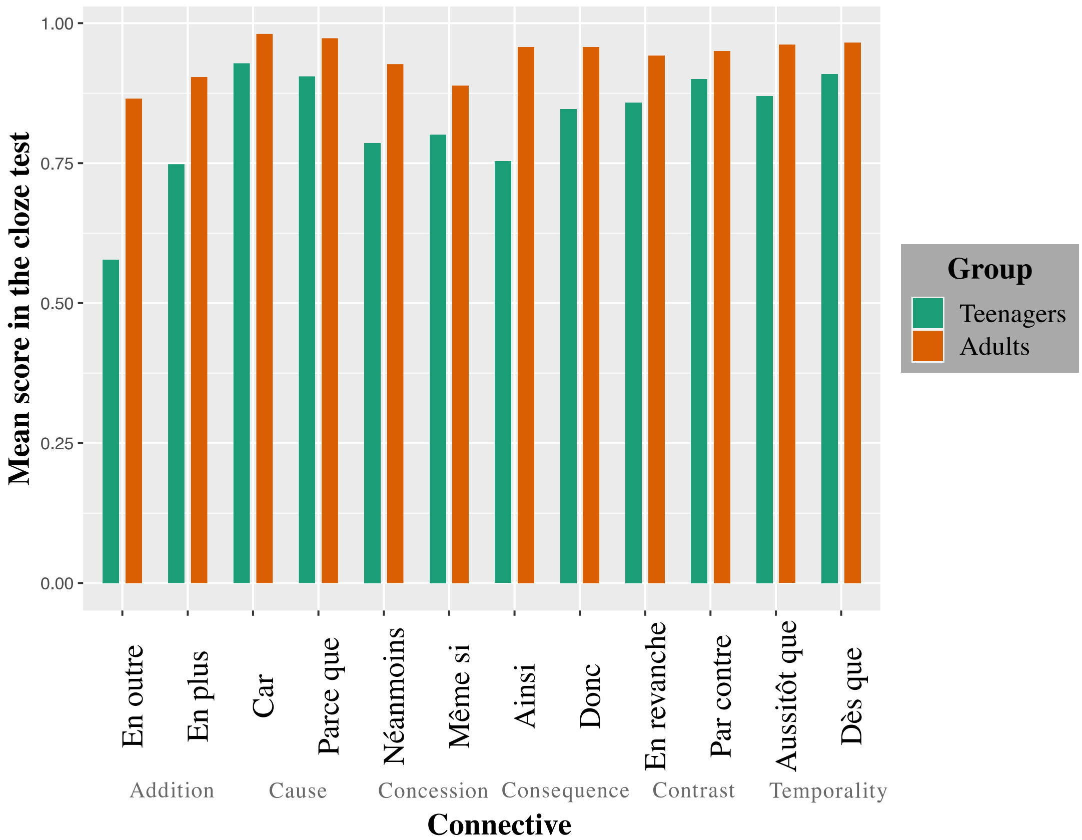


Table 3. Output of the full model and the final reduced model for all participants

Full model					Final reduced model				
Variable	**Estimate**	**SE**	**z**	**Pr(>|z|)**	**Variable**	**Estimate**	**SE**	**z**	**Pr(>|z|)**
All participants									
(Intercept)	4.48	0.29	15.42	<0.001	(Intercept)	4.26	0.30	14.25	**<0.001**
GROUP					GROUP				
Teenagers	-1.39	0.18	-7.92	<0.001	Teenagers	-1.13	0.19	-5.82	**<0.001**
MODE					MODE				
Written	-0.43	0.18	-2.36	0.018	Written	-0.01	0.23	-0.03	0.973
COHERENCE RELATION					COHERENCE RELATION				
Addition	-1.96	0.31	-6.24	<0.001	Addition	-1.96	0.31	-6.26	**<0.001**
Concession	-1.27	0.31	-4.04	<0.001	Concession	-1.27	0.31	-4.04	**<0.001**
Consequence	-1.14	0.31	-3.64	<0.001	Consequence	-1.15	0.31	-3.64	**<0.001**
Contrast	-0.44	0.32	-1.39	0.165	Contrast	-0.44	0.32	-1.39	0.164
Temporality	-0.34	0.32	-1.08	0.280	Temporality	-0.35	0.32	-1.08	0.279
					GROUP*MODE				
					Teenagers*Written	-0.47	0.17	-2.87	**0.004**

4.2.2.2. Teenagers

The final reduced model for teenagers based on the step-down selection of predictors included Vocabulary Level, the ART-F-CL, Age, and Mode (see Table 4). Adding interaction between mode and the ART-F-CL did not improve the model's fit (Δx^2=0.29, Δdf=1, p=.589). The estimates of the final model revealed that Vocabulary Level, the ART-F-CL, and Age were the most important predictors for the performance in the cloze test. Moreover, connectives mostly bound to writing tended to be slightly more challenging for teenagers than the ones used in speech, as demonstrated by an estimated decrease of 0.49±0.25 SE. The only discourse relation where this trend was not attested is causality, for which the written connective *car* had a very similar score (*M*=.93, 95% CI [.86, .99]) to that of the spoken connective *parce que* (*M*=.91, 95% CI [.83, .98]).

The overall prediction accuracy of the random forest analysis was 86%. This analysis supported the mixed logistic regression analysis and showed that vocabulary level had the most impact on the performance with connectives, followed by the score in ART-F-CL, age and to a lesser extent mode (see Figure 2 for the visualisation of the hierarchy of variable importance).

4.2.2.3. Adults

The variables selected for the final reduced model included Vocabulary Level and the ART-F (see Table 4). In contrast to the group of teenagers, Age and Mode were not significant predictors of the performance in the connective task within the group of adults. The prediction accuracy of the random forest analysis was 94%. Similar to the random forest analysis within teenagers, vocabulary level had the biggest impact on the performance with connectives (see Figure 2 for the visualisation of the hierarchy of variable importance).

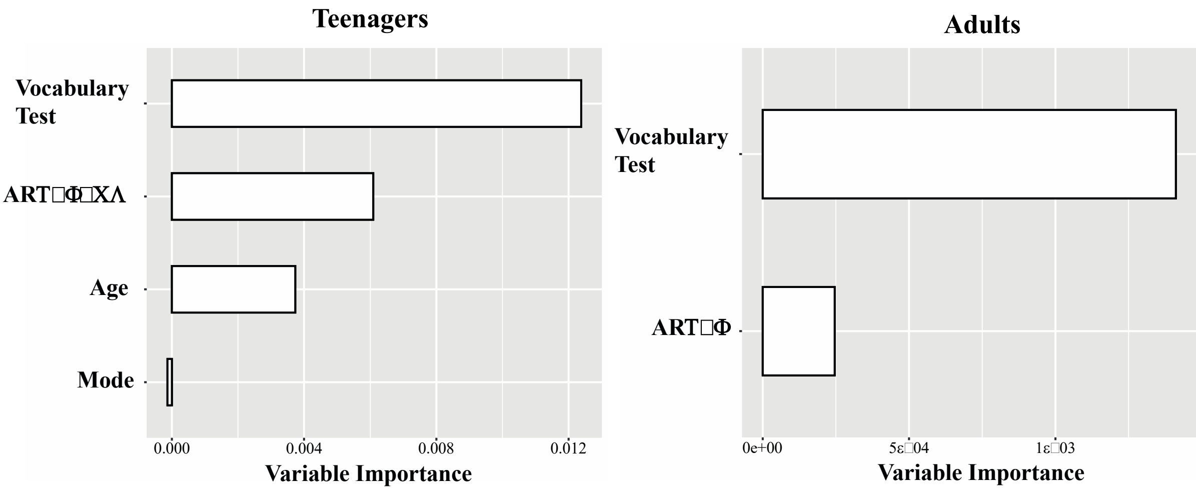
Table 4. Output of the full model and the final reduced model for the group of teenagers and adults

Variable	Full model				Final reduced model			
	Estimate	**SE**	**z**	**Pr(>|z|)**	**Estimate**	**SE**	**z**	**Pr(>|z|)**
Teenagers								
(Intercept)	-0.69	0.55	-1.26	0.209	-0.47	0.53	-0.89	0.376
Vocabulary test**	5.23	0.78	6.74	<0.001	5.37	0.77	6.93	**<0.001**
ART-F-CL*	0.94	0.17	5.44	<0.001	0.96	0.17	5.51	**<0.001**
Age**	1.86	0.64	2.92	0.004	1.63	0.62	2.63	**0.008**
Subjective Exposure to print*	0.17	0.12	1.40	0.161				
Written mode	-0.49	0.25	-1.96	0.050	-0.49	0.25	-1.96	**0.050**
Adults								
(Intercept)	1.45	1.47	0.99	0.324	2.33	0.33	7.15	**<0.001**
Vocabulary test**	14.22	3.04	4.68	<0.001	14.16	2.77	5.11	**<0.001**
ART-F*	0.49	0.16	3.09	0.002	0.51	0.14	3.72	**<0.001**
Age*	0.21	0.38	0.56	0.576				
Subjective Exposure to print*	0.12	0.34	0.34	0.733				
Written mode	-0.04	0.32	-0.11	0.912				

*Centred values

**Centred and residualized values

Figure 2. The impact of each predictor variable on the dependent variable according to the random forest analysis for teenagers and adults



4.3. Discussion

4.3.1. Competence with monofunctional connectives

The goal of this chapter was to provide new evidence on the level of competence with monofunctional connectives by French speaking teenagers, and to assess how factors related to the linguistic properties of connectives and inter-individual differences could explain variability in the level of connective mastery. The results show that, on average, teenagers have a good command of monofunctional connectives signalling different coherence relations, and used both in oral and written modes.

Although the overall performance with connectives was high, one connective represented a particular difficulty for all the participants. Both teenagers and adults had the lowest accuracy score for the additive connective *en outre* ‘in addition (more formal)’. The fact that another additive connective *en plus* ‘in addition (less formal)’ was mastered much better suggests that the difficulty with *en outre* is not related to the complexity of the additive coherence relation, but rather to particular characteristics of this connective. As shown in the preliminary questionnaire assessing the degree of orality of the connectives included in this study, *en outre* received the lowest orality scores (Table 1). In other words, native speakers may consider it atypical for oral speech and do not have a clear intuition about its usage. The reason for that may come from the fact that they had not been exposed enough to the written, formal, administrative contexts where this connective is used.

The general statistical analysis also showed that connectives encoding additive, concessive, and consequence relations were on average more challenging to use in a cloze test than other types of connectives. However, this finding should be considered with caution for several reasons. First of all, this result is difficult to explain in terms of the CCR model, as the mentioned relations represent three different levels of complexity varying from the simplest additive relation to the most complex concessive relation. Therefore, the potential difficulty of connectives, signalling these relations, cannot be associated with their cognitive complexity. Secondly, drawing generalisations about the difficulty of coherence relations based on only two connectives per coherence type can be misleading. This is particularly the case when considering the consequence relation, as only one of the two consequence connectives, namely *ainsi* ‘therefore’, received a lower score. Finally, although some connectives received lower scores than others, on average, these scores were not low. Except for the additive connective *en outre*, which was probably the most challenging and received a score of .58 among the group of teenagers and .87 among adults, the accuracy scores of other additive, concessive, and consequence connectives were above .75 for teenagers and above .89 for adults.

As for the role of the mode in which connectives are typically used, the findings diverged between teenagers and adults. For teenagers, the connectives typically used in writing were slightly more challenging than the ones used in speech. However, this was not the case for adults who, on average, used equally well both types of connectives. The difference between teenagers' performance with oral and written connectives might stem from the overall higher frequency of oral connectives. However, this logic does not apply to all the connectives tested in this chapter. The written connectives *en outre*, *car*, and *en revanche* are infrequent in spoken corpora, but in the written ones, they have higher frequency that their oral counterparts (see Table 1 for frequencies). Hence, at least for these connectives, the difference in scores may stem from the fact that teenagers had not been exposed enough to the written contexts, in which these connectives are used more frequently, while adults probably have had enough exposure to this mode throughout their lifespan. In addition, these findings may also suggest that an appropriate use of written connectives is not acquired and developed in the same way as the use of oral connectives. Since exposure to the connectives mostly bound to the written language comes exclusively through reading, getting access to such connectives may require more effort. Therefore, it is important that school curricula devote more time and resources to teach this type of connectives as part of written language competence.

An overall performance of teenagers with discourse connectives was good, but still inferior to that of adults, even though I tested only monofunctional connectives in a relatively simple experimental context of isolated sentence pairs. It is likely that, within a more ecological text cloze test, the difference in the performance between adults and teenagers would be even greater, as it was observed in Chapter 3. This finding is in line with previous research on language development in older children (Berman, 2004; Nippold, 2008) suggesting that adult-level language proficiency is acquired far beyond puberty and that proficiency with connectives continues to develop even after the high-school years, as high-school students are not yet similar to adults in their ability to use appropriate connectives.

4.3.2. Individual variation in the mastery of connectives

For both groups of speakers, vocabulary level was the strongest predictor of the appropriate usage of connectives in French. Although for teenagers, age was also selected among relevant predictors of the performance in the cloze test, it was less important than lexicon size. To put it differently, a higher vocabulary level still significantly contributes to a better usage of monofunctional connectives both in teenage years and in adulthood, even though connectives may differ from common lexical items in that they encode both concepts and procedures (see, e.g., Wilson, 2011). This finding does not necessarily contradict the idea that connectives function as processing instructions. It rather gives evidence for their intermediate nature as specific lexical items, expressing procedural meaning.

The degree of exposure to written language, as measured by the ART-F and the ART-F-CL, is another prominent factor predicting variation in the competence with connectives by teenagers and adults. This finding corroborates previous research on this topic in adults (Zufferey & Gygax, 2020a) and extends its validity on younger cohorts. The relation between exposure to print and the competence with connectives indicates that long-term reading habits, as revealed by the ART (Scholman et al., 2020), may foster the acquisition of linguistic experience, required for an accurate use of connectives in discourse. The notion of linguistic experience includes a number of linguistic skills, such as metacognitive analysis of texts (McBride-Chang & Chang, 1995), vocabulary knowledge (see, e.g., Stanovich et al., 1995), and reading comprehension (see, e.g., Spear-Swerling et al., 2010). The particularity of the ART test is that its score is related to this complex set of competences, but cannot be reduced to any single one of them.

In contrast to previously developed ARTs, based on popular personal readings among a specific age group of speakers from a specific region (see, e.g., Spear-Swerling et al., 2010), the novel ART-F-CL, based on classical authors, is not really attached to a specific geographical area, is more polyvalent and easier to develop. Since a list of classic literature can be found in school curriculum guidelines and catalogues of big bookstore chains, the ART-CL can be developed without launching a preliminary study on reading preferences among teenagers of a certain linguistic and geographical zone. This makes this test much handier and faster to implement, but presumably no less efficient in capturing individual variation in exposure to print. However, future research should provide further validation of the new version of the ART.

In contrast, the subjective evaluation of exposure to print was reported to be less adequate than the ART tests, as it did not predict all the variation of connective uses. This finding may be due to the fact that this measure of exposure to print is often subject to guessing and to the production of socially desirable answers (see, e.g., Chateau & Jared, 2000; Echols et al., 1996). Hence, it might indicate participants' attitude toward reading rather than their degree of exposure to print (see, e.g., Allen et al., 1992).

Separate analyses conducted with teenagers and adults to assess their mastery of connectives demonstrated that age was an important predictor for the correct usage of connectives in the cloze test for the group of teenagers, but not for adults. This indicates that increasing cognitive maturation and linguistic experience, related to age, are still essential for the proper mastery of connectives in teenage years, even though age is less important than vocabulary level and degree of exposure to print, as measured by the ART-F-CL. This finding shows that later language development, such as the developing ability to use a broad range of connectives, differs from early language acquisition and happens in a slower and qualitatively different way than in early childhood (Nippold, 1993). Furthermore, this result corroborates the initial hypothesis, based on the Ullman's (2001) declarative-procedural model of language, and suggests that connectives are part of procedural knowledge, as age becomes less relevant for the mastery of connectives during teenage years when procedural learning abates. Therefore, the fact that chronological age is gradually becoming less important for the performance with connectives reveals the intermediate position of teenage years between childhood, when it is one of the most prominent predictors of this performance, and adulthood, when its affect vanishes.

4.4. Conclusion

The study, reported in this chapter, contributes to the research on linguistic development during teenage years. It highlights that by the end of high school, French-speaking teenagers have still not attained an adult level mastery of connectives from written and oral modes, even though all the connectives were strictly monofunctional and were tested in a relatively simple experimental context. This finding suggests that linguistic proficiency continues to develop far beyond puberty and during late teenage years.

The accuracy of responses on the connective cloze test of the present study was mainly predicted by lexicon size and degree of exposure to print, and only to a lesser extent by students' age and the mode in which connectives are typically used. This finding highlights the need to increase students' exposure to print during teenage years, as greater exposure to written language should provide students with real examples of connectives, used as processing instructions. Moreover, more frequent exposure to print should also contribute to expanding teenagers' lexicon, which represents another strong predictor of competence with connectives, on top of students' exposure to print, as evidenced by the results from this chapter.

In order to support the claim about linguistic development in teenage years and to make it more generalisable, I will replicate in Chapter 5 the study on French monofunctional connectives in Russian. Replicating the current study in a different language will allow to verify whether the found effect of coherence relations types could be truly attributed to the particularities of these relations, or it was the specific connectives chosen to represent those relations that were slightly more challenging to use. By choosing Russian, I do not only aim to verify the generalisability of the results obtained in French, but also to contribute to the research on connectives in a language that has not been studied much in this context and will include the factor of polyfunctionality as an additional predictor of the connective mastery.

Lastly, the fact that connectives like *en outre* are challenging for teenagers and adults indicates that certain connectives still represent particular difficulty even for adult native speakers, and should be studied more in depth. Therefore, in Chapter 7, I assess the use of more infrequent connectives that are bound to the written mode, as well as the use of connectives that signal less common and conventional coherence relations unlike those presented in this chapter. Furthermore, I also examine whether it is possible to enhance the development of the mastery of connectives through training or it is mostly through a long-term exposure to print and the contexts, in which connectives are used, that processing instructions can be acquired.

5. The Mastery of Mono- and Polyfunctional Connectives in Russian: The Role of Vocabulary Size and Exposure to Print[1]

The study presented in Chapter 4 demonstrated that French-speaking teenagers aged from 12 to 19 had a good command of 12 monofunctional connectives, expressing the relations of addition, concession, contrast, temporality, cause, and consequence, and typically used in oral or written language. It also showed that the mastery of connectives was predicted by teenagers' general vocabulary size and level of exposure to print, rather than by their chronological age. This finding is important, as it shows that language development during teenage years, involving the developing ability to use a wide variety of connectives, appears to be qualitatively different from the earlier stages of language

[1] Results from this study have been submitted for publication in Tskhovrebova, Zufferey and Tribushinina (2023, submitted).

acquisition (Nippold, 1993). In fact, in childhood years, age constitutes one of the major predictors of the acquisition of connectives (see, e.g., Blything et al., 2015).

Yet, these results do not automatically imply that the effect found in French can be generalized to other languages that have different inventories of connectives, and to other educational cultures. The study reported in this chapter will extend this line of research by examining the predictors of connective mastery in Russian-speaking teenagers, as only few studies have thus far examined the acquisition of discourse connectives in Russian (see for a few examples Knjazev, 2007; Mak et al., 2020; Tribushinina, Mak et al., 2017; Tribushinina, Dubinkina & Sanders, 2015; Tribushinina, Valcheva & Gagarina, 2017).

The study in Chapter 4 focused on the use of monofunctional connectives in order to overcome the design bias related to alternative functions of polyfunctional connectives, like it was the case in Chapter 3 (see Section 3.5). However, it is not completely clear whether the teenagers' performance on the cloze test was quite high because monofunctional connectives are generally easier than polyfunctional ones or because of the removed design bias. In other words, it is not clear whether polyfunctionality per se, – i.e., the existence of other functions, – is a factor of difficulty in the absence of the design bias. To examine the role of polyfunctionality more in depth, in the present chapter, I assess the use of both mono- and polyfunctional connectives, signalling the same type of coherence relations as in Chapter 4 and typically used in the written and oral modes. However, in order to address the design bias associated with the polyfunctionality of certain connectives and to avoid the possibility of having several correct answers, I ensure that among answer options there are only connectives with non-competing functions. On the one hand, considering that the task design controls for competing alternative functions of polyfunctional connectives and that only their dominant functions are targeted in the current study, the use of connectives in the sentence cloze task may not be more difficult than the use of monofunctional connectives. On the other hand, if polyfunctional connectives are revealed to be more challenging than monofunctional ones, it would mean that that polyfunctionality in itself is a factor affecting the use of connectives in teenage years.

Similar to Chapter 4, I intend to explore also other connective- and speaker-related factors that may account for the differences in the use of discourse connectives. For instance, I aim to study whether the connective mode (written or oral) predicts teenagers' performance in the cloze task also in the Russian language. Bearing in mind the results for French, I hypothesize that Russian-speaking teenagers should probably have more difficulties in using written connectives than oral ones, as massive exposure to written language comes later than exposure to oral language. In addition, by examining the same coherence relations types as in Chapter 4, I intend to validate whether the observed impact of coherence relation types can genuinely be attributed to the characteristics inherent to these relations, or if it was simply the specific connectives chosen to represent those relations that posed a slightly greater difficulty in their usage.

To assess whether the use of connectives by Russian-speaking teenagers is modulated by their individual differences in a similar manner to French-speaking teenagers, the same two background measures of individual differences were examined, namely degree of exposure to print and general vocabulary knowledge. In my third hypothesis, I predict that teenagers who have a larger vocabulary and those who are more exposed to print are more likely to use discourse connectives accurately, as connectives constitute a specific part of the lexicon (Crosson & Lesaux, 2013) and it is mostly through exposure to the written language that the widest variety of connectives can be acquired (see, e.g., Crible & Cuenca, 2017).

Finally, previous studies provide contrasting results on the role of age for the development of the competence with connectives. For primary school children, age has been consistently found to predict better comprehension and usage of connectives, as older children typically perform better than younger ones (see, e.g., Blything et al., 2015; Cain & Nash, 2011; Pyykkönen & Järvikivi, 2012). In contrast, the existing evidence on teenage years suggests that academic background is a stronger predictor of connective use in a sentence cloze task (Tskhovrebova, Zufferey & Gygax, 2022). Therefore, there is a need to assess the role of age for the mastery of connectives also by Russian-speaking teenagers, especially in comparison to other measures of linguistic competence, namely vocabulary level and exposure to print. My fourth hypothesis is that, similarly to previous studies on teenage years (Tskhovrebova, Zufferey & Gygax, 2022; Tskhovrebova, Zufferey & Tribushinina, 2022), the factor of age should play a less important role for the mastery of connectives in comparison to individual differences in linguistic experience, as reflected by vocabulary level and exposure to print.

5.1. Method

5.1.1. Participants

One hundred and twenty-three native speakers of Russian, aged 11 to 17 (M_{age}=13.68, SD=1.87), participated in this study. Their native-level competence in Russian as well as absence of language disorders were validated by their teachers of Russian. The experiment was carried out in seven schools in Saint-Petersburg, Russia, and included classes from the 5th to the 11th grade. A group of adults (N=51, M_{age}=33.37, SD=8.08, Range 19–52) was also recruited for the experiment via the crowdsourcing platform Prolific© (Prolific, Oxford, UK, www.prolific.co) in order to establish the baseline of performance. All teachers responsible for the classes of teenagers, as well as adult participants, gave their informed consent for taking part in the experiment.

5.1.2. Materials

All the materials were created following the procedure described in Chapter 4 in an experiment with French-speaking teenagers. All materials, data, and code of this study are accessible on the OSF repository[2].

5.1.2.1. Sentence cloze test

Choice of Connectives. Six types of common coherence relations (Sanders et al., 1992), namely addition, cause, concession, consequence, contrast, and temporality, were selected for this experiment. Each coherence relation was represented by two connectives – one that is more common in oral speech and one that is more prevalent in written language. Moreover, six connectives that were included in the task were polyfunctional, namely *hotia* 'even if', *no* 'but', *odnako* 'however', *da i* [3] 'moreover', *vpročem* 'nevertheless', *sledovatel'no* 'therefore', and another six were monofunctional, namely *potomu čto* 'because', *tak čto* 'so', *kak tolko* 'as soon as', *krome togo* 'moreover', *tak kak* 'because', *edva* 'as soon as'. The number of functions that each connective can encode was determined based on the dictionaries of Yefremova (2000) and Yevgen'eva (1999). Moreover, I also conducted a corpus analysis in order to trace the dominance of the tested functions for these polyfunctional connectives. For each connective, I

[2] https://osf.io/gqpy7/?view_only=1f466fe2e9cc42aba5b53b721d9475ba
[3] Russian connectives were transliterated from the Cyrillic to the Latin alphabet for convenience throughout the chapter.

annotated 50 sentences, randomly extracted from the oral subcorpus of the *Russian National Corpus* (https://ruscorpora.ru; Grishina & Savchuk, 2009), as well as 50 sentences extracted from the written subcorpus. The results of the corpus analysis showed that for the majority of the polyfunctional connectives, the tested function was dominant both in written and oral corpora (see Table 1 for the distribution of different functions). It appears that for the connective *vpročem* 'nevertheless', it is not possible to distinguish one dominant function, as the concessive use tends to be more frequent in the written mode, and the contrastive function is more frequent in the oral mode.

Table 1. Number of occurrences of different functions of the polyfunctional connectives in written and oral corpora

		Concession	Consequence	Contrast	Addition	Total
da i 'moreover'				9	91	100
	Oral			2	48	50
	Written			7	43	50
hotia 'even if'		85		15		100
	Oral	44		6		50
	Written	41		9		50
no 'but'		6		80	14	100
	Oral	4		39	7	50
	Written	2		41	7	50
odnako 'however'		8		90	2	100
	Oral	5		43	2	50
	Written	3		47		50
sledovatel'no 'therefore'			79		21	100
	Oral		40		10	50
	Written		39		11	50
vpročem 'nevertheless'		35		40	25	100
	Oral	8		32	10	50
	Written	27		8	15	50

To determine which connectives were more bound to oral speech and which ones to written language, I conducted a corpus analysis of connective frequencies and administered a questionnaire to gauge native speakers' judgments. I calculated the connective frequencies in oral speech based on the *Russian National Corpus*, as it is large (13.4 million words) and contains speech from a wide variety of genres and degrees of formality, such as everyday conversation and public speech. The connective frequencies in writing were calculated based on the written sub-corpora of the *Russian National Corpus*, including journalistic, literary, scientific, and technical texts. Those connectives that had higher frequency in oral than in written corpora were classified as *oral*; and those with a higher frequency in the written subcorpus were categorized as *written*.

In addition to the corpus study, I recruited online 109 adult native Russian speakers to verify whether each of the selected connectives was common for an informal oral conversation, such as the one at a dinner with friends. The participants had to make their evaluation on a scale from 0 to 20. If they reckoned that a connective was never used in informal oral conversation, they were to choose the answer 0; and if they believed that it was used in such contexts very often, they were asked to choose the answer 20. For each pair of connectives representing the same coherence relation, the connective with a higher total was labelled as *oral* and that with a lower one as *written*. The results from the judgement test were congruent with the categorization based on the corpus analysis. Participants who fulfilled the judgement task did not take part in the main experiment. Table 2 shows the distribution of the 12 selected connectives per coherence relation, modality, and polyfunctionality.

Table 2. Distribution of connectives per type of coherence relation, modality, and polyfunctionality (Poly) with their mean subjective orality rate (Mor) and frequency per million words in oral (Freq OR) and written (Freq WR) corpora

Relation	Mode	Connective	Translation in English	M_{OR} (SD)	Freq OR	Freq WR	Poly
Addition	oral	*da i*	moreover (less formal)	14.09 (5.24)	232.18	163.89	+
	written	*krome togo*	moreover (more formal)	10.48 (5.09)	93.16	307.10	–
Cause	oral	*potomu čto*	because (less formal)	17.67 (3.16)	2565.03	453.94	+
	written	*tak kak*	because (more formal)	12.06 (5.28)	191.90	286.23	–
Concession	oral	*hotia*	even if	16.02 (4.09)	630.96	536.00	+
	written	*vpročem*	nevertheless	8.77 (5.35)	19.00	204.45	+
Consequence	oral	*tak čto*	so	15.96 (3.93)	474.69	192.43	–
	written	*sledovatel'no*	therefore	7.63 (5.19)	28.51	60.92	+
Contrast	oral	*no*	but	18.68 (2.39)	5999.09	4427.84	+
	written	*odnako*	however	9.16 (5.51)	51.02	738.67	+
Temporality	oral	*kak tolko*	as soon as (less formal)	12.81 (4.62)	57.53	54.99	–
	written	*edva*	as soon as (more formal)	6.92 (4.78)	10.74	107.71	–

Design of the Cloze Task. Participants were asked to fill in a blank between two sentences with a correct connective, by choosing between four options. The blank was delimited with double slashes '//________//' instead of punctuation marks so that punctuation on the border between two sentences did not affect the choice of a connective. There were 60 pairs of sentences in the task. Each coherence relation was represented by 10 items, half of which tested a connective typical to oral speech and another half the one mostly used in writing. I tested oral and written modalities separately in order to avoid more common oral connectives being always selected instead of written ones. To do so, I presented only oral connectives as answer options in sentences targeting oral connectives and only written connectives in sentences targeting written connectives. Examples (3) and (4) illustrate how this principle was applied to the relation of consequence.

(3) The correct answer: written connective *sledovatel'no* ‘therefore’

Саша пропустил много лекций //________// ему будет непросто на экзамене.

‘Sasha has missed a lot of lectures //________// he will have a hard time at the exam.’

Answer options: (a) vpročem ‘nevertheless’; (b) krome togo ‘moreover (more formal)’; (c) sledovatel'no ‘therefore’; (d) edva ‘as soon as (more formal)’

(4) The correct answer: oral connective *tak čto* ‘so’

Маша не спала всю ночь //________// на утро у нее сильно болела голова.

‘Masha stayed up all night //________// she had a bad headache in the morning.’

Answer options: (a) hotia ‘even if’; (b) da i ‘moreover (less formal)’; (c) tak čto ‘so’; (d) kak tolko ‘as soon as (less formal)’

To verify whether the linguistic context in which connectives were used was well suited for all the tested connectives, I asked a different group of 40 adult native Russian speakers, recruited via an online platform (Prolific, Oxford, UK, www.prolific.co), to judge the acceptability of the task sentences in the Russian language on the scale from 0 to 10, where 0 stood for *absolutely not acceptable* and 10 stood for *absolutely acceptable*. To the 60 items from the cloze test with correctly inserted connectives, I added 24 fillers. Half of the fillers included wrong connectives, as in (5), and another half included lexico-grammatical mistakes, such as wrong usage of phrasal expressions (6) and verb government errors (7). In the example (6), for instance, the phrasal expression *язык не* *поворачивался* ‘the tongue did not turn’ was replaced by the wrong expression *язык не поднимался* ‘the tongue did not rise’. The sentence (7), in its turn, included an error in the government of the verbal expression *уделять внимание* ‘pay attention’ that should be followed by the noun *child* in the dative case (*ребёнку*). Instead, it was followed by the noun *child* in the accusative case with the preposition *на* ‘on’ (*на ребёнка*).

The results of the acceptability judgement task showed that the overall acceptability rate of the cloze test items (*M*=8.33, SD=2.36, Range 0–10), as well as the acceptability of the specific items testing oral (M=8.72, SD=2.02, Range 0–10) and written (M=7.95, SD=2.61, Range 0–10) connectives were quite high. This suggests that both types of connectives were tested in a context that was suitable for them.

(5) Пошёл снег, потому что мы зашли домой.

‘It started snowing because we came home.’

(6) Маша была так напугана, что у неё язык не поднимался ничего сказать в ответ.

‘Masha was so frightened that she had no tongue to say anything back.’

(7) Антонина не достаточно уделяет внимание на ребёнка из-за большой занятости на работе.

‘Antonina doesn't pay enough attention to the child because she is very busy at work.’

I also ensured that there was only one possible answer for each item, as connectives whose primary or secondary function could interfere with the targeted connective were not included in the choice of answers. For instance, in sentences testing concessive connectives *vpročem* ‘nevertheless’ and *hotia* ‘even if’, I did not propose as answer options connectives *odnako* ‘however’ and *no* ‘but’ that can signal contrastive and concessive relations. I calculated the score for this task as a proportion of correct answers per connective.

5.1.2.2. Vocabulary test

The participants' vocabulary size was measured with a newly developed vocabulary level test based on Nation and Beglar (2007). The task involved choosing out of six options the word that corresponded best to a given definition. There were four groups of words (30 items each), selected from frequency lists of the *Russian National Corpus* (https://ruscorpora.ru; Lyashevskaya & Sharoff, 2009) and representing the first, second, third, and fourth 5000-word families. The participants fulfilled the task by starting from the first group of words, having the highest frequency, and by finishing with the fourth group of words, having the lowest frequency. Among the words included in each group, there were 18 nouns, 6 verbs, and 6 adjectives. The reliability of the test, as measured by Cronbach's alpha, was high both for teenagers, .99, 95% CI [.986–.991], and adults, .91, 95% CI [.85–.93]. Hence, the total vocabulary score was computed as the percentage of correct responses per participant.

5.1.2.3. Author recognition tests

To assess the teenagers' level of exposure to print, I administered an author recognition test (ART) developed specifically for this study. The new version of the ART (*ART-RU-CL*) included 40 names of classic authors and 40 filler non-author names, that were presented in a random order. The classical authors' names were chosen based on the classifications provided by four national bookstore chains. Participants were asked to choose all the authors’ names that they knew. To avoid guessing, the participants were instructed to check only those names about which they were sure, as not all the names belonged to real authors, and they would lose one point per each wrongly selected name. One point was attributed for each correctly checked author, and *-1* for each wrongly chosen one. To calculate the final score, I made a sum of correct and wrong answers, where the minimum total score was *-40* and the maximum was *40*.

Another version of the ART (*ART-RU*) was developed for the group of adults since this measure is sensitive to the age of participants (see, e.g., Allen et al., 1992; Cunningham & Stanovich, 1990). This test followed design principles developed by Stanovich and West (1989) in the original ART, and included the names of modern and contemporary authors who won literary prizes or are bestselling. The test structure and the score calculation procedure were the same as in the ART-RU-CL. The reliability of both ART tests was high, as indicated by their Cronbach's alphas greater than .90 (ART-RU-CL: .94, 95% CI [.92–.95]; ART-RU: .92, 95% CI [.86–.94]).

Finally, all the participants also provided a subjective evaluation of their exposure to print. More specifically, they had to evaluate their reading habits on the scale from 0 to 10, where 0 stood for *never* and 10 for *every day*.

5.1.3. Procedure

All the participants performed the tasks in the same order, starting with the connective cloze task and then proceeding to the author recognition test and the vocabulary level test. They could not return to previous questions and make changes as soon as they clicked the button leading to the next question. Teenagers fulfilled all the tasks online via a link that they received directly on their classroom computers. It took them approximately one hour to finish the whole test battery. Adults also completed the tasks online, but via the Prolific website (https://www.prolific.co), and spent around 40 minutes on it.

5.1.4. Analysis

I used a generalized mixed-effects logistic regression model to analyse binary responses (right or wrong) on the connective test in the R software (R Core Team, 2020). To examine whether there was a general difference between teenagers and adults in the performance on the connective task, I first analysed the results of all participants together. After a global analysis, I separately analysed the results of teenagers and adults to assess the role of the predictors of individual variation within each group. The measures of inter-individual variation were not included in the analysis for all participants because different versions of the ART were used to assess the degree of exposure to print in teenagers and adults.

First, I centred all the predictors of individual variation and then created a full model with the *glmer* function of the *lme4* package (Bates et al., 2015), including all the relevant variables. The model for all participants included Coherence Relation, connective Mode (Oral versus Written), Polyfunctionality (Monofunctional versus Polyfunctional), and Group (Adults versus Teenagers).

The model for teenagers included degree of Exposure to Print (ART-RU-CL), Vocabulary Level, Subjective Evaluation of Exposure to Print, Mode, Polyfunctionality, and Age. When I checked the data for multicollinearity, I observed that Vocabulary Score correlated with ART-RU-CL (*rho*=.73, 95% CI [.64, .81], *p*<.001) and Age (*rho*=.33, 95% CI [.16, .49], *p*<.001), as well as Age and ART-RU-CL also correlated with each other (*rho*=.33, 95% CI [.16, .48], *p*<.001). To avoid multicollinearity, I first residualized Vocabulary Score by ART-RU-CL and Age, and then, Age by ART-RU-CL, using the *umx_residualize* function of the *umx* package (Bates, 2021).

In the full model for adults, there were the same variables as in the one for teenagers, except for the ART-RU that was included instead of the ART-RU-CL, adapted for teenagers. As it was the case in the teenagers' data, Age correlated with ART-RU-CL (*rho*=.30, 95% CI [.01, .53], *p*<.001); and Vocabulary Score correlated with ART-RU (*rho*=.55, 95% CI [.31, .72], *p*<.001) and Age (*rho*=.43, 95% CI [.17, .64], *p*<.001). To avoid multicollinearity, I also residualized Vocabulary Score by ART-RU and Age, as well as Age by ART-RU.

After having built the full models, with the *drop1* function of the *stats* package (R Core Team, 2020), I selected only those predictors that were relevant for the reduced models. When the factors of Mode and Coherence Relation were selected among other relevant predictors for the reduced model, I checked whether adding an interaction between these two factors improved the final model's fit. Comparison between the models without and with an interaction was done with the *anova* function of the *stats* package. The final reduced models were returned with the *summary* function of the *lmerTest* package (Kuznetsova et al., 2014). The significance threshold was set at *p*<.05.

In the separate models for teenagers and adults, I also performed a random forest analysis with the predictors that were kept in the final reduced models (Strobl et al., 2009). I chose to complement regression analysis with this method because it can deal even with highly correlated variables (such as the measures of individual variation tested in this chapter), as it does not rely on any assumptions about the distribution of data. This method is also robust because it computes the importance of each variable on the basis of a large number of regression trees (Strobl et al., 2009).

5.2. Results

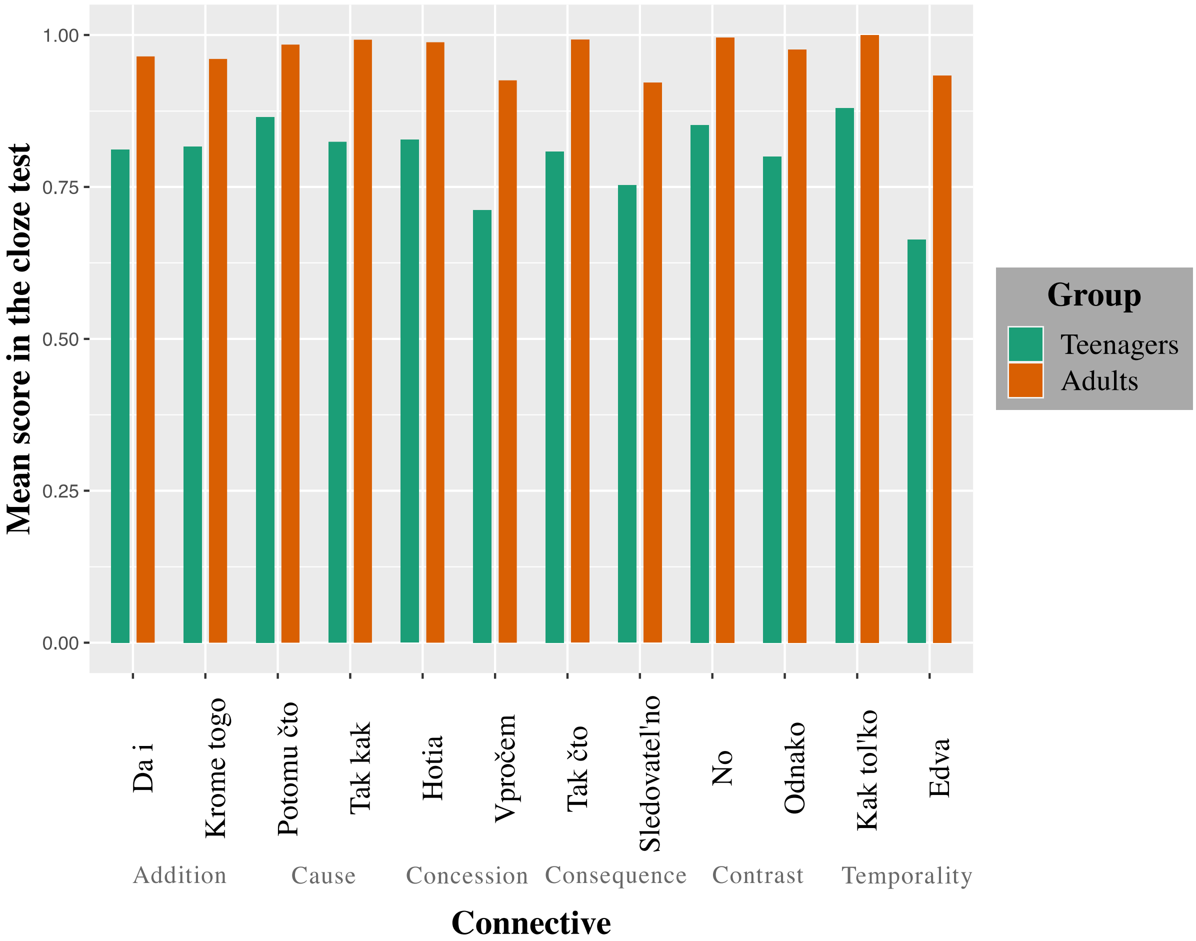
5.2.1. All participants

The final reduced model for all participants included fixed effects of Group, Coherence Relation, and Mode, as well as a three-way interaction between Group, Coherence Relation, and Mode. Adding, first, an interaction between Mode and Coherence Relation (Δx^2=24.93, Δdf=5, *p*=<.001) and then between Group, Mode, and Coherence Relation (Δx^2= 27.89, Δdf=11, *p*=.003) improved the models' fit. Connective Polyfunctionality was not among relevant predictors of the performance in the connective task. The output of the statistical analysis showed that teenagers scored significantly lower than adults in the cloze task, as the factor of Group accounted for an estimated decrease of 2.06±0.62 SE (see Figure 1 and Table 3).

Moreover, there was a significant interaction between Group, Mode, and Coherence Relation for the relations of temporality, consequence, and concession (see Table 3 for the estimates). This means that the written connectives *edva* 'as soon as', *sledovatel'no* 'therefore', and *vpročem* 'nevertheless', encoding temporal, consequence, and concessive relations, respectively, received the lowest scores across both groups of participants (see Table 4 for the mean scores per connective within each age group). Finally, the relation of temporality was particularly challenging for the group of teenagers, as an interaction between teenagers and temporality accounted for an estimated decrease of 14.15±4.64 SE. However, teenagers overall scored quite high, namely between .80 and .88 for the majority of connectives, except for the above-mentioned written connectives *vpročem* 'nevertheless' (*M*=.71, 95% CI [.60, .83]), *sledovatel'no* 'therefore' (*M*=.75, 95% CI [.64, .86]), and *edva* (*M*=.66, 95% CI [.54,

.78]). The general accuracy level of adults, in contrast, was above 90% and varied between .92 and 1.

Figure 1. Distribution of mean scores per connective in sentence cloze task among teenagers and adults



Note. The translations of the Russian connectives are as follows: *da i* ‘moreover (less formal)’, *krome togo* ‘moreover (more formal)’, *potomu čto* ‘because (less formal)’, *tak kak* ‘because (more formal)’, *hotia* ‘even if’, *vpročem* ‘nevertheless’, *tak čto* ‘so’, *sledovatel'no* ‘therefore’, *no* ‘but’, *odnako* ‘however’, *kak tolko* ‘as soon as (less formal)’, *edva* ‘as soon as (more formal)’.

Table 3. Output of the full models and the final reduced models for all groups of participants

Variable	**Full model**					**Final reduced model**			
	Estimate	**SE**	**z**	**Pr(>|z|)**		**Estimate**	**SE**	**z**	**Pr(>|z|)**
All participants									
(Intercept)	5.50	0.38	14.47	<0.001	(Intercept)	5.13	0.61	8.46	<0.001
GROUP					GROUP				
Teenagers	-2.22	0.36	-6.17	<0.001	Teenagers	-2.06	0.62	-3.33	0.001
MODE					MODE				
Written	-0.85	0.15	-5.51	<0.001	Written	0.70	0.87	0.81	0.418
COHERENCE RELATION					COHERENCE RELATION				
Addition	-0.23	0.30	-0.76	0.449	Addition	-0.88	0.66	-1.34	0.182
Concession	-0.37	0.37	-1.00	0.318	Concession	0.36	0.80	0.45	0.653
Consequence	-0.54	0.30	-1.80	0.071	Consequence	0.71	0.89	0.80	0.424
Contrast	0.21	0.38	0.55	0.582	Contrast	1.43	1.12	1.28	0.200
Temporality	-0.67	0.27	-2.51	0.012	Temporality	14.34	4.64	3.09	0.002
POLYFUNCTIONALITY					GROUP*MODE				
Polyfunctional	-0.39	0.26	-1.50	0.134	Teenagers*Written	-1.19	0.86	-1.39	0.165
					GROUP*COHERENCE RELATION				
					Teenagers*Addition	0.25	0.63	0.40	0.693

					Teenagers*Concession	-0.79	0.79	-1.00	0.316
					Teenagers*Consequence	-1.37	0.88	-1.56	0.118
					Teenagers*Contrast	-1.59	1.10	-1.44	0.151
					Teenagers*Temporality	-14.15	4.64	-3.05	0.002
					MODE*COHERENCE RELATION				
					Written*Addition	-0.75	1.02	-0.74	0.461
					Written*Concession	-2.69	1.10	-2.45	0.014
					Written*Consequence	-3.20	1.16	-2.76	0.006
					Written*Contrast	-2.52	1.38	-1.83	0.068
					Written*Temporality	-16.63	4.64	-3.59	<0.001
					GROUP*MODE*COHERENCE RELATION				
					Teenagers*Written*Addition	1.31	0.99	1.32	0.187
					Teenagers*Written*Concession	2.06	1.07	1.93	0.054
					Teenagers*Written*Consequence	3.14	1.14	2.77	0.006
					Teenagers*Written*Contrast	2.43	1.36	1.79	0.074
					Teenagers*Written*Temporality	14.96	4.63	3.23	0.001
Teenagers									
(Intercept)	-3.41	0.58	-5.88	<0.001	(Intercept)	-2.84	0.40	-7.06	<0.001
Vocabulary test**	6.98	0.60	11.56	<0.001	Vocabulary test**	6.98	0.61	11.47	<0.001
ART-RU-CL*	1.93	0.14	13.86	<0.001	ART-RU-CL*	1.96	0.14	14.06	<0.001
Age*	2.67	0.80	3.34	0.001	Age*	2.29	0.77	2.98	0.003

Subjective Exposure to print*	0.40	0.25	1.64	0.101					
MODE					MODE				
Written	-0.81	0.16	-4.93	<0.001	Written	-0.81	0.17	-4.88	<0.001
POLYFUNCTIONALITY									
Polyfunctional	-0.20	0.16	-1.19	0.234					
Adults									
(Intercept)	5.06	1.06	4.78	<0.001	(Intercept)	3.99	0.57	7.06	<0.001
Vocabulary test**	19.26	5.34	3.61	<0.001	Vocabulary test**	19.76	5.50	3.59	<0.001
ART-RU*	0.81	0.22	3.61	<0.001	ART-RU*	0.71	0.20	3.58	<0.001
Age*	0.81	0.67	1.21	0.225	MODE				
Subjective Exposure to print*	-0.52	0.54	-0.95	0.341	Written	-1.51	0.46	-3.29	0.001
MODE									
Written	-1.53	0.45	-3.36	0.001					
POLYFUNCTIONALITY									
Polyfunctional	-0.50	0.43	-1.15	0.250					

*Centred values

**Centred and residualized values

Table 4. Mean accuracy score per connective among teenagers and adults

Mode	Connective	Translation in English	Teenagers		Adults	
			M	95% CI	M	95% CI
oral	*da i*	moreover	.81	[.71, .91]	.96	[.92, 1.01]
written	*krome togo*	moreover	.82	[.72, .91]	.96	[.91, 1.01]
oral	*potomu čto*	because	.87	[.71, .91]	.98	[.95, 1.02]
written	*tak kak*	because	.82	[.73, .92]	.99	[.97, 1.01]
oral	*hotia*	even if	.83	[.73, .92]	.99	[.96, 1.02]
written	*vpročem*	nevertheless	.71	[.60, .83]	.92	[.86, .99]
oral	*tak čto*	so	.81	[.71, .91]	.99	[.97, 1.01]
written	*sledovatel'no*	therefore	.75	[.64, .86]	.92	[.85, .99]
oral	*no*	but	.85	[.76, .94]	.99	[.98, 1.01]
written	*odnako*	however	.80	[.70, .90]	.98	[.94, 1.01]
oral	*kak tolko*	as soon as	.88	[.80, .96]	1	[1, 1]
written	*edva*	as soon as	.66	[.54, .78]	.93	[.87, 1]

5.2.2. Teenagers

An automatic step-back analysis of data showed that the performance in the connectives test within the group of teenagers was mostly predicted by Vocabulary Level, Exposure to Print, Age, and Mode (see Table 3). Polyfunctionality and Subjective Exposure to Print were not revealed to be significant predictors of teenagers' performance in the connective task. A higher Vocabulary Level accounted for an estimated increase of 6.98±0.61 SE in the connectives task, higher scores in ART-RU-CL were associated with an estimated increase of 1.96±0.61 SE, and Age accounted for an increase of 2.29±0.77 SE estimates (see Table 3). Moreover, connectives from the written mode were on average slightly more challenging than oral ones, as reflected by an estimated decrease of 0.81±0.17 SE.

The random forest analysis (prediction accuracy of 89%) demonstrated that the most important predictor of performance with connectives by teenagers was Vocabulary Level, outranking Exposure to Print, Age, and Mode (Figure 2). The scores on the vocabulary test and the ART across teenagers and adults are reported in Table 5.

Figure 2. The impact of each predictor variable on the dependent variable according to the random forest analyses for teenagers and adults

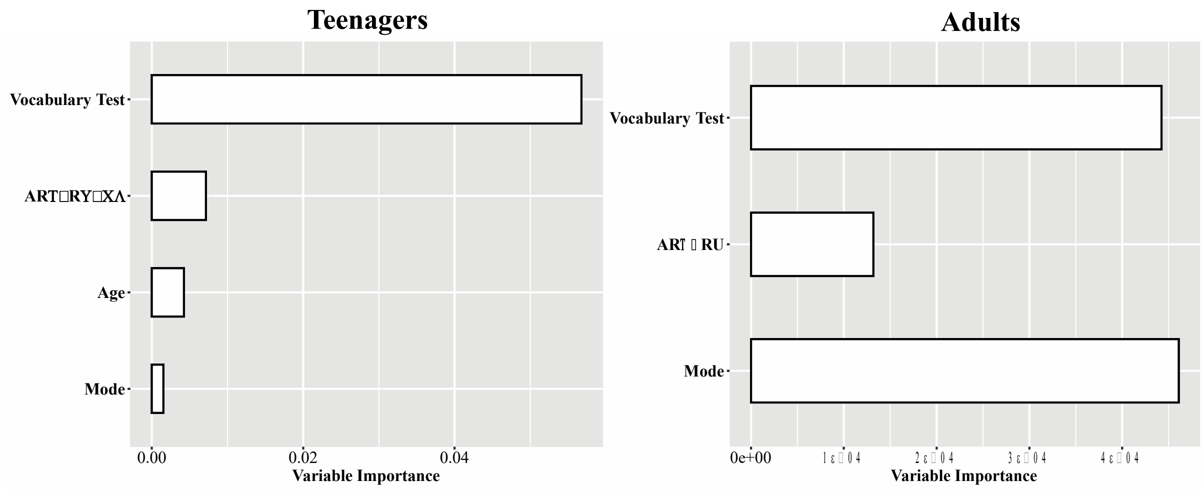


Table 5. Descriptive statistics for background measures among teenagers and adults

Teenagers		Adults		Possible range
M (SD)	**Observed range**	**M (SD)**	**Observed range**	
Vocabulary size				
.73 (.29)	.12-1	.96 (.06)	.79-1	0–1
Author recognition test*				
13 (10.01)	-6-40	10.71 (7.76)	0-32	-40–40
Subjective exposure to print				
5.95 (2.37)	1-10	6.63 (2.07)	2-10	0–10

*A different version of the ART was used for teenagers and adults

5.2.3. Adults

The final reduced model for the group of adults included Vocabulary Level, Exposure to Print, and Mode (see Table 3). Polyfunctionality, Age, and Subjective Exposure to Print were not significant predictors of the performance in the cloze test. Higher scores on the vocabulary level test were associated with an estimated increase of 19.76±5.50 SE in the sentence cloze task; and a greater degree of exposure to print, as measured by the ART-RU, accounted for an estimated increase of 0.71±0.20 SE. In contrast, written mode accounted for a decrease of 1.51±0.46 SE estimates (see Table 3). The random forest analysis had prediction accuracy of 97% and revealed that, for the group of adults, the most important predictors of an accurate use of connectives in the cloze test were Vocabulary Level and Mode, followed by Exposure to Print (see Figure 2).

5.3. Discussion

5.3.1. Factors related to the properties of connectives

This study set out to explore the factors influencing the acquisition of Russian discourse connectives in teenage years. In line with the earlier study on French (see Chapter 4), the present experiment demonstrated that, when used in a monofunctional context, the twelve connectives targeted in this chapter do not pose big difficulties for Russian-speaking teenagers. However, it is possible that the high performance in the cloze test was due to the task design, in which participants had to fill in blanks between pairs of sentences and not in a more ecological context of texts. Indeed, it was shown by Tskhovrebova et al. (2022) that it is cognitively more challenging for both teenagers and adults to apply an appropriate connective within a short text than between two sentences. In future studies, it will therefore be useful to increase the challenge of the task and to examine the use of infrequent connectives, expressing rare coherence relations, in more realistic contexts. Although the current task modality was probably not particularly challenging for teenagers, their scores were still lower than those of adults. This result supports studies by Berman (2004) and Nippold (2008) and suggests that adult competence to use connectives is not completely acquired by age 18 and continues to develop far beyond high school years.

Not all connectives however were mastered equally well by teenagers. There was a significant interaction between a coherence relation type and mode on the performance of teenagers in the cloze test. The written connectives *vpročem* ‘nevertheless’, *sledovatel'no* ‘therefore’, and *edva* ‘as soon as’ received lower scores than their oral counterparts. This result is in line with the prior finding on French-speaking teenagers from Chapter 4 and supports the initial hypothesis on the role of modality, indicating that lack of sufficient exposure to the written modality may be at the heart of the lower performance, at least with some written connectives. In order to improve competence with this type of connectives, teenagers should have more training at school with different registers of written language as well as with connectives that most often appear in this mode.

As for the factor of coherence relation type, it was not the main predictor of the performance in the cloze test. Out of six coherence relations, only the relation of temporality, and more specifically, the written connective *edva* ‘as soon as’, was associated with a certain level of difficulty. This result corroborates the findings from Chapter 3 and suggests that complexity of coherence relations is not a prominent predictor of the performance with connectives in off-line production tasks. In addition, this finding also hints that the effects of coherence relation type found in Chapter 4 were most probably related to the particular features of the connectives selected to represent those relations (such as the additive connective *en outre* ‘moreover’) and not to the type of coherence relations per se.

The score for the temporal written connective *edva* ‘as soon as’ was particularly low in the group of teenagers. It is difficult to propose a solid explanation for this finding. My tentative suggestion would be that this word can also be used as an adverb, translated into English as *barely* and *slightly*. Therefore, these non-connective meanings may interfere and mislead young speakers in their attempts to match an appropriate connective with a temporal coherence relation. However, a more comprehensive examination of how non-connective meanings may affect the performance with the connective functions should be conducted in future studies.

Finally, polyfunctionality was not a significant predictor of the performance in the sentence cloze test across both age groups. This result may indicate that polyfunctionality is not an important source of complexity for native speakers of Russian, starting from early teenage years. The existence of other functions does not make polyfunctional connectives more challenging to use in a cloze task, at least when alternative meanings are controlled for in the experimental design. However, this result may have stemmed from the fact that I targeted dominant functions of the polyfunctional connectives. Future research should explore more closely whether, in a similar experimental design, the performance with non-dominant functions would be equally high. On a more general note, future work would also benefit from creating a standardised inventory of connectives in Russian.

5.3.2. Student-level predictors of the mastery of connectives

The results have demonstrated that, similarly to the findings for French-speaking teenagers in Chapter 4, vocabulary level was the most important predictor of the varying competence with connectives in the cloze task. This result suggests that vocabulary knowledge plays an important role in the use of connectives starting from early teenage years and during adulthood, even though connectives are not typical lexical items, as they encode procedural rather than (or along with) conceptual meaning (see, e.g., Wilson, 2011).

Another factor that predicted differences in performance on the cloze task was the level of exposure to print, as assessed by the newly created Russian versions of the ART. This finding complements the existing research in adults (Scholman et al., 2020; Zufferey & Gygax, 2020a) and French-speaking teenagers (Tskhovrebova, Zufferey & Tribushinina, 2022), showing that the ability of Russian-speaking teenagers and adults to use connectives also varies according to the degree of exposure to print, starting as early as at age 11. It should be noticed that even though the ART does not measure the actual number of books that a person has read in their life, it has been shown to be indicative of a level of general linguistic competence, including sentence-processing skill (Acheson et al., 2008), world and vocabulary knowledge (see, e.g., Stanovich et al., 1995), as well as metacognitive competence (McBride-Chang & Chang, 1995). Since the ART reflects a complex set of language skills, future studies will need to disentangle them in order to better understand variations in the ability to use connectives. Furthermore, it should be mentioned that the comparison of the effects from the ART on the use of connectives by teenagers and adults must be done with caution, as two different versions were used to measure the degree of exposure to print in the two age groups of participants. In general, the question of how to handle the scores from the ARTs, adapted to different age groups of speakers, should be explored more in detail in future research.

Although the ART test is an indirect measure of exposure to print, it was better suited for assessing differences in the use of discourse connectives than the subjective evaluation of exposure to print. The reason for such a difference between indirect and direct measures of exposure to print is that self-report tests are more prone to socially desirable answers (see, e.g., Echols et al., 1996; Wimmer & Ferguson, 2022). Therefore, in future work, the ART tests should be preferred to subjective evaluations of exposure to print.

Finally, age was found to be another factor predicting a better performance in a connective cloze task, but only for the group of teenagers. However, the random forest analysis demonstrated that age did not play as an important role as lexicon size and degree of exposure to print. This finding is not in line with previous studies on children, where mastery of connectives was strongly predicted by age (Blything et al., 2015; Pyykkönen & Järvikivi, 2012), but it corroborates the results of a similar experiment on French-speaking teenagers in Chapter 4. Although this interesting result needs further replications within and across languages, the findings that age plays a less important role than the measures of individual variation both for French- and Russian-speaking teenagers may suggest that later linguistic development is qualitatively different and follows a slower pace than early language acquisition (Nippold, 1993). Appropriate use of a wide variety of connectives, which happens at a later developmental stage, probably requires additional effort and input, on top of age-related cognitive maturation. In addition, this result may also highlight the fact that chronological age is an indirect measure of various competences, such as cognitive skills and language experience, that are subject to individual variation (Kidd et al., 2018). Indeed, the factor of age was not relevant at all for the group of adult speakers, whose mastery of connectives probably does not further develop with chronological age.

5.4. Conclusion

The study presented in this chapter emphasizes the importance of research on linguistic development in teenage years, a period that has received relatively little attention in language acquisition research. The present results corroborate general findings from a similar study on French-speaking teenagers (see Chapter 4 and Tskhovrebova, Zufferey & Tribushinina, 2022) and show that, by age 18, Russian-speaking teenagers have a high command of 12 connectives signalling most common coherence relations, belonging to the written and the oral modes, and having one or several functions. This result suggests that the development of the mastery of connectives in Russian follows a similar course as the one observed in French. In the present experiment, I also reported a strong individual variation among Russian speakers in the ability to match connectives with an appropriate coherence relation. The use of connectives in a cloze task was strongly predicted by teenagers' vocabulary size and level of exposure to print and more marginally by speakers' age, and connective modality. This finding may suggest that, at least starting from 11, age and intrinsic properties of connectives matter less for their mastery than general linguistic experience, as measured by vocabulary level and exposure to print. Moreover, exposure to print and vocabulary size continue to explain individual differences in the performance with connectives even during adult years. Therefore, school curricula should support and promote students' exposure to written texts, which will enable them to enlarge their vocabulary size and to become more familiar with the use of a greater number of connectives, thereby enhancing reading comprehension skills and general academic performance.

In the next chapter, I pursue further the examination of endogenous and exogenous factors influencing the development of competence with connectives. Since connectives are not the only linguistic devices signalling coherence relations between parts of discourse, I explore how the presence of different types of connectives interacts with other, alternative signals indicating coherence relations. More precisely, I try to answer the question whether the sensitivity to non-connective signals, which are less salient and prototypical, is affected by the presence of more prominent signalling devices such as connectives. In addition, I aim to evaluate whether the sensitivity to alternative signals in young speakers also varies depending on individual differences in linguistic competence.

6. Exploring the Sensitivity to Alternative Signals of Coherence Relations by French Speaking Teenagers[1]

Coherence relations between segments of discourse can be signalled by various linguistic means. So far, I have presented studies exploring only one type of signalling devices, namely, connectives. These linguistic elements are among the most studied signals of coherence relations (see, e.g., Bloom et al., 1980; Champaud & Bassano, 1994; Blything et al., 2015), with studies focusing on speakers from various age groups (see, e.g., Blything & Cain, 2016; Evers-Vermeul & Sanders, 2009; Nippold et al., 1992; Peterson, 1986) and linguistic competences (see, e.g., Crosson et al., 2008; van Silfhout, 2015; Volodina & Weinert, 2020).

[1] Results of this study have been submitted for publication in Tskhovrebova, Zufferey and Gygax (2023, submitted).

Many coherence relations, however, are not marked by a connective but rather conveyed implicitly. In the Penn Discourse Treebank, about 41% of the relations are not marked by connectives (Webber et al., 2019). Liu (2019), examining the distribution of signals across four different text genres, namely academic articles, how-to guides, interviews, and news articles, from the Georgetown University Multilayer (GUM) corpus (Zeldes, 2017), also found that connectives signal only 16% of relations as opposed to 84% of relations marked by other signal types. Similarly, in the mono-genre RST Discourse Treebank (Carlson et al., 2002), only about 11% of coherence relations are signalled exclusively by connectives, whereas approximatively 75% of relations are marked by other, alternative types of coherence signals (Das & Taboada, 2018). In fact, Das and Taboada (2018) identified at least seven types of alternative coherence markers in this corpus, such as lexical, semantic, morphological, syntactic, graphical, genre, and numerical features (for other approaches to the annotation of different signal types, see Knaebel & Stede, 2022; Liu & Zeldes, 2019; Zeldes & Liu, 2020). For instance, such syntactic feature as the present participial clause in (1) can signal the relation of manner; and the antonyms in (2) are the semantic signals of a contrastive relation (Das, 2014).

(1) Wyse has done well, <u>establishing a distribution business</u>.

(2) … <u>higher</u> bidding <u>narrows</u> the investor's return, while <u>lower</u> bidding <u>widens</u> it.

Although these corpus studies demonstrate the prevalence of coherence signalling by other means than connectives, less is known about the way readers infer coherence relations from these types of signals (but see Brown & Fish, 1983; Scholman et al., 2020). Moreover, very few studies (except for Crible, 2021; Crible & Demberg, 2021; Crible & Pickering, 2020; Grisot & Blochowiak, 2021; Schwab & Liu, 2020) have assessed how different types of signals interact with each other. Lexical (Schwab & Liu,

2020) and syntactic (Crible & Pickering, 2020) cues, for instance, were found to reinforce inference of a particular coherence relation signalled by polyfunctional connectives, such as *but* or *and*. However, it is not clear whether the interaction between alternative signals and connectives would be the same if the latter were monofunctional, i.e., specialized in marking one type of coherence relation. In comparison to alternative signals, connectives are more salient markers of coherence, as the signalling of coherence relations is their primary function, and they are often used in a prominent clause-initial position. In contrast, alternative signals often occupy less prominent syntactic positions and are not specialized in signalling coherence relations. Therefore, an important question is whether the inference from alternative signals, such as the lexical or semantic features from Das and Taboada (2018), is still generated even in the presence of a stronger cue of coherence relations, such as a connective.

As little as we know about the functioning of alternative signals of coherence relations in adults, even less is known about the sensitivity to these signals in younger populations. In other words, we do not know whether and how young speakers are guided by alternative signals in their production and comprehension of coherence relations. To the best of our knowledge, only Au (1986) examined the sensitivity of preschool children to implicit causality verbs and showed that already at the age of 5, children could perceive whether it is an agent or a patient who is causing a certain event in a sentence. However, teenage years seem not to have been studied, even though this is an important period of linguistic development between the emergence and mastery of language (Berman, 2004). Language development in teenagers continues on lexical, semantic, syntactic, and pragmatic levels (see, e.g., Nippold, 2008). The mastery of connectives, in turn, is at the interface between lexical, syntactic and pragmatic skills, which are actively developing during this period, and therefore occupy a central role in the development of a full-fledged linguistic competence (see Section 2.3.1. for a more detailed description of different factors that render teenage years crucial for linguistic development).

Previous research has shown that, on average, teenagers' competence with any type of connective is inferior to that of adult speakers (Nippold et al., 1992; Tskhovrebova, Zufferey & Gygax, 2022; Zufferey & Gygax, 2020b). Nippold et al. (1992) assessed the competence of English-speaking teenagers aged 12 to 15 and young adults aged 19 to 23 with connectives mostly used in written language, such as *furthermore* and *nevertheless*, in a connective insertion task and a sentence continuation task. The authors found that young adults performed significantly better than teenagers in both tasks. A similar result was obtained in two studies examining the usage of four French connectives bound to the written mode but varying in frequency (Tskhovrebova, Zufferey & Gygax, 2022; Zufferey & Gygax, 2020b). Both studies demonstrated that even high-school students aged 16 to 18 did not reach the performance of adults in the connective cloze task across all connectives, suggesting that proficiency with connectives continues to develop late during teenage years. Moreover, research on competence with connectives in L2, i.e., for speakers with a lower level of linguistic proficiency and who can be in that respect compared to teenagers, shows that language learners also have difficulties detecting incorrect uses, even for very frequent connectives. The study of Wetzel et al. (2022) reported, for instance, that German-speaking learners of French did not react to erroneous uses of the frequent French connective *alors* ‘so’ in a self-paced reading task.

Finally, we know that even adult speakers show variation in their linguistic competence in general (Kidd et al., 2018) and in the sensitivity to alternative signals of list relations in particular (Scholman et al., 2020). Thus, Scholman et al. (2020) demonstrated that the ability of adult speakers to infer the relation of list from the expressions of quantity like *a couple, a few, multiple, and several* varied according to the speakers' degree of exposure to print, as measured by the Author Recognition Test (Stanovich & West, 1989). In the current chapter, I extend the study of alternative list signals on a younger population in order to explore this relation further and to unveil the degree of sensitivity to such signals during teenage years. Considering the findings on the mastery of connectives by less experienced speakers and on individual variations in the sensitivity to alternative signals among adults, I suggest that teenagers may also be less proficient with alternative signals of coherence relations, and thus less sensitive to them when they are used either alone or combined with connectives.

6.1. The role of alternative signals for coherence marking

There are various types of alternative signals that can mark coherence relations. Many authors have studied the role of lexical cues for the inference of causal relations (e.g., Au, 1986; Koornneef & Van Berkum, 2006; Pyykkönen & Järvikivi, 2010; Rohde & Horton, 2014). Pyykkönen and Järvikivi (2010), for instance, showed that in the sentences *John feared Bill because …* and *John frightened Bill because …*, the implicit causality verbs *fear* and *frighten* immediately activate verb-based reference toward either the second or the first participant of the action, respectively. Kehler (1994) and Lascarides & Asher (1993) revealed the importance of morphological features, such as the combination of verb tenses, for signalling order of the occurring events. For example, in (3), the usage of the past simple in the first sentence and past perfect in the second one suggests that the event presented in the second clause (swimming in the lake) preceded the one shown in the first clause (illness).

(3) Jane fell ill. She had swum in the cold lake.

There is also evidence about alternative signals used in other coherence relations. For instance, Crible and Pickering (2020) found a facilitating effect of syntactic parallelism in combination with the connectives *but* or *and* for marking the relation of addition and contrast (4), as sentences with parallel structures were read faster than sentences without parallelism across a series of self-paced reading experiments. Schwab and Liu (2020) observed in a self-paced reading task that the lexical cues *true* and *sure*, as in example (5), helped readers to anticipate the upcoming concessive relation, as reflected by shorter reading times at the post-critical region. Moreover, Crible (2021) demonstrated in a series of four self-paced reading experiments that verbal negation, introduced in the first sentence, facilitates processing of the concessive relation, removing the difference in processing cost between the more complex concessive relation and the less complex result relation.

(4) Nick always eats in low-budget restaurants and/but Grace always eats in fancy places (Crible & Pickering, 2020, p. 8).

(5) James likes to run. True/sure, he has a treadmill in the living room, but he often jogs in parks (Schwab & Liu, 2020, p. 106).

Crible and Demberg (2021) argued that resultative verbs, as in (6), and antonyms, as in (7), respectively generate inferences of consequence and contrast relations. Yet, the inference power of these alternative signals was not as important as that of connectives signalling the same relations.

(6) Males have been proven to be more skilled at sports. It allows them to win in mixed competitions (Crible & Demberg, 2021, p. 320).

(7) The Belgian government decided to create a new tax on solar panels. The French government decided to remove the existing tax (Crible & Demberg, 2021, p. 321).

As for temporal relations, Grisot and Blochowiak (2021) reported in a bilingual French-English corpus study that pluperfects signal backward temporal relations, simple past marks forward temporal relations, and imperfectives convey synchronous temporals. Less is known, however, about the inference generation for additive relations. Still, Scholman et al. (2020) examined expressions of quantity such as *a couple, a few, multiple*, and *several*, and found that they activate in adult speakers the inference of list relation – a particular type of a more generic additive relation (see Sanders et al., 2018 for a detailed description of how different annotation frameworks categorise the list relation). In addition, the corpus study by Péry-Woodley et al. (2017) showed that the relation of list, or enumeration, can be expressed by a variety of enumerative structures of different length and graphical aspect, such as multiparagraph structures and bullet lists. Interestingly, it also showed that these structures often have a similar organization. They predominantly start with a trigger, which often includes a lexical cue. The trigger element is followed by a series of items, which in turn can be followed by a closure element.

These findings are particularly insightful, because additive relations are one of the relations that are the least signalled by connectives and are conveyed by the greatest variety of alternative signals (Das & Taboada, 2018). It even seems that speakers' comprehension of additive relations is hindered when an additive connective is present between two sentences (Kleijn et al., 2019), as in (8).

(8) Not everyone can register in the Donor Register: you must be at least twelve years old and in addition you must be a registered citizen of a Dutch municipality (Kleijn, 2018, p. 216).

This effect is different from other types of relations such as cause or contrast that elicit better comprehension scores when marked by connectives. A possible reason of this hindering effect, as suggested by Kleijn et al. (2019), is that additive connectives draw excessive attention to the coherence relation and elicit an overinterpretation of the intended relation in contrast to a simple juxtaposition. Other signals become therefore interesting to investigate, especially to better understand how additive relations work.

Another important contribution would be to examine the interaction between alternative signals of coherence relations and connectives. Only few studies have attempted to explore this interaction, reporting findings for a limited number of coherence relations, namely contrast (Crible & Demberg, 2021; Crible & Pickering, 2020), consequence (Crible & Demberg, 2021), and concession (Schwab & Liu, 2020). However, more work is needed to describe how this interaction works for other types of coherence relations. In this respect, it would be useful to provide evidence on the interaction between alternative signals and connectives signalling the less studied additive relation. For instance, assessing the interaction between alternative list signals, additive and consequence connectives for readers' propensity to generate inferences of list relations would enable us to evaluate whether these relations are still inferred from alternative signals. Importantly, one could document whether they are inferred even in the presence of stronger coherence signals such as connectives marking the same or a different type of relation. In all, it would constitute an interesting extension to the study by Scholman et al. (2020). Moreover, examining speakers' sensitivity to alternative list signals and their interaction with connectives in teenagers would allow us to fill a gap in the literature on alternative signalling in teenage years and to generalize the results of Scholman et al. (2020) to other age groups.

It is also possible that even connectives conveying the same type of coherence relation but varying in frequency may have a different impact on the generation of inferences. For instance, even adults have difficulties using (Tskhovrebova, Zufferey & Gygax, 2022; Zufferey & Gygax, 2020b) and identifying correct and incorrect uses (Zufferey & Gygax, 2020a) of the infrequent, written connectives *aussi* 'therefore' and *en outre* 'in addition'. Since speakers appear to be less confident about the usage of less frequent connectives, these connectives may also generate weaker inferences of a certain coherence relation, even combined with alternative signals. An overview of research on the competence with connectives in teenage years will allow us to make predictions on the sensitivity to alternative signals in combination with connectives (of different frequency) in this age group.

6.2. The current study

In the current set of experiments, I aim to address the gaps identified in previous research on alternative signals of coherence relations. The main goal of this chapter is to extend the study by Scholman et al. (2020) on a younger cohort of teenagers and to examine their sensitivity to alternative signals of list relation (Experiment 1) in combination with connectives varying in frequency and signalling two types of coherence relations (Experiments 2 & 3). More specifically, I assess French-speaking teenagers' sensitivity to the adjectives of quantity *plusieurs* 'several' and *différents* 'various', and how this sensitivity is modulated by the presence of connectives signalling the relations of addition and consequence. This way, I aim to examine whether a list inference, generated by an alternative signal, is strong enough to trigger list continuations even in the presence of connectives. The additive connectives were chosen because addition does not compete with the logic of the list relation. In fact, additive relations represent a generic type of relations that include several subtypes, among which the relation of list. In contrast, the consequence connectives were selected because consequence represents a separate class of coherence relations, which is competing with the logic of the list relation (see Table 1 for a summary of all the signals used in the set of experiments). I use the following definitions for the three coherence relations included in the experiments:

1. Sentences are linked with a list relation when the second sentence enumerate one or several events related to the content of the first sentence;
2. Sentences are linked with an additive relation when the second sentence expands and elaborates on the content of the first sentence, except for instances of enumeration that are included in the category of list relations;
3. Sentences are linked with a consequence relation when the second sentence describes an event caused by an activity presented in the first sentence.

Table 1. All the connectives and alternative signals used across the three experiments

<table>
<tr><th rowspan="2"></th><th rowspan="2">Alternative signals</th><th colspan="2">Connectives</th></tr>
<tr><th>Additive</th><th>Consequence</th></tr>
<tr><td>Experiment 1</td><td rowspan="3">plusieurs ‘several’
différents ‘various’</td><td>–</td><td>–</td></tr>
<tr><td>Experiment 2</td><td>en plus</td><td>donc</td></tr>
<tr><td>Experiment 3</td><td>en outre</td><td>ainsi</td></tr>
</table>

Based on the results of Scholman et al. (2020), I predict that participants will produce more list continuations after reading items containing adjectives of quantity in all experiments. However, it is possible that teenagers will be less sensitive than adults to such signals, due to a lower level of linguistic competence. I also expect that after reading sentences including both a list signal and an additive connective (Experiments 2 and 3), the proportion of list continuations should not decrease, but rather increase or remain unchanged because an additive connective is not in contradiction with the relation of list. Moreover, I predict that the combination of a list signal and a consequence connective will decrease the percentage of list continuations, as this type of connectives expresses a non-compatible relation of consequence, and this will override the inference generated by a less salient and more polysemous (in the sense that it is not specialized only in coherence marking) alternative signal of list (Experiments 2 & 3). Finally, I expect that the general effect from the less frequent connectives (Experiment 3) will be lower than from the more frequent connectives (Experiment 2).

To identify whether the sensitivity to these signals in young speakers also varies depending on individual differences in linguistic competence, I assess participants' degree of exposure to print, as measured by adapted French versions of the author recognition test (for the teenage version see Tskhovrebova, Zufferey & Tribushinina, 2022; for the adult version see Zufferey & Gygax, 2020a). I use this measure of individual variation, as, in adults, exposure to print was shown to predict readers’ mastery of connectives (Zufferey & Gygax, 2020a) and sensitivity to alternative signals (Scholman et al., 2020). Degree of general exposure to print could therefore be an important factor, modulating the inference generated by alternative signals in combination with connectives, also in teenage populations.

6.3. Experiment 1

6.3.1. Participants

Fifty-three teenagers (M_{age} = 14.18, SD = 1.66, Range: 12–18) and twenty adults (M_{age} = 31.36, SD = 11.35, Range: 21–63) took part in the experiment. All of them were native French speakers. The experiment among teenagers was carried out in secondary and high schools of the French-speaking part of Switzerland, and was performed online via a weblink. Adults were recruited online on the Prolific© platform (Prolific, Oxford, UK).

6.3.2. Materials and procedure

The data that support the findings of this study are openly available in OSF repository[2].

6.3.2.1. Story-continuation task

In this experiment, participants had to write a continuation to a series of pairs of sentences. In each pair, the first sentence introduced an agent and the context it was in, and either included a list signal (the adjectives of quantity *plusieurs* 'several' or *différent* 'various') or not. The second sentence started with a pronoun coreferential with the agent of the first sentence, and developed the situation (see example 9). The second sentence was identical across both *list signal* conditions. I did not simply remove the adjectives of quantity from the first sentence of the condition without list signal, but also changed the verb for several reasons. First, I wanted to avoid list and non-list items to be perceived as repetitions of the same sentence after reading multiple task items in a row, which could be the case if I just omitted the list signal. Second, I wanted to make sure that participants would perceive list and non-list items as different sentences and treat them as such across the whole task, but without making it obvious that the presence or absence of these alternative signals were the focal point of the task. Third, I wanted to ensure that list and

[2] https://osf.io/gqpy7/?view_only=1f466fe2e9cc42aba5b53b721d9475ba

non-list items were similar in terms of the expectations that they would create. The objective was to build neutral sentences in the non-list condition, without any obvious alternative cues of coherence (such as implicit causality verbs, for instance).

The choice of the adjectives *plusieurs* and *différents* was based on several criteria. First, they have the same function in French as previously examined English expressions of quantity from the study of Scholman et al. (2020). Second, both *plusieurs* and *différents* belong to the same part of speech (indefinite adjectives) and are used in the same syntactical position before nouns. Third, they refer to an indefinite number of things or events in contrast to other indefinite adjectives of quantity like *quelques*, which normally is used to describe a little number of things, or *nombreux* and *multiple*, which refer to big numbers of things of events. Finally, both adjectives are frequent in French with 447.03 (for *plusieurs*) and 144.79 (for *différent*) occurrences per million words in large corpora of written French (Jakubíček et al., 2013).

In total, there were 20 items, with two conditions per item (with and without list signal). Each type of list signal was inserted equally frequently across all list conditions. Participants were asked to provide a continuation with at least one sentence that had to be complete, grammatically correct, and contain at least three words. Example (9) illustrates an experimental item in the list and non-list conditions.

(9) List condition:

La comédienne a planifié <u>plusieurs</u> rendez-vous pour la journée. Elle a prévu d'aller voir son agent.

'The actress scheduled <u>several</u> appointments for the day. She planned to meet her agent.'

Non-list condition:

La comédienne se préparait à la maison. Elle a prévu d'aller voir son agent.

'The actress was getting ready at home. She planned to meet her agent.'

Coding Procedure. Continuation sentences were annotated for the analysis as *list*, *additive*, *consequence* or *other*, depending on their relation to the cue sentence. I defined as *list* continuations those sentences that contained an enumeration of one or several events related to the content of the first cue sentence. Examples (10) – (12) illustrate list continuations that participants wrote in the list condition for item (9).

(10) Et elle a prévu de passer d'autres castings.

'And she planned to do more castings.'

(11) Puis elle prévu d'aller faire un coucou à ses grand-parent[3].

'Then she planned to go and say hello to her grandparents.'

(12) Elle doit aussi aller se faire une teinture chez le coiffeur.

'She also has to get her hair dyed.'

There were several completion sentences that expressed not only list, but also temporal (11) or contrast (13) relations at the same time. In such cases, continuations were labelled as *list*, as the focal point of this set of experiments was to identify sentences conveying the idea of enumeration in relation to the prompt.

(13) Task sentences:

La journaliste a fait <u>différents</u> commentaires sur le film. Elle a apprécié le jeu de l'actrice principale.

'The journalist made <u>various</u> comments about the film. She appreciated the acting of the lead actress.'

Continuation:

[3] The faulty original spelling of participants was kept in French examples.

Elle a moins aimé la qualité des dialogues.

'She liked the quality of the dialogues less.'

I coded as *additive* continuations that provided new information or more details about the first or the second task sentences, including exemplification and sub-events, but excluding the instances of enumeration, which were included in the category of list relations. In the continuations (14) and (15), for instance, participants do not list any other activities planned by the actress for the day as in (10) – (12), but rather add a new fact about an actress (14) or her agent (15). For this reason, these continuations were labelled as additive rather than list. I did not make a further distinction about additive and elaboration relations, as it was not relevant for the present set of experiments.

(14) Task sentence (list condition):

La comédienne a planifié <u>plusieurs</u> rendez-vous pour la journée. Elle a prévu d'aller voir son agent.

'The actress scheduled <u>several</u> appointments for the day. She planned to meet her agent.'

Continuation:

Elle a reçu beaucoup d'argent.

'She received a lot of money.'

(15) Continuation:

Celui-ci a annulé au dernier moment.

'The latter cancelled at the last moment.'

When a continuation phrase described an event caused by an activity presented in the task item, it was tagged as a *consequence* relation. Example (16) illustrates a consequence continuation that was written by one of the participants after a task item in the list condition. The fact that the girl's mother gave her an ice cream is considered here to be a consequence of her good performance at school.

(16) Task sentence (list condition):

La fille a reçu plusieurs bonnes notes à l'école. Elle a réussi l'examen d'histoire.

'The girl received several good marks at school. She passed the history exam.'

Continuation:

Donc sa maman l'a récompensé avec une glace.

'So her mother rewarded her with an ice cream.'

All the remaining relations were labelled as *other*. This category included several types of discourse relations, such as temporality, contrast, cause, and goal, which were not further distinguished, as it was not essential for the goals of the present investigation.

If a participant provided several continuation sentences, discourse relations between the provided continuations were not labelled, since it was outside of the scope of the present set of experiments. The focus was on discourse relations between the prompt and the completion sentence. Out of 8114 continuations[4], 10% were annotated together by one of the authors and an independent experienced coder. It is important to mention that, when we deal with the annotation of discourse relations, multiple, non-self-excluding interpretations can be possible and not always all of these interpretations are noticed and taken into account by coders. However, the agreement rate between the two coders on this continuation sample was 95% (κ=.82; Gwet's AC_1=.92), which granted the

[4] The details about annotation are reported for the data from all three experiments.

remaining of the continuations to be divided in half and annotated independently. Note that all instances where one coder had doubts were cross-checked by the other coder.

6.3.2.2. Author recognition tests

To assess teenagers' degree of exposure to print, I used an adapted version of the Author Recognition Test (ART) (Tskhovrebova, Zufferey & Tribushinina, 2022), since the ART is not only sensitive to cultural differences (e.g., Stainthorp, 1997) but also to age (e.g., Cunningham & Stanovich, 1990). This version of the ART (*ART-F-CL*) was based on the names of French-speaking authors who are considered to be classics according to three Swiss and French national libraries and bookshop chains. The list included 40 author names and 40 names of unknown people, which were randomly mixed. The participants had to select only those names that they knew to be authors. The instruction mentioned that some of the names were not authors, and that one point would be removed if the participants checked a wrong name. For each correct answer, the participants were given *1* point, and for each wrong one *-1*. The maximum possible score therefore was *40* and the minimum *-40*, as I computed the general score summing up the points for correct and incorrect answers.

For the adult control group, I used a different version of the ART, which was developed by Zufferey and Gygax (2020a, https://osf.io/yxj8q/) and was based on the names of best-selling and prize-winning authors (*ART-F*). The ART-F replicated the design of the original English ART (Stanovich & West, 1989). The number of items and the calculation of the final score was the same as for the teenage version of the task, described before. The reliability of the tests was quite high, as indicated by their Cronbach's alphas which are close to or greater than .90 (ART-F-CL: .88, 95% CI [.85–.91]; ART-F: .92, 95% CI [.86–.94]).

The participants fulfilled the tasks always in the same order, starting with the story-continuation task and finishing with the ART. Once the participants gave an answer and proceeded to the next question, they could not go back and correct their initial response.

6.3.3. Analysis

Continuation sentences were analysed by fitting generalized mixed-effects logistic regression models on the binary variable (list versus non-list relation), using the R software (R Core Team, 2020). I tested models with the *glmer* function of the *lme4* package (Bates et al., 2015) and made model comparison with the *anova()* function, using a forward-testing approach. I added main and interaction fixed effects one at a time, and each model with an added factor was compared to a previous model that did not have the included factor. P-values of the final model were obtained with the *summary()* function from the *lmerTest* package (Kuznetsova et al., 2014). The statistical significance level was set at <5% and is indicated by bold marking in the corresponding tables. In total, I created three models: the first one only for teenage participants, the second one only for adults, and the third one for both age groups. In addition, for each separate analysis, I performed a pairwise comparison between list signal (absent versus present) and connectives used in the task with the *lsmeans()* function of the *emmeans* package in R (Lenth, 2020).

This analysis at first was performed separately for teenagers and adults, and then together for all participants. Age groups were first analysed separately, given that the primary aim was to shed light on teenage sensitivity to alternative list signals – and given that ART was different across age groups. I also present general analyses considering all groups together, yet without including ART. In order to facilitate reading, I report all the details of the model selections in the Online Appendix. Moreover, separate models for teenagers and adults are also included in the Online Appendix, since the degree of exposure to print, as measured by ART-F-CL and ART-F, did not predict the variation in the sensitivity to list signals.

6.3.4. Results and discussion

The final model included List Signal (absent versus present) and Group (teenagers versus adults) as fixed factors (both main and interaction effects), and Item and Participant as random intercepts (see Table 2 and Figure 1). The results demonstrate that both teenagers and adults were sensitive to list signals, as revealed by an estimated increase of 1.95 ± SE 0.33. However, there was a significant interaction between the factors Group and List signal, demonstrating that teenagers were on average less sensitive to list signals than adults. Finally, in the condition without list signal, the production of list continuations did not vary between the two groups of participants. The two separate analyses within each age group confirmed the effects found in the general analysis and did not reveal any significant inter-individual variation, related to the degree of exposure to print and age.

This result replicates the finding of Scholman et al. (2020) on the sensitivity to list signals, applied both to adult and young speakers of French. In the next experiment, I aim to examine further the effect from the adjectives of quantity. More precisely, I assess whether participants are still sensitive to these alternative signals, even if the task items include both adjectives of quantity and different types of connectives, which are more salient and prototypical signals of coherence relations.

Figure 1. Proportions of different types of relations in the continuation sentences in Experiment 1 (see Table S1 in Online Appendix for the exact values)

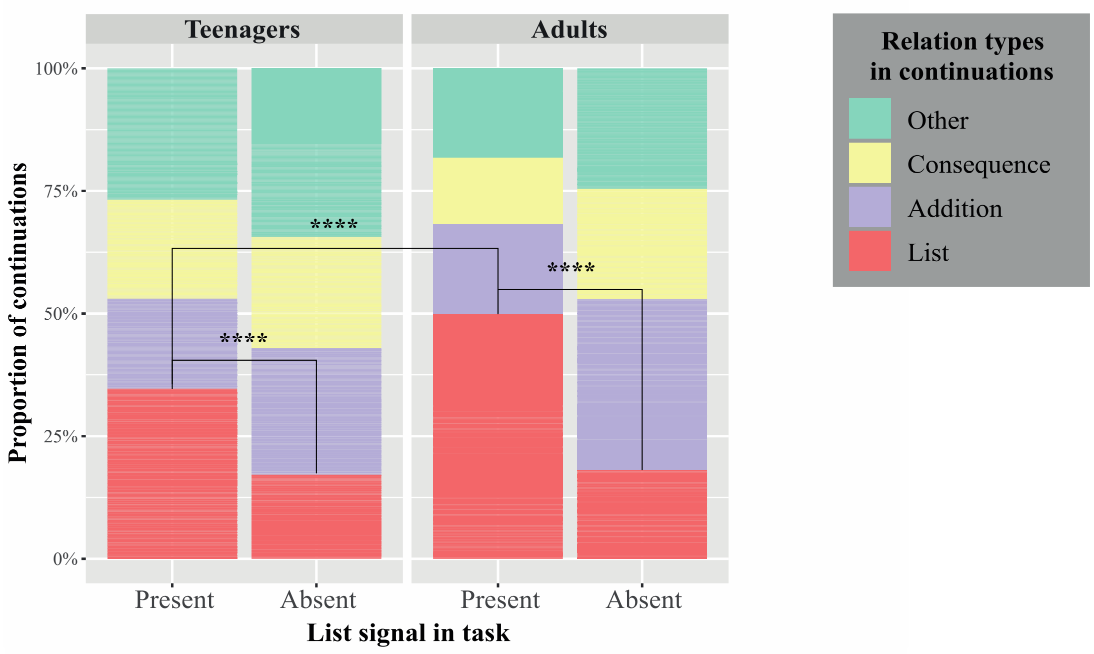


Table 2. Model's estimates for the best fitting models in Experiment 1

	Estimate	SE	z	p
	All participants			
(Intercept)	-1.92	0.36	-5.37	**<0.001**
List signal	1.95	0.33	5.85	**<0.001**
Teenagers	-0.22	0.36	-0.63	0.529
List signal*Teenagers	-0.75	0.23	-3.29	**0.001**
	Teenagers			
(Intercept)	-2.18	0.29	-7.51	**<0.001**
List signal	1.20	0.31	3.92	**<0.001**
	Adults			
(Intercept)	-1.96	0.34	-5.76	**<0.001**
List signal	2.00	0.43	4.64	**<0.001**

Note. In the model for all participants, conditional $R^2\Delta$=.36, marginal $R^2\Delta$=.08; for teenagers, conditional $R^2\Delta$=.35, marginal $R^2\Delta$=.05; for adults, conditional $R^2\Delta$=.40, marginal $R^2\Delta$=.13

6.3.5. Additional analysis of the distribution of connectives in list continuations

I noticed that in this experiment, where connectives were not included in the prompt passage, participants added their own connectives in 80% of list continuations. When I fitted the generalized mixed-effects logistic regression model on the binary variable (absence versus presence of the connective in the list continuation), adding the factors of List Signal (absence versus presence in the task item) and Group (adults versus teenagers) did not improve the model's fit (list signal: $\Delta\chi^2$ = 0.22, Δdf = 1, p < .638; group: $\Delta\chi^2$ = 0.22, Δdf = 1, p < .638). In other words, the insertion of the connective in list continuations was not predicted by the presence or the absence of the adjectives of quantity in the task sentence for both groups of participants.

However, it seems that the position in which teenagers and adults used connectives in their productions was different. In sentence-initial position, teenagers used connectives in 70% of cases and adults in 31%. In contrast, in sentence-medial or sentence-final position, teenagers included connectives only in 14% of continuations, whereas adults used them 50% of the time. Among teenagers, the most popular connective was sentence-initial *et* 'and' (48%), followed by sentence-medial *aussi* 'also' (11%), and sentence-initial *mais* 'but' (6%), and *puis* 'then' (5%). Adults used most often sentence-medial connectives *aussi* 'also' (24%) and *également* 'also' (18%), sentence-initial *et* 'and' (10%), sentence-medial *ensuite* 'then' (7%), and sentence-initial *puis* 'then' (5%).

6.4. Experiment 2

6.4.1. Participants

Fifty-four French native speaking teenagers (M_{age} = 14.44, SD = 1.62, Range: 12–17) and twenty-two adults (M_{age} = 26.10, SD = 7.17, Range: 18–43) participated in the second experiment. The recruitment modalities of both groups of participants were the same as in Experiment 1.

6.4.2. Materials and procedure

The ART tests were the same as in Experiment 1, while the story-continuation task was slightly modified. Participants were proposed to fulfil almost identical story-continuation task to the one in the first experiment, with the only difference that the second sentence was this time followed by a connective. The selected connectives *en plus* and *donc* respectively encode a relation of addition and consequence and are frequently used in French (respectively, 279.30 and 3'318.41 occurrences per million words[5]). Adding connectives allowed us to examine whether participants' sensitivity to list signals was modulated by the presence of a connective. Moreover, by including different types of connectives, I also aimed to study their effect on the generation of inference for the upcoming coherence relation. The additive connective is not in

[5] The connectives' mean frequency was calculated by averaging their frequencies in oral and written language. The frequency in oral speech was calculated based on the oral sub-corpus of *Orféo* (Benzitoun et al., 2016). The frequency in writing was obtained based on three different corpora, namely Le Monde, the French part of Europarl (Koehn, 2005), and Frantext (ATILF, 1998-2022), respectively representing journalistic, argumentative and literary genres.

contradiction with the logic of enumeration conveyed by the lexical signal, as this connective encodes a more generic additive relation, and can also introduce a list relation (more specific). The connective *en plus* was a particularly suitable candidate for this experiment, as it is frequent, monofunctional, and specialized in signalling additive coherence relations (Roze et al., 2012). In contrast, I expected that the connective of consequence should decrease the production of list continuations, as this connective cannot be used to introduce a list relation. Examples (17) and (18) illustrate the items used in Experiment 2.

(17) List condition:

La comédienne a planifié plusieurs rendez-vous pour la journée. Elle a prévu d'aller voir son agent. En plus, …

'The actress scheduled several appointments for the day. She planned to meet her agent. In addition, …'

Non-list condition:

La comédienne se préparait à la maison. Elle a prévu d'aller voir son agent. En plus, …

'The actress was getting ready at home. She planned to meet her agent. In addition, …'

(18) List condition:

Le médecin avait plusieurs lieux de travail. Il avait un cabinet à l'hôpital central. Donc, …

'The doctor had several places of work. He had an office at the central hospital. So, ...'

Non-list condition:

Le médecin était spécialisé dans les traitements contre le cancer. Il avait un cabinet à l'hôpital central. Donc, …

'The doctor specialized in cancer treatment. He had an office at the central hospital. So, …'

6.4.3. Analyses

I started by making the same statistical analysis as in Experiment 1. However, in order to compare the effects from list signals and connectives between the task without connectives (Experiment 1) and the one with frequent connectives (Experiment 2), I made an additional comparative analysis separately for each connective.

6.4.4. Results

6.4.4.1. Sensitivity to connectives and list signals in Experiment 2

Our final model included Connective (*en plus* versus *donc*) as a fixed factor and Item and Participant as random intercepts (see Table 3). This result shows that, in contrast to the connective *donc*, the additive connective *en plus* predicted a greater number of list continuations, independently of the presence of the list signal and the age group.

Figure 2. Proportions of different types of relations in the continuation sentences in Experiment 2

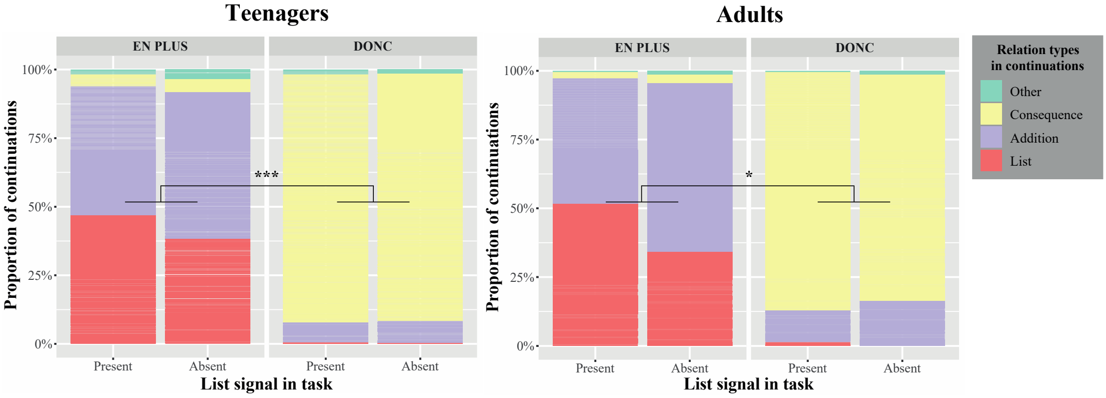


Table 3. Model's estimates for the best fitting model in Experiment 2

	Estimate	SE	z	p
All participants				
(Intercept)	-6.49	0.52	-12.42	**<0.001**
En plus	6.09	0.54	11.27	**<0.001**
Teenagers				
(Intercept)	-6.75	0.64	-10.50	**<0.001**
En plus	6.34	0.64	9.88	**<0.001**
Adults				
(Intercept)	-6.52	0.91	-7.19	**<0.001**
En plus	6.07	0.93	6.52	**0.026**

Note. In the model for all participants, conditional $R^2\Delta$=.71, marginal $R^2\Delta$=.59; for teenagers, conditional $R^2\Delta$=.74, marginal $R^2\Delta$=.63; for adults, conditional $R^2\Delta$=.75, marginal $R^2\Delta$=.56

6.4.4.2. Additional comparative analysis between Experiment 1 and Experiment 2 for the connective *en plus*

The final model for the analysis, comparing the production of list continuations after the task items without a connective (from Experiment 1) and after the items with the additive connective *en plus* (from Experiment 2), included List Signal (absent versus present), Connective (no connective versus *en plus*), and Group (adults versus teenagers) as fixed factors (both main and interaction effects), and Item and Participant as random intercepts (see Table 4 and Figure 3). The results from this analysis demonstrate that there was a main effect of List signal and of the connective *en plus* for the production of list continuations. However, when a list signal was present in the cue sentence, adults were on average more sensitive to it than teenagers. Furthermore, there was also a significant interaction between List signal and *en plus*. The post-hoc pairwise comparison revealed that there were significantly more list continuations after the sentences with list signals, both when the connective *en plus* was present (log odds ratio=0.57, SE=0.28, $p=.045$) and absent (log odds ratio=1.51, SE=0.26, $p<.0001$). As a result, there was no significant change in the production of lists after the adjectives of quantity between the sentences followed by *en plus* and the sentences not followed by a connective (log odds ratio=0.13, SE=0.25, $p<.598$). However, when the adjectives of quantity were absent in the task sentences, there was a significant increase in the number of list relations in participants' responses after the sentences including the connective *en plus* (log odds ratio=1.07, SE=0.25, $p<.0001$).

The separate models for teenagers and adults had similar effects as the general model for all participants (see Table 4). The only difference was that teenagers produced significantly more list continuations after the sentences with list signals when the connective *en plus* was not present in the task (log odds ratio=1.14, SE=0.26, $p<.0001$).

In contrast, adults wrote significantly more list continuations after the sentences with list signals both when the connective *en plus* was present (log odds ratio=0.85, SE=0.43, p=0.046) and absent (log odds ratio=1.91, SE=0.38, p <.0001). In other words, it seems that teenagers were sensitive to the adjectives of quantity only in the task items that were not followed by a connective, while adults were sensitive to the alternative signals in both conditions, independently of the connective *en plus*.

Figure 3. Proportions of different types of relations in the continuation after the task items without a connective (from Experiment 1) and after those with the connective *en plus* (from Experiment 2) across teenagers and adults

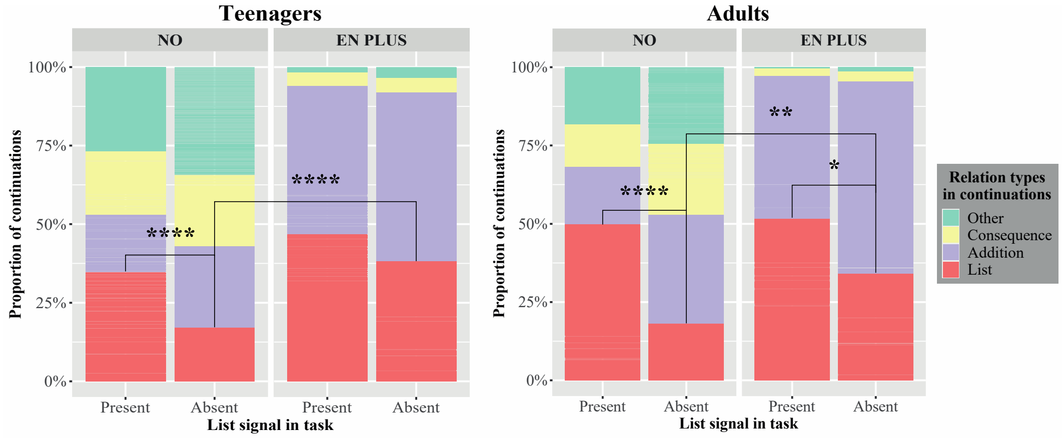


Table 4. Model's estimates for the best fitting model in the additional analysis, comparing the production of list continuations after the task items without a connective (from Experiment 1) and after those with the connective *en plus* (from Experiment 2)

	Estimate	SE	z	p
	All participants			
(Intercept)	-1.85	0.34	-5.51	**<0.001**
List signal	1.87	0.30	6.27	**<0.001**
En plus	0.87	0.42	2.08	**0.038**
Teenagers	-0.22	0.35	-0.63	0.530
List signal**En plus*	-1.10	0.30	-3.64	**<0.001**
List signal*Teenagers	-0.73	0.22	-3.25	**0.001**
En plus*Teenagers	0.40	0.49	0.82	0.413
List signal**En plus**Teenagers	0.33	0.35	0.94	0.349
	Teenagers			
(Intercept)	-2.09	0.26	-8.07	**<0.001**
List signal	1.14	0.26	4.40	**<0.001**
En plus	1.35	0.30	4.54	**<0.001**
List signal**En plus*	-0.82	0.22	-3.70	**<0.001**
	Adults			
(Intercept)	-1.89	0.31	-6.03	**<0.001**
List signal	1.91	0.38	5.02	**<0.001**
En plus	0.81	0.33	2.45	**0.015**
List signal**En plus*	-1.06	0.37	-2.83	**0.005**

Note. In the model for all participants, conditional $R^2\Delta$=.34, marginal $R^2\Delta$=.07; for teenagers, conditional $R^2\Delta$=.35, marginal $R^2\Delta$=.06; for adults, conditional $R^2\Delta$=.37, marginal $R^2\Delta$=.10

6.4.4.3. Additional comparative analysis between Experiment 1 and Experiment 2 for the connective *donc*

The final model for the analysis, comparing the production of list continuations after the task items without a connective (from Experiment 1) and after the items with the consequence connective *donc* (from Experiment 2), included List Signal (absent versus present), Connective (no connective versus *donc*), and Group (adults versus teenagers) as fixed factors (both main and interaction effects), and Item and Participant as random intercepts (see Table 5 and Figure 4). This analysis reveals that the presence of the connective *donc* in Experiment 2 significantly decreased the proportion of list continuations in comparison to the task sentences without this connective from Experiment 1. In other words, both groups of participants were responsive to list signals only after the sentences without the connective *donc*, while the presence of the consequence connective almost completely prevented participants from writing list relations in their productions. The analyses within each age group confirmed the overall effects obtained in the general analysis.

Figure 4. Proportions of different types of relations in the continuation after the task items without a connective (from Experiment 1) and after those with the connective *donc* (from Experiment 2) across teenagers and adults

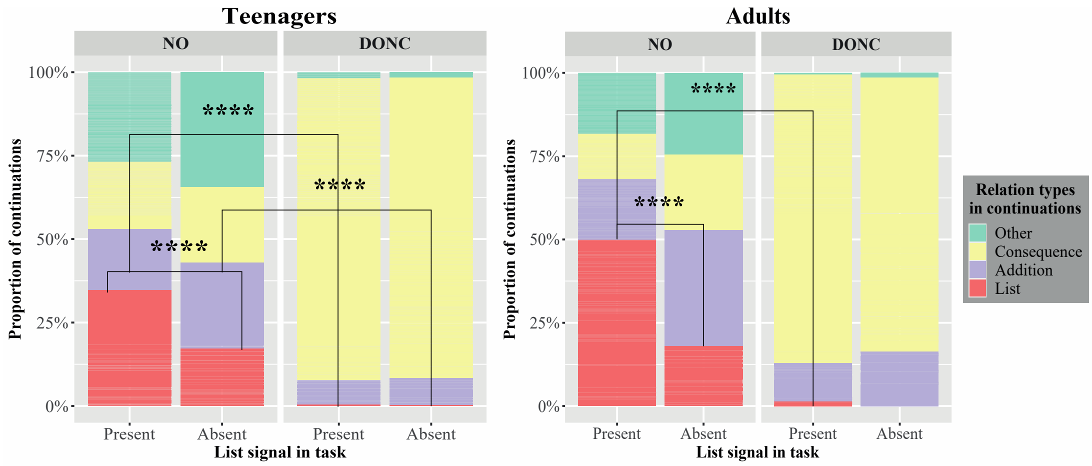


Table 5. Model's estimates for the best fitting model in the additional analysis, comparing the production of list continuations after the task items without a connective (from Experiment 1) and after those with the connective *donc* (from Experiment 2)

	Estimate	SE	z	p
	All participants			
(Intercept)	-1.92	0.36	-5.31	**<0.001**
List signal	1.95	0.34	5.80	**<0.001**
Donc	-17.05	7.89	-2.16	**0.031**
Teenagers	-0.23	0.36	-0.63	0.528
List signal* *Donc*	11.84	7.88	1.50	0.133
List signal*Teenagers	-0.75	0.23	-3.24	**0.001**
*Donc**Teenagers	12.33	7.88	1.57	0.118
List signal* *Donc* *Teenagers	-12.63	7.89	-1.60	0.109
	Teenagers			
(Intercept)	-2.18	0.29	-7.45	**<0.001**
List signal	1.20	0.30	3.94	**<0.001**
Donc	-4.88	1.06	-4.62	**<0.001**

	Estimate	SE	z	p
List signal* *Donc*	-0.70	1.25	-0.56	0.577
	Adults			
(Intercept)	-2.04	0.36	-5.59	**<0.001**
List signal	2.06	0.40	5.21	**<0.001**
Donc	-16.90	86.55	-0.20	0.845
List signal* *Donc*	11.90	86.54	0.14	0.891

Note. In the model for all participants, conditional $R^2\Delta$=.76, marginal $R^2\Delta$=.68; for teenagers, conditional $R^2\Delta$=.56, marginal $R^2\Delta$=.40; for adults, conditional $R^2\Delta$=.90, marginal $R^2\Delta$=.86

6.4.4.4. Discussion

The results of this experiment revealed that the presence of the connectives *en plus* ‘in addition’ and *donc* ‘so’ affected the sensitivity to the adjectives of quantity of French speakers. The consequence connective *donc* completely overrode the inference from the alternative signals of the list relation in both groups of participants. As for the additive connective *en plus,* the effects were not the same for the two age groups. Teenagers were sensitive to the adjectives of quantity only in the items that were not followed by the connective *en plus*, while adults remained sensitive to the alternative signals, independently of the additive connective.

In the next experiment, I examine whether the frequency of the connectives, following the task items, may be an additional factor, affecting the sensitivity to the alternative signals. More precisely, I assess whether the presence of the less frequent additive and consequence connectives *en outre* ‘in addition’ and *ainsi* ‘therefore’, respectively, would produce the same effects on the generation of list inferences as the equivalent frequent connectives. It was found in previous studies, for instance, that certain infrequent connectives, such as *en outre* ‘in addition’ and *aussi* ‘therefore’, are particularly challenging both for teenagers and adults (Tskhovrebova, Zufferey & Gygax, 2022; Zufferey & Gygax, 2020a, 2020b). This difficulty may stem from the fact that infrequent connectives are mostly used in written modality, and extensive exposure to the written language happens later than that to the oral language, coming with schooling process (Nippold, 2004, 2008). It is only starting from secondary school that teenagers become autonomous readers and start to be exposed to written texts of various genres (Nippold, 2004, 2008). As a result, connectives that appear mostly in writing, and thus have on average lower frequency, may be mastered less well than those that are often used in oral language.

Hence, in Experiment 3, I included the connectives *en outre* and *ainsi* to assess their effect of on the generation of list inferences, as these connectives are mostly used in writing and have a lower frequency. The additive connective *en outre* can be considered as equivalent to *en plus*, as it signals the same coherence relation, but has much lower average frequency (46.52 versus 279.30 occurrences per million words, respectively). The consequence connective *ainsi* can be considered as equivalent to *donc*, but it is much less frequent (178.61 versus 3'318.41 occurrences per million words, respectively). I include the connective *ainsi* instead of the previously tested *aussi*, as the latter is polyfunctional and can convey both relation of addition and that of consequence (Roze et al., 2012). Including two monofunctional connectives (*en outre* and *ainsi*) allowed us to disentangle two coherence relations and avoid possible confusions.

6.5. Experiment 3

6.5.1. Participants

In the third experiment, I recruited 50 French native speaking teenagers (M_{age} = 14.34, SD = 1.94, Range: 12–19) and 21 adults (M_{age} = 28.64, SD = 10.43, Range: 20–57). The recruitment process of both groups of participants were the same as in Experiment 1.

6.5.2. Materials and procedure

The ART tests were again the same as in Experiment 1, while the story-continuation task slightly differed. Experiment 3 was almost identical to Experiment 2, and differed only in the choice of connectives. Instead of more frequent connectives, the cue passage included one of the two less frequent connectives, namely *en outre* ‘in addition’ and *ainsi* ‘therefore’.

6.5.3. Analyses

Statistical analyses were the same as in Experiment 2.

6.5.4. Results

6.5.4.1. Sensitivity to connectives and list signals in Experiment 3

The final model for all participants included List Signal, Connective (*en outre* versus *ainsi*), and Group as fixed factors (main and interaction effects), Item and Participant as random intercepts, and Connective as random slope by Participant (see

Table 6 and Figure 5). This result shows that, similar to the Experiment 2, the additive connective *en outre* predicted a greater number of list continuations than the consequence connective *ainsi.* However, in contrast to the Experiment 2, teenagers on average wrote fewer list continuations after *en outre* than adults. The separate analyses for teenagers and adults confirmed the trends from the general analysis and did not reveal variation, predicted by the ARTs.

Figure 5. Proportions of different types of relations in the continuation sentences in Experiment 3

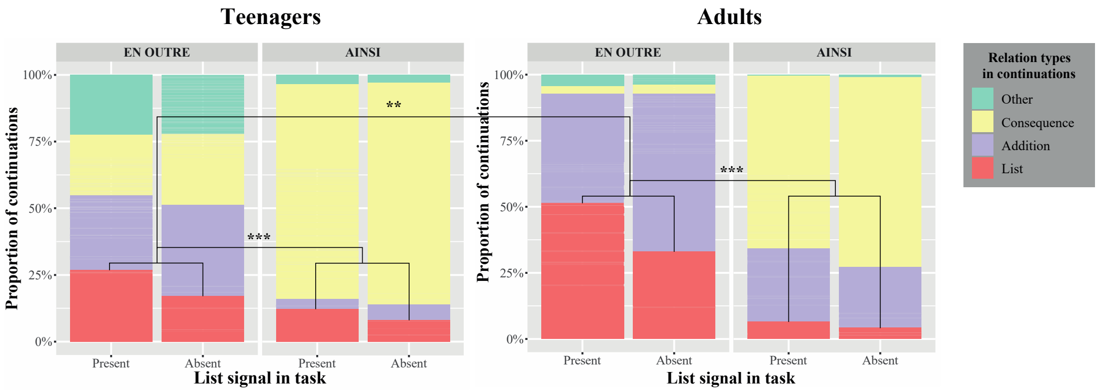


Table 6. Model's estimates for the best fitting model in Experiment 2

	Estimate	SE	z	p
	All participants			
(Intercept)	-7.23	1.33	-5.43	**<0.001**
En outre	6.31	1.35	4.66	**<0.001**
List signal	1.45	0.91	1.59	0.112
Teenagers	2.40	1.24	1.93	**0.054**
*En outre**List signal	-0.44	0.97	-0.45	0.655
*En outre**Teenagers	-3.46	1.27	-2.73	**0.006**
List signal*Teenagers	-0.73	0.92	-0.80	0.427
*En outre**List signal*Teenagers	0.39	0.97	0.40	0.691

	Estimate	SE	z	p
	Teenagers			
(Intercept)	-4.45	0.67	-6.61	**<0.001**
En outre	2.50	0.71	3.51	**<0.001**
List signal	0.70	0.38	1.85	0.064
*En outre**List signal	-0.05	0.49	-0.10	0.923
	Adults			
(Intercept)	-10.43	2.46	-4.24	**<0.001**
En outre	9.37	2.50	3.75	**<0.001**
List signal	1.84	1.15	1.60	0.109
*En outre**List signal	-0.67	1.27	-0.53	0.597

Note. In the model for all participants, conditional $R^2\Delta$=.66, marginal $R^2\Delta$=.29; for teenagers, conditional $R^2\Delta$=.50, marginal $R^2\Delta$=.13; for adults, conditional $R^2\Delta$=.35, marginal $R^2\Delta$=.17

6.5.4.2. Additional comparative analysis between Experiment 1 and Experiment 3 for the connective *en outre*

The final model for the analysis, comparing the production of list continuations after the task items without a connective (from Experiment 1) and after the items with the additive connective *en outre* (from Experiment 3), included List Signal (absent versus present), Connective (no connective versus *en outre*), and Group (adults versus teenagers) as fixed factors (both main and interaction effects), and Item and Participant as random intercepts (see Table 7 and Figure 6). This analysis showed that both groups of participants overall produced more list continuations after the items including the adjectives of quantity. Moreover, teenagers were on average less responsive to the presence of list signals than adults across both experiments.

The separate analysis within the group of teenagers showed that teenagers produced significantly more list continuations after the sentences with list signals when the connective *en outre* was absent in the task (log odds ratio=1.16, SE=0.27, p <.0001). In addition, the presence of the alternative signals and the additive connective *en outre* significantly decreased the production of lists in comparison to the sentences that included only the alternative signals (log odds ratio=-0.69, SE=0.29, p=0.016). In contrast, the analysis within the group of adults demonstrated that adults wrote significantly more list continuations after the sentences with list signals both when the connective *en outre* was present (log odds ratio=0.85, SE=0.43, p=0.046) and absent (log odds ratio=1.91, SE=0.38, p <.0001). However, the proportion of lists in the adult productions did not significantly change between the sentences with the connective *en outre* and those without any connective, both when adjectives of quantity were present (log odds ratio=-0.21, SE=0.49, p=0.662) and absent (log odds ratio=0.73, SE=0.39, p=0.061) in the task items.

To summarize, similarly to the Experiment 2, teenagers were more sensitive to the adjectives of quantity in the task items that were not followed by a connective, while adults were sensitive to the alternative signals in both conditions, independently of the connective *en outre.* Moreover, the presence of the connective *en outre* together with the alternative signals significantly reduced the proportion of list productions by teenagers, but did not affect the proportion of lists, produced by adult speakers.

Figure 6. Proportions of different types of relations in the continuation after the task items without a connective (from Experiment 1) and after those with the connective *en outre* (from Experiment 3) across teenagers and adults

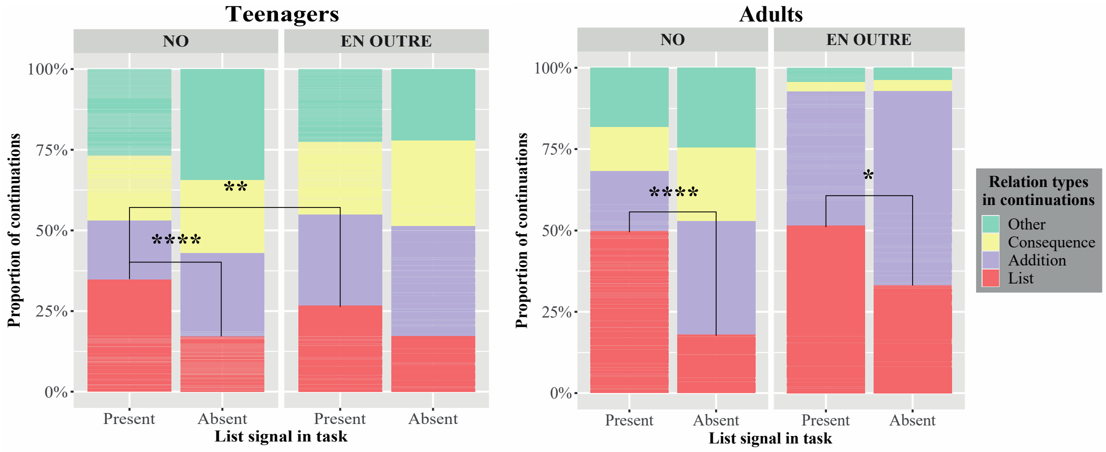


Table 7. Model's estimates for the best fitting model in the additional analysis, comparing the production of list continuations after the task items without a connective (from Experiment 1) and after those with the connective *en outre* (from Experiment 3)

	Estimate	**SE**	**z**	**p**
	All participants			
(Intercept)	-1.86	0.33	-5.57	**<0.001**
List signal	1.89	0.30	6.30	**<0.001**
En outre	0.78	0.42	1.86	0.063
Teenagers	-0.21	0.34	-0.62	0.535
List signal**En outre*	-1.00	0.31	-3.18	**0.001**
List signal*Teenagers	-0.73	0.22	-3.24	**0.001**
En outre *Teenagers	-0.85	0.50	-1.69	0.091
List signal**En outre* *Teenagers	0.35	0.37	0.95	0.342
	Teenagers			
(Intercept)	-2.09	0.26	-8.18	**<0.001**
List signal	1.16	0.27	4.35	**<0.001**
En outre	-0.01	0.30	-0.02	0.983

	Estimate	SE	z	p
List signal**En outre*	-0.69	0.25	-2.79	**0.005**
	Adults			
(Intercept)	-1.93	0.34	-5.61	**<0.001**
List signal	1.97	0.40	4.97	**<0.001**
En outre	0.73	0.39	1.88	0.061
List signal**En outre*	-0.94	0.40	-2.33	**0.020**

Note. In the model for all participants, conditional $R^2\Delta$=.66, marginal $R^2\Delta$=.29; for teenagers, conditional $R^2\Delta$=.28, marginal $R^2\Delta$=.03; for adults, conditional $R^2\Delta$=.41, marginal $R^2\Delta$=.10

6.5.4.3. Additional comparative analysis between Experiment 1 and Experiment 3 for the connective *ainsi*

The final model for the analysis, comparing the production of list continuations after the task items without a connective (from Experiment 1) and after the items with the additive connective *ainsi* (from Experiment 3), included List Signal (absent versus present), Connective (no connective versus *ainsi*), and Group (adults versus teenagers) as fixed factors (both main and interaction effects), and Item and Participant as random intercepts (see Table 8 and Figure 7). This analysis shows that the presence of the alternatives signals significantly increased the production of lists for all the participants. However, overall, the presence of the consequence connective *ainsi* almost completely prevented participants from writing list continuations. The two separate within-group analyses confirmed general trends revealed in the analysis for all participants. The only difference was that when the consequence connective *ainsi* was present in the task sentences, adult speakers were not sensitive to the alternative list signals (log odds ratio=1.09, SE=0.81, p=0.176). In contrast, teenagers responded to the presence of the adjectives of quantity and produced slightly more lists even in the presence of the consequence connective *ainsi* (log odds ratio=0.78, SE=0.40, p=0.049).

I noticed however that not all participants who produced list continuations after the connective *ainsi* interpreted it as a consequence connective. Out of 115 continuations, *ainsi* was treated as a connective of consequence in only 11 of them. In the other 104 continuations, participants started their sentence with *que* and, this way, used it as an additive conjunction *ainsi que* 'as well as'. In other words, some participants changed the connective intended in the task. As a result, it is complicated to interpret the effects of *ainsi* as a connective of consequence on list inference generation.

Figure 7. Proportions of different types of relations in the continuation after the task items without a connective (from Experiment 1) and after those with the connective *ainsi* (from Experiment 3) across teenagers and adults

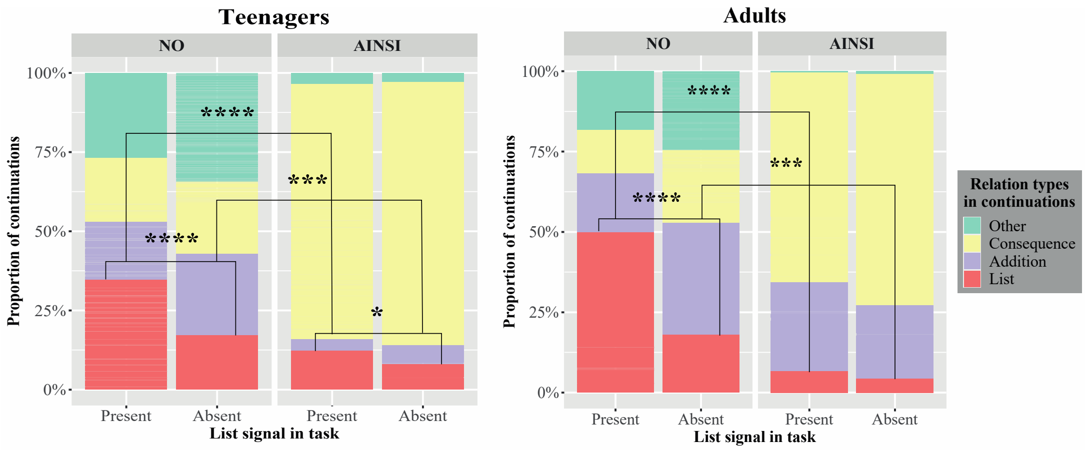


Table 8 Model's estimates for the best fitting model in the additional analysis, comparing the production of list continuations after the task items without a connective (from Experiment 1) and after those with the connective *ainsi* (from Experiment 3)

	Estimate	SE	z	p
	All participants			
(Intercept)	-1.93	0.45	-4.28	**<0.001**

	Estimate	**SE**	**z**	**p**
List signal	1.96	0.32	6.11	**<0.001**
Ainsi	-3.50	0.90	-3.91	**<0.001**
Teenagers	-0.31	0.49	-0.64	0.525
List signal**Ainsi*	-0.84	0.74	-1.13	0.261
List signal*Teenagers	-0.74	0.23	-3.18	**0.001**
Ainsi *Teenagers	2.09	0.98	2.14	**0.032**
List signal**Ainsi* *Teenagers	0.37	0.80	0.47	0.641
	Teenagers			
(Intercept)	-2.23	0.32	-6.90	**<0.001**
List signal	1.20	0.28	4.26	**<0.001**
Ainsi	-1.49	0.47	-3.16	**0.002**
List signal**Ainsi*	-0.42	0.33	-1.26	0.209
	Adults			
(Intercept)	-2.09	0.48	-4.37	**<0.001**
List signal	2.12	0.39	5.37	**<0.001**
Ainsi	-3.31	0.95	-3.48	**<0.001**
List signal**Ainsi*	-1.02	0.76	-1.34	0.180

Note. In the model for all participants, conditional $R^2\Delta$=.48, marginal $R^2\Delta$=.17; for teenagers, conditional $R^2\Delta$=.43, marginal $R^2\Delta$=.09; for adults, conditional $R^2\Delta$=.62, marginal $R^2\Delta$=.32

6.5.4.4. Comparative analysis between Experiment 2 and Experiment 3

In order to see whether connectives with different frequencies had a different impact on the generation of list inference, I performed an analysis contrasting the results from Experiment 2, which included more frequent connectives, and from Experiment 3, which assessed less frequent connectives. However, given the issue in the interpretation of results after the connective *ainsi*, I excluded all the results for both connectives of consequence (*ainsi* and *donc*) from this analysis and focused only on the two connectives of additive relations (*en plus* and *en outre*).

The statistical procedure remained the same as in previous analyses. I also made three separate models for different groups of participants (for teenagers, adults, and all participants together) and reported the details of model selection for teenagers and adults in the Online Appendix (see Table S9). Similar to previous analyses, the measures of exposure to print did not predict the variation in the sensitivity to list signals. Finally, treatment contrasts were applied to the factor of Connective, where *en plus* was set as a reference level for comparison in all three models.

The final model for all participants included List Signal, Connective, and Group as fixed factors (both main and interaction effects), and Item and Participant as random intercepts (see Table 9). Comparing the results from all participants revealed that teenagers produced significantly fewer list continuations than adults after the prompt including the connective *en outre*. However, all other interactions were not statistically significant. The separate within-group analyses demonstrated that connective frequency played a role only for the group of teenagers, as they produced significantly fewer list continuations after the less frequent connective *en outre* than after the more frequent *en plus*, both when list signals were absent (log odds ratio=1.30, SE=0.28, z=4.67, p=<.0001) or present (log odds ratio=1.13, SE=0.27, z=4.16, p=<.0001) in the task. As for the group of adults, the frequency of connectives did not affect the proportion of list continuations.

Table 9. Model's estimates of the best fitting models in the analysis, comparing the task with the more frequent connective *en plus* 'in addition' (from Experiment 2) and the task with less frequent connective *en outre* 'in addition' (from Experiment 3)

	Estimate	**SE**	**z**	***p***
	All participants			
(Intercept)	-0.82	0.32	-2.56	**0.011**
List signal	0.88	0.31	2.85	**0.004**
En outre	-0.09	0.41	-0.21	0.830
Teenagers	0.19	0.33	0.57	0.569
List signal**En outre*	0.11	0.32	0.34	0.735
List signal*Teenagers	-0.41	0.26	-1.56	0.119
*En outre**Teenagers	-1.24	0.49	-2.53	**0.011**
List signal**En outre**Teenagers	0.06	0.39	0.16	0.870

Note. In the model for all participants, conditional $R^2\Delta$= .30, marginal $R^2\Delta$= .07

6.5.4.5. Discussion

The results of the Experiment 3 were similar to those from the Experiment 2. It was shown that the presence of the consequence connective *ainsi*, similar to the more frequent consequence connective *donc*, almost completely overrode the inference from the alternative signals of the list relation in both groups of participants. As for the additive connective *en outre,* the effects again were not the same for the two age groups. Teenagers were sensitive to the adjectives of quantity only in the task items that were not followed by the additive connective *en outre*, while adults remained sensitive to the alternative signals, independently of the additive connective. In general, the presence of *en outre* significantly decreased the production of list continuations in teenagers in comparison to the sentences not followed by any connective and to those followed by the more frequent additive connective *en plus*.

To sum up, the findings from Experiments 2 and 3 demonstrated that the combination of list signals with additive connectives did not significantly increase the production of list continuations, but rather decreased (*en outre*) or left unchanged (*en plus*). Given that these connectives signal a more generic additive relation, they can be used to express the relation of list, but are not limited to it. As a result, when a connective expressing a more generic additive relation is used together with an alternative signal of a more specific list relation, it does not significantly improve the inference for a more specific list signal. This effect may stem from the fact that the inference of a more generic additive relation, coming from a more salient and monofunctional signal such as connective, competes with the inference of the list relation, coming from a less prominent and non-monofunctional alternative signal. I make in the next section an additional analysis aiming to assess whether participants were more sensitive to the additive connectives and produced significantly more additive continuations in the conditions that included additive connectives *en plus* and *en outre*.

6.6. Analysis of additive continuations after the sentences with additive connectives *en plus* and *en outre*

6.6.1. Comparative analysis between Experiment 1 and Experiment 2 for the connective *en plus*

In order to examine whether more additive continuations were produced after the items including the additive connective *en plus*, I created a statistical model, comparing the proportion of additive continuations after the sentences without any connective (from Experiment 1) and those followed by the connective *en plus* (from Experiment 2). The results of both age groups were analysed together, as I did not need to include the measures of exposure to print in the analysis. The details on model selections can be found in the Online Appendix also for this analysis.

Results show that both groups of participants indeed produced significantly more additive continuations after the sentences containing the additive connective *en plus* than after the sentences without any connective (see Table 10 for the model's estimates and Figure 8). The sensitivity to the frequent additive connective *en plus* was not significantly different between the two age groups. Moreover, in the sentences without any connective, participants produced more additive continuations when the adjectives of quantity were absent (log odds ratio=0.73, SE=0.25, p=0.003). In the sentences including the additive connective, the presence of adjectives of quantity did not affect the proportion of additive continuations (log odds ratio= 0.47, SE=0.28, p=0.086). Finally, when both types of signals were absent in the task sentences, adults on average wrote more additive continuations than teenagers (log odds ratio= 0.48, SE=0.21, p=0.022).

Figure 8. Proportions of different types of relations in the continuation after the task items without a connective (from Experiment 1) and after those with the connective *en plus* (from Experiment 2) across teenagers and adults

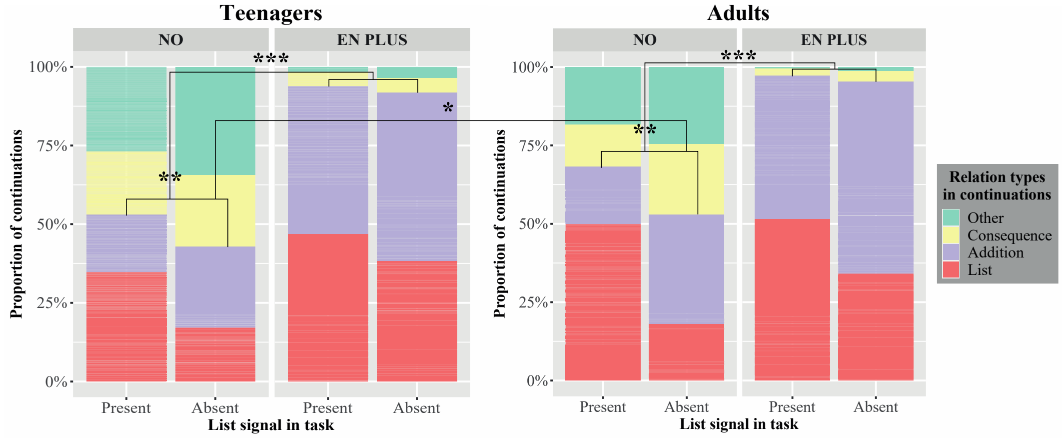


Table 10. Model's estimates of the best fitting model in the analysis, comparing additive continuations after the task items without a connective (from Experiment 1) and after those with the connective *en plus* (from Experiment 2)

	Estimate	**SE**	**z**	***p***
(Intercept)	-0.82	0.30	-2.70	**0.007**
List signal	-1.00	0.31	-3.18	**0.001**
En plus	1.34	0.38	3.55	**<0.001**
Teenagers	-0.57	0.29	-1.98	**0.048**
List signal**En plus*	0.27	0.34	0.79	0.427
List signal*Teenagers	0.52	0.22	2.35	**0.019**
*En plus**Teenagers	0.18	0.42	0.42	0.672
List signal**En plus**Teenagers	-0.07	0.34	-0.20	0.838

Note. Conditional $R^2\Delta$ =.31, marginal $R^2\Delta$ = .10

6.6.2. Comparative analysis between Experiment 1 and Experiment 3 for the connective *en outre*

In order to examine whether more additive continuations were produced after the items including the additive connective *en outre*, I created a statistical model, comparing the proportion of additive continuations after the sentences without any connective (from Experiment 1) and those followed by the connective *en outre* (from Experiment 3).

Results show that there were also significantly more additive continuations after the sentences containing the additive connectives *en outre* than after the sentences without connectives (see Table 11 for the model's estimates and Figure 9). However, adults were more sensitive to the less frequent additive connective *en outre*, as they produced significantly more additive sentences than teenagers after the task items including this connective (log odds ratio=1.03, SE=0.26, $p<0.001$). Finally, as in the analysis for the connective *en plus*, after the items without any connective, participants produced more additive continuations when the adjectives of quantity were absent (log odds ratio= 0.76, SE=0.26, $p=0.004$). In contrast, after the sentences including the additive connective *en outre*, the presence of adjectives of quantity did not affect the proportion of additive continuations (log odds ratio=0.53, SE=0.29, $p=0.073$).

Figure 9. Proportions of different types of relations in the continuation after the task items without a connective (from Experiment 1) and after those with the connective *en outre* (from Experiment 3) across teenagers and adults

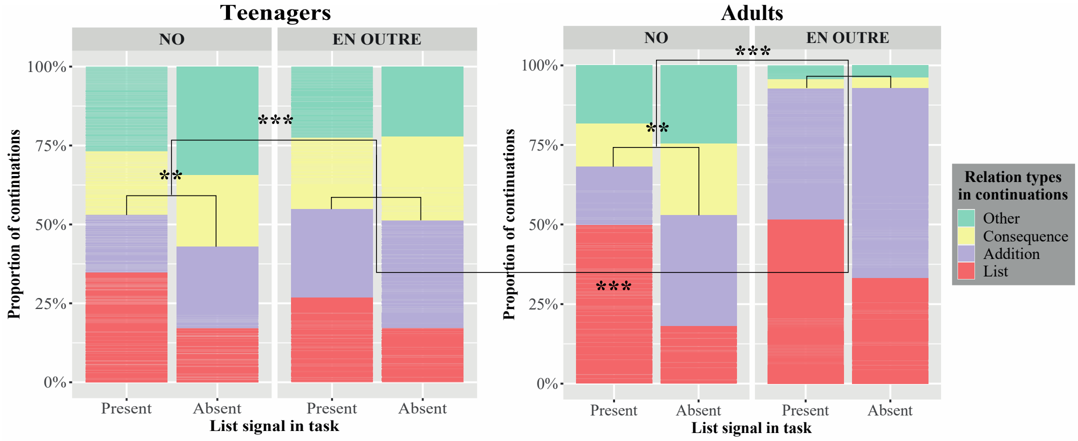


Table 11. Model's estimates of the best fitting model in the analysis, comparing additive continuations after the task items without a connective (from Experiment 1) and after those with the connective *en outre* (from Experiment 3)

	Estimate	**SE**	**z**	***p***
(Intercept)	-0.79	0.28	-2.81	0.005
List signal	-1.01	0.30	-3.37	**<0.001**
En outre	1.18	0.34	3.47	**<0.001**
Teenagers	-0.56	0.27	-2.10	0.036
List signal**En outre*	0.16	0.31	0.53	0.597
List signal*Teenagers	0.51	0.22	2.34	0.019
*En outre**Teenagers	-0.79	0.40	-1.99	0.047
List signal**En outre**Teenagers	0.13	0.36	0.37	0.709

Note. Conditional $R^2\Delta$ =.25, marginal $R^2\Delta$ =.05

6.7. Qualitative analysis of data from the three experiments

6.7.1. Distribution of elliptic continuation sentences

Across all experiments and conditions, teenagers produced more elliptic continuations that lacked subject or verb. Out of 5604 continuations written by teenagers, 670 (12%) were elliptic; whereas only 38 (2%) of the 2509 completions created by adults had an ellipse. Most ellipses were found in list continuations across both age groups (503 (75%) in teenagers and 36 (95%) in adults). Note that some participants analysed *ainsi* not as a connective of consequence, but as an additive connective *ainsi que*, by adding *que* in their continuation sentence (see example 19). Since all such instances were elliptic, this accounted for most elliptic sentences produced by adults and an important part of ellipses produced by teenagers (see Table 12). However, even when no connective was present in the prompt, the proportion of elliptic sentences written by teenagers was still greater than that of adults (54% versus 25%).

(19) Task sentences:

L'enfant a surpris ses parents. Il voulait comprendre pourquoi le ciel était bleu. Ainsi, …

‘The child surprised his parents. He wanted to understand why the sky was blue. Therefore, …’

Continuation:

que pourquoi la neige est-t-elle blache.

‘as well as why the snow is white.’

Table 12. Raw number (and proportion) of elliptic sentences in list continuations across all three experiments and all age groups

	No connective	***Ainsi***	***Donc***	***En outre***	***En plus***	**Total N**
Teenagers	271 (.54)	86 (.17)	0	11 (.02)	135 (.27)	503
Adults	9 (.25)	21 (.58)	0	0	6 (.17)	36

6.7.2. Distribution of list signal types per experimental condition

I also looked at the percentage of list continuations across two types of list signals (i.e., *plusieurs* and *différent*) to check whether different signals generated a different number of list continuations. It seems that on average, across all three experiments, when a connective was present in the cue sentence, the proportion of list continuations was quite similar for both signal types across all age groups (see Table 13). The distribution appears to be slightly different for task sentences followed by additive connectives and for the sentences without connectives in the cue passage. In such cases, after the signal *différent*, there was a greater percentage of list continuations, and it seems that the difference between the two signal types was more pronounced for teenagers than for adults. The effects of the signal type should, however, be interpreted with caution, as the design of the three experiments did not allow a systematic assessment of this variable.

Table 13. Proportion of list continuations per signal type among teenagers and adults

	Teenagers	Adults	Total
	Without connective		
Différents	.55	.52	.54
Plusieurs	.45	.48	.46
	En plus		
Différents	.53	.47	.51
Plusieurs	.47	.50	.48
	Donc		
Différents	0	0	0
Plusieurs	.1	.3	.1
	En outre		
Différents	.37	.45	.41
Plusieurs	.30	.43	.35
	Ainsi		
Différents	.16	.05	.11
Plusieurs	.17	.07	.13

6.8. General discussion

In the current set of experiments, I examined whether native French-speaking teenagers were sensitive to signals of list relation, expressed by the adjectives of quantity *plusieurs* ‘several’ and *différents* ‘various’ (Experiments 1, 2 & 3). I also assessed whether this sensitivity was modulated by the presence of another signal of coherence relation, namely connectives of additive and consequence relations, varying in frequency (Experiments 2 & 3). Finally, I systematically contrasted the results obtained by teenagers with those of a control group of adults, and assessed whether their performance in the main task was modulated by their linguistic competence, as measured by the author recognition test.

6.8.1. Sensitivity to alternative signals of list relation

Both groups of participants were sensitive to list signals, as they produced more continuations expressing a list relation when one of the adjectives of quantity was present in the first sentence of the task that did not include connectives (Experiment 1). However, teenagers' receptiveness to alternative list signals was still inferior to that of adults. This finding might indicate that sensitivity to alternative signals develops with age and the increasing linguistic experience that is normally associated to it. It is possible that teenagers are less sensitive to alternative signals than adults because they have not yet mastered non-sentence-initial usage of coherence markers. Indeed, when teenagers used connectives in their own productions, they preferred to use them in sentence-initial position and only rarely used them in other positions. In contrast, adults produced connectives in different syntactic positions and even did so more frequently in non-sentence initial positions.

In addition, the fact that linguistic experience and level of linguistic proficiency develop with age is reflected in the types of continuations produced by teenagers and adults, as teenagers were more likely than adults to provide elliptic sentences. This may of course indicate that teenagers took the task less seriously and paid less attention to it. However, it may also mean that they have not yet mastered well all the particularities of written language, which namely tends to avoid ellipses (Menzel, 2016). Another indication of the fact that teenagers may not master the written modality is the usage of connective *et* ‘and’ in sentence-initial position produced in their own sentences. Whereas in oral speech it is perfectly normal to use this connective in the beginning of the sentence, in written language it is not stylistically appropriate, as coordinating conjunctions are not possible in sentence-initial position according to reference grammars (Riegel et al., 2021).

The distribution of list continuations after each of the list signals suggested that the adjective *différents*, on average, was associated with more list inferences than adjective *plusieurs*. The difference between two list signals was more pronounced for teenagers than for adults, especially in the task without connectives and after the prompt passages followed by the additive connectives *en plus* and *en outre*. This might indicate that the usefulness of alternative signals to indicate coherence relation may vary across lexical items. However, this observation should be further studied in experiments that will be specifically designed to test the effect of different types of list signals, as the current set of experiments compared only two words.

6.8.2. Sensitivity to list signals combined with connectives

When the task combined both alternative signals and connectives, I found different effects in the production of list continuation sentences. First of all, the difference between the proportion of list continuations after the sentences including and not including the list signal was not the same within three experiments. After the cue sentences with connectives *en plus* and *en outre*, teenagers and adults produced more list continuations when a list signal was present than when it was absent. However, the observed effects in teenagers were not as strong as those in adults, suggesting that teenagers are probably even less sensitive to the alternative list signals when a more salient signal like connective is also present in the sentence.

The presence of list signals together with the connectives of consequence *donc* and *ainsi* did not have any effect on the generation of list inference. Indeed, after the task passages followed by the connectives of consequence*,* the list relation was almost completely absent in the continuation sentences produced by both teenage and adult participants. Presumably, this means that connectives signalling the relation of consequence create a much stronger mental inference of this relation than do the alternative list signals for the relation of list. However, the results for *ainsi* should be considered with caution, as in a significant number of cases, it was interpreted as a different type of signal (*ainsi que*), used for marking addition.

Secondly, I found different effects on the production of list relations not only within, but also between the experiments. I observed for instance that in the condition without list signal, there were significantly more list continuations after *en plus* in comparison to the task without connectives in all age groups. In other words, this means that even alone the additive connective *en plus* can generate inference of the list relation. However, when both *en plus* and the list signal were present in the cue sentence, it did not significantly reinforce the inference of a list relation. As demonstrated the analysis of additive continuations, *en plus* can generate not only an inference of list relation, but also that of an additive relation. Therefore, when both types of signals are present in the sentence, the additional additive function of the connective may compete with the inference of list relation from the alternative signal. Alternatively, and in line with findings of Crible and Demberg (2021), this effect may be due to a stronger inference power of connectives as a type of coherence signal compared to list signals within the related segments.

In the condition without list signal, after the more infrequent additive connective *en outre,* the proportion of list continuations produced by teenagers was the same as in the task without connectives; while when combined with the list signals, there were even fewer list continuations in comparison to the same condition in the task without connectives. As for adults, although they produced slightly more list sentences after the connective *en outre*, their proportion was not significantly higher than in the task without connectives in both conditions (with and without list signal). As far as the comparison of connectives was concerned, teenagers produced significantly more list continuations after the more frequent additive connective *en plus* than after the infrequent connective *en outre*. In contrast, there was no such difference between the effects of the two additive connectives for adults. Taken together, these findings suggest that *en outre* does not facilitate the inference of the list relation and may even hinder this inference, especially in the case of young speakers. Indeed, teenagers may be less familiar with the less frequent connective *en outre.* Hence, it is less easy for them to infer a more specific list relation. In addition, similar to the connective *en plus*, a more generic additive meaning triggered by *en outre* may override the more specific list meaning. Nevertheless, the fact that an important number of list relations was produced even in the presence of more salient, stronger, and prototypical signals of coherence such as the additive connectives *en plus* and *en outre*, shows that alternative list signals are an important source for inferring a list relation. These signals start to be perceived and to affect discourse inferences as early as at the age of 12 and their impact increases with age.

In contrast to Scholman et al. (2020), I did not find an effect of the author recognition test on the sensitivity to alternative list signals both for teenagers and adults. Although the French versions of the ART were strong predictors for the use of connectives in other studies (Tskhovrebova, Zufferey & Tribushinina, 2022; Zufferey & Gygax, 2020b), it probably requires further validation in French. As a matter of fact, the French version of this test included 80 items, while the English ART, used by Scholman et al. (2020), consisted of 130 items, which might have rendered this version a more sensitive measure. Moreover, the overall performance of both groups of participants was not very high on the measures of exposure to print (teenagers: *M*=6.44, SD=6.08, observed range: -11 to 28, possible range: -40 to 40; adults: *M*=8.89, SD=5.43, observed range: -1 to 23, possible range: -40 to 40). This may have created a floor effect that did not allow us to track individual variation. Finally, the lack of effect of ART scores in the present experiments may also suggest that exposure to print does not necessarily reflect individual differences in the ability to infer an intended coherence relation. It is possible that this ability constitutes a specific type of linguistic competence that should be assessed with a more sensitive measure.

6.9. Conclusion

Taken together, the results presented in this chapter suggest that expressions of quantity, such as *plusieurs* 'several' and *different* 'various', are an important source for the inference of the list relation as early as in teenage years, even though the sensitivity to these alternative signals still develops into adulthood. The fact that the combination of an alternative signal with the additive connective *en plus* did not significantly increase the inference of a list relation in both age groups indicates that a more generic additive relation, signalled by this connective, may compete with a more specific relation of list. Furthermore, it seems that the inference of list relation in teenagers is inhibited by a less frequent additive connective *en outre*, and is almost completely hindered by both types of consequence connectives. Ultimately, the degree of exposure to print, as measured by the ART in the presented data, does not predict the individual differences in the sensitivity to the adjectives of quantity as signals of the list relation. More globally, the presented set of experiments shows that the examination of how different types of coherence signals combine with each other opens many new avenues of enquiry for future research. This type of research sheds light onto the linguistic devices that can reinforce or inhibit the generation of a certain coherence relation, and thus, allows to understand the functioning of this relation better.

Nonetheless, the present experiments had several limitations that should be taken into account in follow-up research. It is important to point out that I examined the effect of alternative signals and connectives on production data and can only speculate about the comprehension level of the coherence relations included in this study. Moreover, one of the most important limitations is related to the design of the experiments, as they involved between-participant design. Future research should therefore address the issue of the design and focus on comprehension measures in order to complement the present findings.

As for the interpretations of the results on the interaction between adjectives of quantity and additive connectives, it should be noticed that since the relation of list is a subtype of the relation of addition, it is possible that in some continuations both relations simply co-existed, without necessarily competing with each other. Furthermore, the analysis of the connective insertions in the participants' productions have hinted that, perhaps, temporal connectives, such as *ensuite* and *puis*, may be even better suited for marking list relations and should be analysed in future studies.

Finally, an important contribution to future research would be to unveil other types of alternative signals that can generate coherence inferences when used alone or together with connectives, and to continue the examination of other linguistic competences that may better explain individual variation in speakers' sensitivity to alternative signalling.

In the next chapter, I pursue the examination of factors affecting the use of connectives during teenage years. Chapter 4 revealed that the general level of the mastery of monofunctional connectives is high, independently of how frequently these connectives are used in oral speech and in writing. However, the fact that the additive connective *en outre*, which is infrequent and mostly used in formal, administrative, and written contexts, received significantly lower scores than other connectives among teenagers and adults shows that some connectives are still challenging even for adults. In the next chapter, I aim to unveil other potential sources of difficulty for the mastery of certain connectives, by assessing the use of really infrequent written connectives and the connectives that convey less usual coherence relations, in contrast to those presented so far. Moreover, I also investigate whether the development of the competence with such connectives can be accelerated through different types of training.

7. The Role of Training for the Mastery of Connectives in French[1]

In Chapters 4 and 5, I showed that the overall level of competence with connectives in a controlled environment is quite high independently from the mode in which these connectives are typically used. However, previous experiments also revealed that there are still some connectives that represent a particular difficulty for speakers of different ages, such as the French connectives *en outre* ‘moreover’ (Chapters 3, 4 & 6) and *aussi* ‘therefore’ (Chapter 3). This finding suggests that there are probably other potential sources of difficulty, which have not been detected or fully examined so far, that may render certain connectives challenging even for adult speakers.

One of such factors may be the frequency of a connective in language use. As suggested by the scores obtained in the study presented in Chapter 3, the less frequent

[1] Results from this study have been submitted for publication in Tskhovrebova, Wetzel, Gygax and Zufferey (2023, submitted).

connectives *en outre* and *aussi* remain difficult for teenagers and adults, even in the controlled context of cloze tests. These low scores call for a more in-depth study of connectives with comparable or even lower frequencies, in order to determine whether it was the frequency of *en outre* and *aussi* that affected the performance of participants, or whether other characteristics of these connectives, such as a restricted context of use for *en outre* and polyfunctionality for *aussi*, played a role and were confounded with the factor of frequency in that particular experiment.

The studies presented in previous chapters also demonstrated that the type of coherence relation is not a strong predictor of the ability to use connectives in cloze tests. This is likely due to the fact that by the age 12, all the common coherence relations have already been acquired, and the difference of cognitive complexity between them is not reflected in off-line measures. Another important limitation from previous studies is that the connectives tested so far signalled such coherence relations as addition, cause, consequence, temporality, contrast, and concession, which are prototypical, exist in many different languages, and are listed and described by various frameworks of coherence relations, even though the way they are labelled may differ (see, e.g., PDTB, Prasad et al., 2008; SDRT, Asher & Lascarides, 2003; RST DT, Carlson et al., 2003). One of the factors of difficulty that may affect the use of connectives is whether they convey a coherence relation that is less prototypical. These relations are often not straightforward to describe, and may even not have a designated label. For instance, the major inventory of the French connectives Lexconn (Roze et al., 2012) does not provide a label for the connectives *au fur et à mesure que* and *suivant que.* Indeed, these connectives signal relations that do not appear to correspond to any common category. *Au fur et à mesure que* signals a cause that intensifies with time, as in example (1); and *suivant que* is used to introduce a series of at least two conditions, as in example (2).

(1) L'économie du pays se détériore au fur et à mesure que les mois avancent.

'The country's economy is deteriorating CONNECTIVE the months go by.'

(2) La taxation change suivant que la personne a une voiture à essence ou électrique.

'The taxation changes CONNECTIVE the person has a petrol or an electric car.'

In contrast, the connective *or* has a categorisation in Lexconn, but it does not represent quite precisely one of the functions of this polyfunctional connective. One function of *or* is to mark concession, a quite prototypical relation, but another function is to signal a logical continuation during narration as in (3), which is more specific than the more conventional additive relation.

(3) Le médecin a peur des virus. Or il y a beaucoup de patients grippés en ce moment.

'The doctor is afraid of viruses. CONNECTIVE there are a lot of flu patients at the moment.'

Moreover, the three connectives mentioned not only signal less prototypical coherence relations, but also are quite unique, as they cannot be substituted by equivalent and more frequent connectives. This is not the case of more common coherence relations that often can be expressed by several connectives, varying in frequency. It is therefore possible that connectives marking less prototypical coherence relations may require additional cognitive effort, and may therefore be more difficult to use and understand, especially for less experienced speakers who may not have had ample exposure to them.

Another question that I aim to address in the current chapter is whether and how it is possible to improve the ability to use connectives in writing. Written language is different from oral speech in numerous respects, including the functioning of connectives, as a larger variety of connectives with more precise functions is used in written discourse than in oral speech (see, e.g., Crible & Cuenca, 2017). Hence, learning to use connectives in writing may be similar in some way to learning a second language that possesses its own features, and for which exposure is limited and linked to a context of formal instruction in the classroom.

Research on second language learning often refers to the concepts of implicit and explicit knowledge that are involved in the process of language learning (see, e.g., Ellis, 1994; Ellis et al., 2009; Rebuschat, 2015, VanPatten & Smith, 2022; Wong & Simard, 2015). Implicit knowledge is an underlying knowledge of language with all its components, such as phonology, morphology, lexicon, syntax, and pragmatics. This knowledge is called implicit because it is usually unconscious and cannot be easily explained by speakers, despite the fact that they are able to use it. The implicit knowledge, or mental representation, of language is formed based on linguistic input. In consequence, to enhance the process of acquisition of implicit knowledge about language, it is important to receive a sufficient amount of input combined with instruction, helping to extract and process data from input (see, e.g., Wong & Simard, 2015).

The explicit knowledge is a conscious awareness of how certain elements of the language work, and hence this type of knowledge involves metalinguistic reasoning. As suggested by Ellis (1994), it is possible that explicit knowledge about the functioning of language, received through instruction in the form of rules, may transform into an implicit knowledge. For this transformation to happen, readers should practice how to apply these rules, so that the process of application of the rules becomes automatic. Providing readers with an instruction about the functioning of connectives may help them to better identify the connectives in the input, to create the relationship between their form and function, and with practise to internalise their functioning in the mental representation. Therefore, I hypothesise that after participants undergo training about the functioning of connectives, they may not yet completely internalise the received knowledge, but they should probably better apply the explained rule in the experimental context and as a result should have higher scores on the cloze tests fulfilled after the training.

I also suggest that the performance with connectives in the cloze test should become even better, if the proposed training involves a more active engagement with the rule. Indeed, the explanation of how connectives function in discourse may not be sufficient for retention of new information about connective functioning. According to the Involvement Load Hypothesis (ILH) introduced by Hulstijn and Laufer (2001), learning new lexicon depends on the degree of involvement in its processing. The concept of involvement includes three main components, such as need, i.e., the source of motivation to fulfil a task, search of the meaning of an unknown word, and evaluation of how the word fits its context. The greater the involvement load, the better the retention of an unfamiliar word (see Liu & Reynolds, 2022, for a review of studies supporting this hypothesis). Based on the logic of the ILH, I expect that a more active engagement with the rule, involving search for the meaning of a connective and providing feedback on whether it was used correctly or not, should result in a better performance on the cloze test than just a passive reading of the rule.

Yet, the amount of input is still the primary source that enables readers to build an implicit knowledge about language, especially in L1 (see, e.g., VanPatten & Smith, 2022). Therefore, I expect that readers with a greater degree of exposure to written language input, as measured by the Author Recognition Test (ART) (Stanovich & West, 1989), should on average have a more advanced general competence with connectives, independently of the training activity, as their knowledge of connectives functioning should be better internalised and automatised.

7.1. Method

All materials, data, and code are available on the OSF repository[2].

7.1.1. Participants

A group of 228 native French-speaking teenagers aged 12 to 20 (*M*= 15.81, SD= 2.14) participated in this study. The experiment was held in the classes of four schools in the French-speaking part of Switzerland. In order to determine the baseline of competence with connectives, I also recruited 60 adult native French speakers aged 19 to 46 (M_{age}= 28.22, SD= 6.53) via the crowdsourcing platform Prolific© (Prolific, Oxford, UK, www.prolific.co). The informed consent for participation in the study was provided by adult participants as well as by all the teachers responsible for teenagers.

7.1.2. Materials and procedure

7.1.2.1. Sentence cloze test

Choice of connectives. Two types of infrequent connectives were selected for this study. The first type consists of connectives that do not have more frequently used synonyms and that signal uncommon coherence relations, such as logical continuation (*or*), cause that intensifies with time (*au fur et à mesure que*), and series of conditions (*suivant que*). The second type includes connectives that are unfrequent in corpus data but that have more frequent synonyms and signal more common relations (see, e.g.,

[2] https://osf.io/gqpy7/?view_only=1f466fe2e9cc42aba5b53b721d9475ba

Prasad et al., 2008), such as consequence (*aussi*), exception (*hormis que*), causality (*étant donné que*), addition (*en outre*), and condition (*dans la mesure où*). Table 1 summarises different types of connectives used in this study.

The connective frequencies were calculated based on a large web-crawled corpus of French (Jakubíček et al., 2013). Three connectives, namely *dans la mesure où, aussi,* and *or*, had two functions, and it is their non-dominant functions that were assessed in the current study. The conditional function of *dans la mesure où* represents about 20% of the occurrences in corpora in contrast to the causal one, which represents all the other uses of this connective. The consequence function of *aussi* has approximately 10% of the occurrences in corpora in comparison to its dominant function of addition (the distribution is taken from Zufferey & Gygax, 2020a). For the connective *or*, the function of logical continuation constitutes about 30% of occurrences compared to the concessive one. The frequencies of *dans la mesure où* and *or* were calculated by randomly selecting 100 occurrences and manually annotating each function. The proportion of each function was extrapolated to the whole corpus. Since *aussi* can express consequence only in sentence-initial position, I used this syntactical constraint to search for the frequency of this function directly in the main corpus, without making the additional annotation.

Table 1. Summary of the assessed connectives

Connective	Relation	Approximate translation in English	Frequency per million words	A more frequent synonym
		Common coherence relations		
dans la mesure où	condition	in so far as	4.15	*si* ‘if’ (1,457.7)
aussi	consequence	therefore	27	*donc* ‘so’ (348.02)
hormis que	exception	except	0.11	*sauf que* ‘except’ (15.35)
étant donné que	cause	given that	6.4	*parce que* ‘because’ (294.07)
en outre	addition	moreover	49.6	*en plus* ‘also’ (118.49)
		Uncommon coherence relations		
or	logical continuation	no equivalent	27.88	–
au fur et à mesure que	a cause that intensifies with time	as	4.45	–
suivant que	series of at least two conditions	depending on	0.76	–

Structure of the test. The participants were asked to fill in gaps between two sentences with an appropriate connective. Instead of commonly used punctuation marks such as commas or full stops, the gap between sentences was demarcated by double slashes ‘//________//’ so that punctuation between the two sentences did not affect the choice of a connective. The test included 5 sentences per connective, amounting to a total of 40 sentences. To ensure that only one option was possible, not all connectives were always proposed as answer options. For instance, *aussi* ‘therefore’ was never proposed among the answer options for the sentences targeting the additive connective *en outre*, because *aussi* can signal both consequence and additive relations. Similarly, *dans la mesure où* ‘in so far as’, which can mark both condition and cause, was not proposed among answers for the sentences targeting the causal connective *étant donné que* ‘given that’. The sentences that targeted connectives encoding common coherence relations had four answer options, randomly selected from the five connectives marking these relations included in the experiment (*aussi*, *hormis que*, *étant donné que*, *en outre*, and *dans la mesure où*). The sentences targeting the connectives signalling uncommon coherence relations included only three answer options, namely the connectives *or, au fur et à mesure que,* and *suivant que*, so that connectives belonging to the same category of uncommon coherence relations were the only possible choices for these connectives.

Participants completed three cloze tests over three sessions, so that it was possible to assess whether their performance on these tests was affected by a training activity carried out in-between session 1 and 2. The difference between session 2 and 3 was meant to assess whether improvements would be lasting in time. The second session was held one week after the first one. It included a training activity on the use of the tested connectives, followed by the cloze test, evaluating the immediate effect of the training on the performance with connectives. The third session included only the cloze test and took place four weeks after the second one in order to assess a long-term effect of the training on the use of connectives. For each participant, the sentences included in the cloze test differed between the three sessions (i.e., were randomly selected) so that they were always actively engaged in the task and could not simply reproduce previously given answers.

7.1.2.2. Training session

One of two types of trainings was randomly allocated to participants. The first type of training was passive. It included a rule describing the function of a connective, followed by three example sentences with the same connective, as illustrated in example (4).

(4) Example of a passive training for the connective *dans la mesure où*

The rule:

Le connecteur *dans la mesure où* peut introduire une **condition** et est utilisé entre deux phrases de la même manière que le connecteur ***si***.

'The connective *dans la mesure où* can introduce a condition and is used between two sentences in the same way as the connective *si*'

Examples:

- Tu peux manger ce dessert **dans la mesure où (= si)** tu en laisses un morceau à ton frère.

  'You can eat this dessert in so far as (=if) you leave a piece for your brother.'
- Le port du masque restera obligatoire dans des espaces publics **dans la mesure où (= si)** la situation sanitaire ne se normalise pas.

  'The wearing of masks will remain mandatory in public spaces in so far as (= if) the health situation does not normalise.'
- Tout le monde peut publier son contenu sur ce site **dans la mesure où (= si)** ce contenu n'incite pas à la violence.

  'Anyone can publish their content on this website in so far as (= if) this content does not promote violence.'

When a connective had a more frequent synonym, it was mentioned in the rule and in the examples. This training was labelled as passive because participants simply had to read the information about each connective one after the other, without actively engaging with it.

The second type of training was active. Participants had to first deduce the function of a connective and, where possible, elicit synonyms, by answering step-by-step questions. Secondly, after deducing the rule, they had to put this rule in practice and choose an appropriate continuation for a sentence followed by a connective, as in example (5a). There were two sentences to complete, each of which was followed by the correct answer (5b).

(5) Example of an active training for the connective *dans la mesure où*

a) You will see a sentence followed by a connector and three continuation options. Please choose the option that best fits the content.

Le port du masque restera obligatoire dans des espaces publics dans la mesure où …

'The wearing of masks will remain mandatory in public spaces in so far as...'

- la situation sanitaire ne se normalise pas.

'the health situation does not normalise.'

- personne n'est plus contaminé par le virus.

'no one is infected with the virus anymore.'

- il faut qu'on en achète encore trois boîtes.

'we need to buy three more boxes.'

b) The correct answer:

Le port du masque restera obligatoire dans des espaces publics dans la mesure où la situation sanitaire ne se normalise pas[3].

'The wearing of masks will remain mandatory in public spaces in so far as the health situation does not normalise.'

[3] Bolding, italics, and highlights in the examples correspond to the style of the materials, provided to participants.

7.1.2.3. Author recognition tests

To measure the degree of exposure to written language input, I use two French versions of the Author Recognition Test (ART) (see Stanovich & West, 1989, for the original English ART). Teenagers' degree of exposure to print was assessed with an adapted version of the Author Recognition Test (*ART-F-CL*), introduced in Section 4.1.2.2. An adult version of the ART (*ART-F*), developed by Zufferey and Gygax (2020a), was used to measure the adults' level of exposure to print. The reliability of the two ART tests was quite high (like in previous studies), as indicated by their Cronbach's alphas (ART-F-CL: .88, 95% CI [.86–.90]; ART-F: .90, 95% CI [.84–.93]).

All the tasks were administered online via a weblink. During the first session, the participants started with the connective cloze test and proceeded to the ART. During the second session, the participants fulfilled the training activity and afterwards did the second cloze test. Finally, during the third session, they just completed the third connective cloze test (see Table 2 for the distribution of different tasks between the three sessions). Once the participants gave an answer and proceeded to the next question, they could not go back and correct their initial response. There was no time limit for each session, but they had to be completed without interruptions.

Table 2. Distribution of tasks between the three experimental sessions

	Session 1 (week 1)	**Session 2 (week 2)**	**Session 3 (week 6)**
1	Cloze test	Training activity	Cloze test
2	ART	Cloze test	–

7.1.3. Analysis strategy

The accuracy of responses (*1*=right, *0*=wrong) in the cloze tests was analysed with a generalised mixed-effects logistic regression model in the R software (R Core Team, 2020). First, the results of all participants were analysed together in order to assess whether training affected performance on the cloze tests between the three sessions, and whether teenagers overall had a different performance than adults. Second, two separate analyses were performed for teenagers and adults to assess the role of inter-individual variation in exposure to print within each group. Separate analyses were performed, as teenagers and adults completed different and adapted versions of the ART. After assessing the general progression in the performance on the cloze tests over three sessions and across the different groups, additional analyses were conducted in order to examine whether one type of training (active or passive) better predicted the performance of participants on the cloze tests during Sessions 2 (short term) and 3 (long term).

The models were built with the *glmer* function of the *lme4* package (Bates et al., 2015) and the model comparison was done with the *anova()* function, using a forward-selection approach. When comparing models, I assessed the contribution of random slopes to the models' fit by using log likelihood tests, when the random slopes were justified by the design (as suggested by Barr et al., 2013). The p-values of the final model were obtained with the *summary()* function from the *lmerTest* package (Kuznetsova et al., 2014).

7.2. Results

7.2.1. All participants (without the ARTs)

7.2.1.1. Performance with connectives between the three sessions

The model fit kept improving after adding Connective ($\Delta\chi^2$=91.39, Δdf=7, p<.001), a two-way interaction between Connective and Session (Session 1, 2 & 3) ($\Delta\chi^2$=185.39, Δdf=16, p<.001), and Group (Adults versus Teenagers) ($\Delta\chi^2$=129.59, Δdf=1, p<.001) as fixed factors. Since the category of coherence relation (common versus uncommon) is nested within connectives, it is the variable Connective that was included in the final model. Finally, including Connective as a random slope by Participant prevented the model from converging. As a result, the final model included Connective, Session, and Group as fixed factors, Item and Participant as random intercepts. Treatment contrasts were applied for the unordered factors of Connective and Group. The causal connective *étant donné que* was set as reference level for comparing the scores associated to the different connectives, and Adults were set as reference for the categorical variable Group.

The results revealed that teenagers performed on average significantly lower than adults across all connectives (see Table 3 for the estimates and Figure 1 for the visualisation of the results). However, two connectives received particularly low mean scores across the three experimental sessions, namely the connectives *aussi* 'therefore' ($M_{Teenagers}$=.38, 95% CI [.26, .50]; M_{Adults}=.81, 95% CI [.71, .91]) and *en outre* 'moreover' ($M_{Teenagers}$=.38, 95% CI [.26, .50]; M_{Adults}=.66, 95% CI [.54, .78]). For all the other connectives, teenagers scored between .71 and .84, and adults between .91 and .97.

As for the difference between the three experimental sessions, the post hoc pairwise comparison (see Table 4 for the statistics) demonstrated that the participants overall scored significantly higher between Sessions 1 and 2 only for the connectives *aussi* ‘therefore’, *en outre* ‘moreover’, and *or* (see Table 5 for the mean scores per connective). In contrast, the scores of all the other connectives significantly decreased, except for the connective *suivant que* ‘depending on’, which received a similar score.

However, the comparison of the scores between Sessions 1 and 3 revealed that the long-term effect of training remained only for the connective *or* that also received higher scores during the last session than during the first session. The scores for all the other connectives remained unchanged (*aussi* ‘therefore’*, dans la mesure où* ‘in so far as’*, hormis que* ‘except’*, au fur et à mesure que* ‘as’) or decreased (*étant donné que* ‘given that’, *en outre* ‘moreover’, *suivant que* ‘depending on’).

Figure 1. Distribution of mean scores per connective in sentence cloze test across the three sessions among teenagers and adults

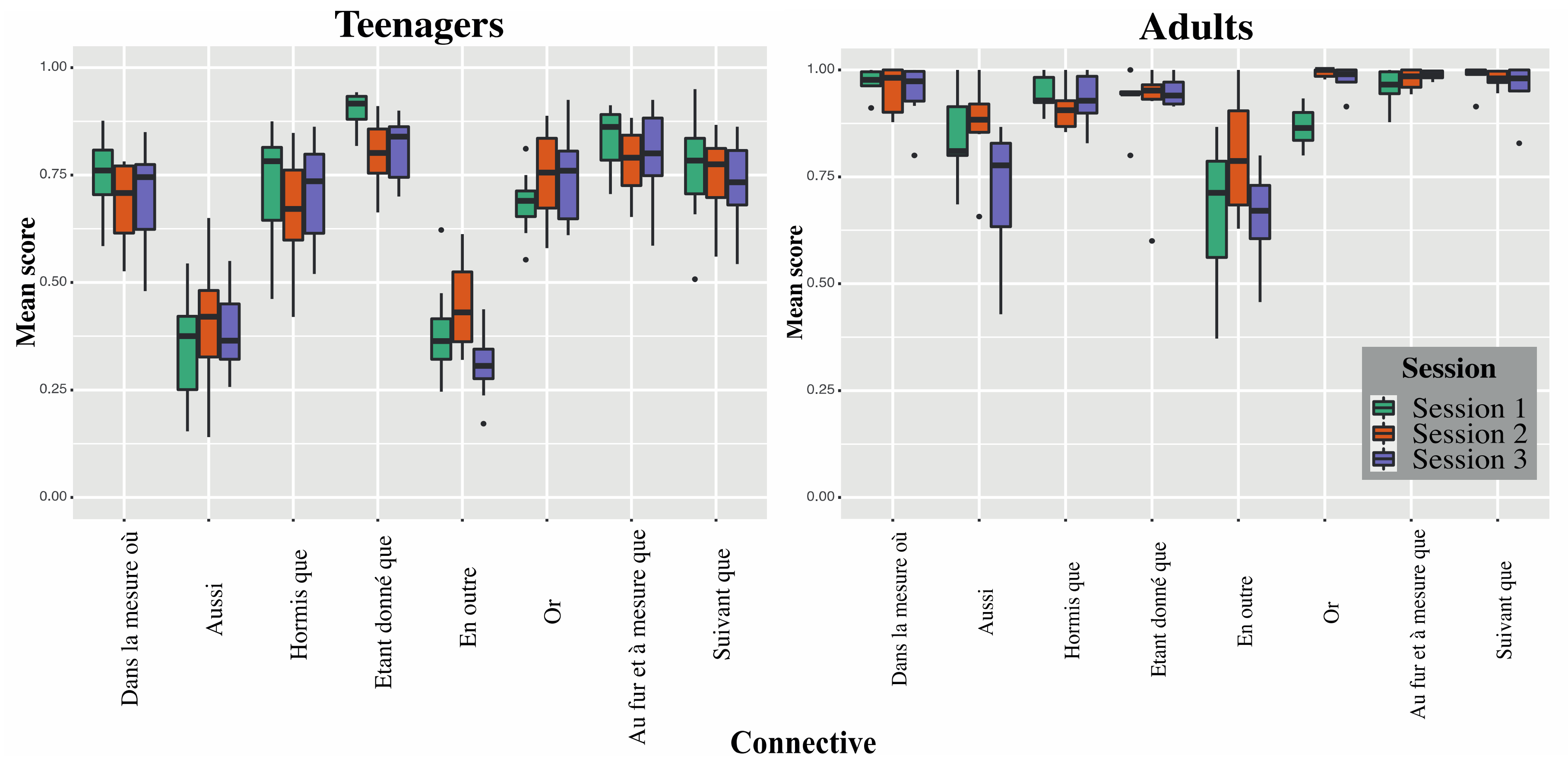


Table 3. Estimates for the best fitting model for all participants

	Estimate	SE	z	p
	All participants			
(Intercept)	4.16	0.21	20.27	<0.001
CONNECTIVE				
aussi 'therefore'	-2.95	0.22	-13.60	<0.001
dans la mesure où 'in so far as'	-1.09	0.22	-4.94	<0.001
en outre 'moreover'	-3.09	0.22	-14.22	<0.001
hormis que 'except'	-1.25	0.22	-5.68	<0.001
or	-1.53	0.22	-6.98	<0.001
au fur et à mesure que 'as'	-0.53	0.22	-2.37	.018
suivant que 'depending on'	-0.90	0.22	-4.04	<0.001
SESSION				
Session 2	-0.73	0.13	-5.71	<0.001
Session 3	-0.62	0.13	-4.66	<0.001
TEENAGERS	-1.79	0.14	-12.75	<0.001
CONNECTIVE*SESSION				
aussi * Session 2	0.97	0.16	6.21	<0.001
dans la mesure où * Session 2	0.44	0.16	2.66	.008
en outre * Session 2	1.14	0.16	7.30	<0.001
hormis que * Session 2	0.50	0.16	3.07	.002
or * Session 2	1.42	0.17	8.59	<0.001
au fur et à mesure que * Session 2	0.48	0.17	2.75	.006
suivant que * Session 2	0.68	0.17	4.04	<0.001
aussi * Session 3	0.64	0.16	3.98	<0.001
dans la mesure où * Session 3	0.43	0.17	2.52	.012
en outre * Session 3	0.29	0.16	1.77	.076
hormis que * Session 3	0.60	0.17	3.57	.000
or * Session 3	1.17	0.17	6.93	<0.001

Table 4. Pairwise comparison between connectives and experimental sessions

	Estimate	SE	z	p
Etant donné que				
Session 1 - Session 2	0.73	0.13	5.71	<0.001
Session 1 - Session 3	0.62	0.13	4.66	<0.001
Session 2 - Session 3	-0.12	0.12	-0.99	.324
Aussi				
Session 1 - Session 2	-0.23	0.09	-2.65	.008
Session 1 - Session 3	-0.02	0.09	-0.23	.819
Session 2 - Session 3	0.21	0.09	2.37	.018
Dans la mesure où				
Session 1 - Session 2	0.30	0.10	2.90	.004
Session 1 - Session 3	0.19	0.11	1.81	.071
Session 2 - Session 3	-0.11	0.10	-1.03	.303
En outre				
Session 1 - Session 2	-0.41	0.09	-4.58	<0.001
Session 1 - Session 3	0.33	0.09	3.56	<0.001
Session 2 - Session 3	0.74	0.09	8.02	<0.001
Hormis que				
Session 1 - Session 2	0.23	0.10	2.33	.020
Session 1 - Session 3	0.02	0.10	0.15	.882
Session 2 - Session 3	-0.22	0.10	-2.15	.032
Or				
Session 1 - Session 2	-0.69	0.10	-6.59	<0.001
Session 1 - Session 3	-0.55	0.11	-5.26	<0.001
Session 2 - Session 3	0.13	0.11	1.20	.230
Au fur et à mesure que				
Session 1 - Session 2	0.26	0.12	2.19	.029
Session 1 - Session 3	0.08	0.12	0.69	.488
Session 2 - Session 3	-0.17	0.12	-1.46	.145
Suivant que				
Session 1 - Session 2	0.05	0.11	0.49	.625
Session 1 - Session 3	0.24	0.11	2.17	.030
Session 2 - Session 3	0.18	0.11	1.71	.088

Table 5. Mean correctness scores per connective, experimental session, and group

	Teenagers			Adults		
	Mean	**SD**	**95% CI**	**Mean**	**SD**	**95% CI**
			dans la mesure où			
Session 1	0.74	0.44	[0.63, 0.85]	0.96	0.20	[0.90, 1]
Session 2	0.71	0.45	[0.59, 0.82]	0.92	0.28	[0.85, 0.99]
Session 3	0.71	0.45	[0.59, 0.82]	0.96	0.20	[0.90, 1]
			aussi			
Session 1	0.36	0.48	[0.24, 0.48]	0.83	0.37	[0.74, 0.93]
Session 2	0.40	0.49	[0.27, 0.52]	0.87	0.34	[0.78, 0.95]
Session 3	0.39	0.49	[0.27, 0.51]	0.75	0.43	[0.64, 0.86]
			hormis que			
Session 1	0.72	0.45	[0.61, 0.84]	0.92	0.27	[0.86, 0.99]
Session 2	0.69	0.46	[0.57, 0.81]	0.91	0.29	[0.83, 0.98]
Session 3	0.73	0.45	[0.61, 0.84]	0.91	0.29	[0.84, 0.98]
			étant donné que			
Session 1	0.89	0.31	[0.81, 0.97]	0.95	0.22	[0.89, 1]
Session 2	0.80	0.40	[0.70, 0.90]	0.95	0.21	[0.90, 1]
Session 3	0.83	0.38	[0.74, 0.93]	0.94	0.24	[0.88, 1]
			en outre			
Session 1	0.38	0.49	[0.26, 0.51]	0.65	0.48	[0.53, 0.78]
Session 2	0.45	0.50	[0.33, 0.58]	0.74	0.44	[0.63, 0.85]
Session 3	0.30	0.46	[0.19, 0.42]	0.62	0.49	[0.50, 0.74]
			or			
Session 1	0.68	0.47	[0.57, 0.80]	0.89	0.32	[0.81, 0.97]
Session 2	0.78	0.41	[0.68, 0.89]	0.99	0.10	[0.96, 1]
Session 3	0.77	0.42	[0.66, 0.87]	0.97	0.16	[0.93, 1]
			au fur et à mesure que			
Session 1	0.84	0.37	[0.75, 0.93]	0.95	0.23	[0.89, 1]
Session 2	0.80	0.40	[0.70, 0.90]	0.97	0.17	[0.93, 1]
Session 3	0.82	0.39	[0.72, 0.92]	0.99	0.12	[0.96, 1]
			suivant que			
Session 1	0.77	0.42	[0.66, 0.88]	0.99	0.10	[0.96, 1]
Session 2	0.76	0.42	[0.66, 0.87]	0.97	0.16	[0.93, 1]
Session 3	0.74	0.44	[0.63, 0.85]	0.95	0.22	[0.89, 1]

7.2.1.2. The role of the type of training

Analysis for Session 2. The correctness of responses during Session 2 was analysed in order to examine whether the type of training (active or passive) fulfilled by participants predicted the performance on the cloze test, completed just after the training. The model's fit improved after adding Connective ($\Delta\chi^2$=58.98, Δdf=7, p<.001) and then Correctness Score during Session 1 ($\Delta\chi^2$=45.75, Δdf=1, p<.001). In contrast, after including Type of Training (active versus passive), it did not show a better fit ($\Delta\chi^2$=0.636, Δdf=1, p=.425). Given that adding Group and random slopes prevented the model from converging, the final model included Connective and Correctness Score during Session 1 as fixed factors and Item and Participant as random intercepts (see Table 6).

Analysis for Session 3. The correctness of responses during Session 3 was analysed to verify whether the type of training had a long-term effect, or a "sleeping" effect (meaning that it did not occur during Session 2, but could appear during Session 3), and predicted the performance on the cloze test, fulfilled four weeks after the training activity. The final model was similar to the one comparing Sessions 1 and 2 and included Connective and Correctness Score during Session 1 as fixed factors and Item and Participant as random intercepts (see Table 6). The model's fit kept improving after first including Connective ($\Delta\chi^2$=74.21, Δdf=7, p<.001) and then Correctness Score during Session 1 ($\Delta\chi^2$=33.94, Δdf=1, p<.001). Similarly, the model did not show a better fit after I added Type of Training ($\Delta\chi^2$=0.09, Δdf=1, p=.769), and stopped converging when Group and random slopes were included.

Results. The results revealed that following an active or a passive training activity did not predict the performance in the cloze tests during Sessions 2 and 3. To put it differently, the type of training activity did not affect the use of connectives neither immediately after the training nor four weeks after. It was the type of connective that explained most of the variation in the performance on the cloze tests. Finally, these analyses also showed that participants who initially scored better during Session 1, continued having higher scores also during Sessions 2 and 3 (see Table 6 for the statistics of an estimated increase).

Table 6. Estimates of the final models, comparing the contribution of the type of training between Sessions 1, 2, and 3

	Estimate	SE	z	p
Session 1 versus Session 2				
(Intercept)	1.80	0.23	7.80	<0.001
aussi 'therefore'	-1.97	0.29	-6.81	<0.001
dans la mesure où 'in so far as'	-0.61	0.29	-2.10	0.035
en outre 'moreover'	-1.90	0.29	-6.57	<0.001
hormis que 'except'	0.01	0.29	0.02	0.984
or	-0.69	0.29	-2.39	0.017
au fur et à mesure que	-0.03	0.29	-0.12	0.909
suivant que	-0.18	0.29	-0.63	0.526
Correctness score in Session 1	0.45	0.07	6.86	<0.001
Session 1 versus Session 3				
(Intercept)	1.93	0.23	8.22	<0.001
aussi 'therefore'	-2.27	0.30	-7.65	<0.001
dans la mesure où 'in so far as'	-0.71	0.30	-2.38	0.018
en outre 'moreover'	-2.81	0.30	-9.41	<0.001
hormis que 'except'	-0.03	0.30	-0.11	0.916
or	-0.67	0.30	-2.26	0.024
au fur et à mesure que	-0.28	0.30	-0.94	0.349
suivant que	-0.54	0.30	-1.80	0.072
Correctness score in Session 1	0.40	0.07	5.91	<0.001

7.2.2. Teenagers

The fit of the model, analysing teenagers alone, kept improving after adding Connective ($\Delta\chi^2$=96.81, Δdf=7, p<.001), a two-way interaction between Connective and ART-F-CL ($\Delta\chi^2$=130.9, Δdf=8, p<.001), and finally Session ($\Delta\chi^2$= 6.49, Δdf=2, p=.039). Since adding random slopes did not allow the model to converge and including Age did not improve the model's fit ($\Delta\chi^2$=0.79, Δdf=1, p=.375), the final model for teenagers included Connective, ART-F-CL, and Session as fixed factors and Item and Participant as random intercepts (see Table 7).

These results revealed that the teenagers' performance in the cloze test was predicted not only by the type of connective but also by individual variation in the degree of exposure to print, as measured by the ART-F-CL. The interaction effect between the ART-F-CL and Connective showed that the performance on the ART-F-CL explained particularly well the use of certain connectives. For instance, teenagers with a higher degree of exposure to print were 0.35 times more likely to give a correct answer for the connective *étant donné que* 'given that', 0.99 times for the connectives *en outre* 'moreover', and 1.32 times for *au fur et à mesure que* 'as' than teenagers with a lower score on the ART-F-CL.

Table 7. Estimates of the best fitting model for teenagers

	Estimate	**SE**	**z**	**p**
	Teenagers			
(Intercept)	0.36	0.32	1.11	0.267
CONNECTIVE				
aussi 'therefore'	-2.30	0.32	-7.31	<0.001
dans la mesure où 'in so far as'	-1.28	0.31	-4.10	<0.001
en outre 'moreover'	-1.45	0.31	-4.69	<0.001
hormis que 'except'	-1.29	0.31	-4.14	<0.001
or	-0.96	0.31	-3.07	0.002
au fur et à mesure que 'as'	-0.98	0.32	-3.08	0.002
suivant que 'depending on'	-1.10	0.31	-3.52	<0.001
ART-F-CL	0.71	0.13	5.53	<0.001
SESSION				
Session 2	-0.03	0.04	-0.65	0.513
Session 3	-0.10	0.04	-2.48	0.013
CONNECTIVE*ART-F-CL				

	Estimate	**SE**	**z**	**p**
aussi * ART-F-CL	-0.10	0.11	-0.91	0.362
dans la mesure où * ART-F-CL	0.17	0.11	1.53	0.126
en outre * ART-F-CL	-0.46	0.11	-4.11	<0.001
hormis que * ART-F-CL	0.17	0.11	1.50	0.133
or * ART-F-CL	0.11	0.11	0.98	0.329
au fur et à mesure que * ART-F-CL	0.34	0.12	2.88	0.004
suivant que * ART-F-CL	0.21	0.12	1.86	0.063

7.2.3. Adults

The fit of the model for the group of adults kept improving after adding Connective ($\Delta\chi^2$=42.14, Δdf=7, p<.001), a two-way interaction between Connective and ART-F ($\Delta\chi^2$=20.87, Δdf=8, p=.007), and finally Session ($\Delta\chi^2$=18.05, Δdf=2, p<.001). Since adding random slopes did not allow the model to converge and Age did not improve the model's fit ($\Delta\chi^2$=1.54, Δdf=1, p=0.215), the final model included Connective, ART-F, and Session as fixed factors and Item and Participant as random intercepts (see Table 8).

The results for adults were similar to those of teenagers. The adults' performance in the cloze test was also predicted by the variation in the ART-F. However, the interactions between the ART-F and connectives were different. Adults with a higher degree of exposure to print were 3.70 times more likely to produce a correct answer for the connective *aussi* 'therefore' and 2.27 times for likely for the connective *suivant que* 'depending on' than adults with a lower ART-F score. The scores on the ART tests for both groups of participants are reported in Table 9.

Table 8. Estimates of the best fitting model for adults

	Estimate	**SE**	**z**	**p**
	Adults			
(Intercept)	2.59	0.85	3.04	0.002
CONNECTIVE				
aussi ‘therefore’	-3.08	0.87	-3.56	<0.001
dans la mesure où ‘in so far as’	0.24	1.00	0.24	0.812
en outre ‘moreover’	-3.45	0.85	-4.07	<0.001
hormis que ‘except’	-0.91	0.93	-0.98	0.326
or	-0.32	1.01	-0.31	0.754
au fur et à mesure que ‘as’	-0.53	1.06	-0.51	0.613
suivant que ‘depending on’	-1.70	1.07	-1.58	0.114
ART-F	0.27	0.32	0.84	0.400
SESSION				
Session 2	0.37	0.11	3.51	<0.001
Session 3	-0.06	0.10	-0.55	0.584
CONNECTIVE*ART-F				
aussi * ART-F	0.62	0.31	1.99	0.047
dans la mesure où * ART-F	-0.13	0.37	-0.36	0.719
en outre * ART-F	0.36	0.30	1.21	0.226
hormis que * ART-F	0.12	0.34	0.35	0.728
or * ART-F	0.23	0.38	0.61	0.540
au fur et à mesure que * ART-F	0.38	0.41	0.93	0.352
suivant que * ART-F	1.02	0.43	2.34	0.019

Table 9. Descriptive statistics for the ARTs by group

	M (SD)	Observed range	Possible range
Teenagers	10.19 (6.03)	-1–29	-40–40
Adults	10.80 (6.63)	1–33	

Note. A different version of the ART was used for teenagers and adults

7.2.4. Qualitative analysis of the results from the active training

The analysis of the participants' performance during the active training revealed that the task of completing the grammatical rule and the continuation exercise resulted in different scores. For most of the connectives, namely *aussi* 'therefore'*, hormis que* 'except'*, étant donné que* 'given that'*, or, au fur et à mesure que* 'as'*, suivant que* 'depending on', teenagers found it more challenging to build the rule than to complete the continuation, as hinted by lower mean scores on the rule than on continuation exercise (see Table 10). This finding may suggest that an implicit knowledge of the use of these connectives is stronger than an explicit metalinguistic knowledge, as speakers better use these connectives in context than reason about their linguistic functions. In contrast, *en outre* 'moreover' is the only connective that received lower scores on the continuation exercise than on the rule reconstruction exercise. This may stem from the fact that this connective signals a less complex coherence relation of addition that participants understand well and therefore can explain better its functioning from the metalinguistic point of view. However, the fact that speakers still find it difficult to complete a continuation task, although they can successfully reconstruct the rule about its functioning, suggests that their metalinguistic understanding was not transferred to an implicit ability to use this connective. Finally, the connective *dans la mesure où* 'in so far as' seems to be well acquired on both implicit and explicit level, as it received comparable scores for the rule and continuation tasks.

Interestingly, even though adults generally performed better on the active training activity, they showed the same pattern as teenagers in the performance between the two tasks for most of the connectives, except for *au fur et à mesure que* 'as', which received comparably high scores for the rule and continuation exercise. This finding shows that probably also adult speakers have different levels of implicit and explicit knowledge about the use and functioning of certain connectives.

Table 10. Mean correctness score in the active training activity per type of task (reconstruction of the rule versus continuation exercise), connective, and group

	Teenagers		**Adults**	
	Rule	**Continuation exercise**	**Rule**	**Continuation exercise**
dans la mesure où	.81	.83	.91	.95
aussi	.49	.58	.66	.78
hormis que	.20	.45	.20	.50
étant donné que	.72	.85	.93	1.00
en outre	.81	.53	.98	.78
or	.22	.38	.20	.35
au fur et à mesure que	.60	.84	.97	.95
suivant que	.34	.83	.50	.93

7.3. Discussion

In this chapter, I presented a study assessing the use of infrequent connectives signalling more and less common coherence relations in a sentence cloze test, and the role of two types of training activities for improving this competence in French speaking teenagers. Moreover, I also examined whether the use of these connectives varied depending on teenagers' degree of exposure to written language, as measured by the ART.

As for the general effect of training, only three connectives, namely *aussi* 'therefore', *en outre* 'moreover', and *or*, benefitted from it, and received higher scores on the cloze test during Session 2, held immediately after the training, than during Session 1. However, the long-term effect of the training activity remained only for the connective *or*, which received higher scores also during Session 3, held four weeks after the training, than during Session 1. It is possible that the training was efficient on a short term for *aussi* ‘therefore’ and *en outre* ‘moreover’ because the implicit knowledge of these connectives was the lowest, as revealed by the mean scores on the cloze test already in Session 1 (M_{aussi}=0.36 and $M_{en\ outre}$=0.38). Therefore, activating of the explicit knowledge system by explaining the functioning of these connectives helped teenagers to perform better on the second cloze test, but apparently was not enough to solidify the implicit knowledge and perform on the same level also during the third cloze test. In comparison, probably since the initial knowledge of the connective *or* was higher (*M*=0.68), it provided a more solid basis for the training, assuring that its effect stays longer and does not disappear four weeks later.

In addition, this analysis also demonstrated that the type of coherence relation (more or less common) did not predict the difference in the scores between the tested connectives nor the difference related to the training activity. This finding suggests that the relations of condition, consequence, exception, cause, and addition, which have a clear label, a more frequently used equivalent connective, and therefore can be more easily described, are not better mastered by teenagers and are not necessarily easier to train.

The analysis of the effect of two types of training activities revealed that following an active or a passive training did not result in different scores on the cloze test neither immediately after the training (Session 2) nor four weeks later (Session 3). This means that both types of training tasks had the same effect on the ability to insert an appropriate connective; although as mentioned before, the general effect of the training was positive only for few connectives. It is all the more intriguing that the degree of exposure to print, in contrast, predicted an important part of the inter-individual variation in the ability to use connectives. This result suggests that, perhaps, the competence with connectives depends more on a long-term exposure to the written language rather than on training that provides relevant input for the acquisition to take place, activating the explicit knowledge that is not converted into implicit knowledge, at least not on the one proposed in this study. It is apparently not sufficient to explain the connectives' meanings in a passive or active way, like it can be the case for lexical items encoding just a conceptual meaning (Hulstijn & Laufer, 2001). The ability to use them most probably comes from the exposure to an extensive repeated input that allows to internalise the procedural meaning of connectives over time. In this respect, connectives are closer to the acquisition of grammar than the lexicon. However, it is possible that combining exposure to written language with the repeated form-focused instruction activities may facilitate the process of acquisition of connective functions in a longer perspective. Further work needs to be carried out to establish which alternative instruction techniques are better suited to enhance the intake from the input to which speakers are exposed.

Finally, the existing difference between an explicit and an implicit knowledge of connectives was also hinted at in the qualitative analysis of the performance in the active training. This analysis revealed that the participants performed better on the continuation insertion task than on the rule reconstruction activity for the majority of connectives. This finding suggests that both teenagers and adult speakers have different levels of implicit and explicit knowledge about the use and functioning of certain connectives, as the implicit knowledge of the connective use in context was stronger than an explicit metalinguistic knowledge about their functions. This result indicates that learning how to use connectives in the written language involves explicit and implicit knowledge systems similar to those at play during the learning a second language (see, e.g., Ellis, 1994). To put it differently, learning rules of written language is in some way comparable to learning the second language, at least when it comes to learning the functioning of discourse connectives in writing. The involvement of explicit and implicit knowledge systems in the development of the competence with discourse connectives should be further examined in future research, especially by using on methods, allowing to have a more direct access to the implicit knowledge system.

7.4. Conclusion

The study presented in this chapter provided several important findings regarding the development of the competence with infrequent connectives during teenage years. First, the connectives *aussi* ‘therefore’ and *en outre* ‘moreover’ received the lowest scores on the cloze test, even among the connectives with comparable frequencies. This result suggests that the frequency of these connectives was not the relevant factor explaining the difficulties observed in previous chapters, making them particularly challenging to use, in contrast to what was suggested in previous studies (Chapter 3; Zufferey & Gygax, 2020a, 2020b).

Second, the unique character of the coherence relation type does not predict either the accuracy of responses in the cloze tests. This outcome is in line with previous findings on the diminishing role of the cognitive complexity of coherence relations during teenage years (Chapters 3, 4 & 5). That is, coherence relation type is no longer relevant for the use of connectives in a cloze test by teenagers, whether we look at it from the point of view of the cognitive complexity in the paradigm of Sanders et al. (1992) or from the perspective of prototypicality like in the present study. At least, not when the accuracy of their use is measured via off-line methods. Third, we observed that the training activities had only a limited impact on the performance with connectives. This finding indicates that it is a long-term exposure to the written input, as measured by the ARTs, that allows to acquire and internalise the functioning of connectives in discourse rather than a one-time activation of the explicit knowledge about this functioning.

8. General Discussion and Conclusion

This final chapter summarises and discusses the main findings of the present thesis, which aimed to provide more evidence on the developing competence of teenagers with discourse connectives. After summarizing and discussing the main findings, this chapter highlights some limitations of the present work, and proposes several directions for future research.

8.1. Between childhood and adult years

General patterns of performance with connectives were overall similar between the group of teenagers and adults. Yet, adults outscored teenagers across all the experiments conducted within this thesis. Adults were better at using connectives with different frequencies in text- and sentence-cloze tests (Chapters 3, 4, 5 & 7), and they were also more sensitive to alternative signals of the list relation (Chapter 6). This finding tends to confirm the first hypothesis (H1), formulated in Section 2.4.1. More specifically, it suggests that the ability to match connectives with an appropriate coherence relation as well as the sensitivity to different signals of coherence when inferring a coherence relation are not fully acquired at the end of childhood and develop throughout teenage years and early adulthood. The fact that teenagers show a pattern of results similar to that of adults across different tasks, but do not reach the same scores as adults, underlines the particular developmental status of teenage years (see, e.g., Berman, 2004b; Nippold, 2008). It is indeed a transitional period when a basic knowledge about connectives has probably already been acquired (such as differences between coherence relation types), but a more fine-grained mastery is still being developed. The latter may be due to the fact that a more advanced mastery of connectives comes from a long-term exposure to written contexts where connectives are used. Exposure to such contexts might accumulate with years and progressively form a more advanced proficiency with connectives. The role of exposure to print for building a more advanced mastery of connectives is discussed more in detail in Section 8.3.

8.1.1. Factors accounting for the difficulty to use certain connectives

The studies presented in Chapters 3, 4, 5 and 7 corroborated H2 about the role of complexity of coherence relations (see Section 2.4.2). More precisely, these studies indicated that the coherence relation type, whether it is distinguished based on the CCR model (Sanders et al., 1992) or, like in Chapter 7, based on the singularity of the function that relations fulfil (more or less common), does not predict the use of connectives in cloze tests. To put it differently, when the teenagers' performance with connectives is measured with an off-line task, the level of complexity of a relation does not seem to be a challenge for young speakers. This result does not mean that different relations are necessarily processed in the same way. Rather, an on-line measure could possibly reveal differences between different coherence relation types, just like it was the case in the studies on adult speakers (see, e.g., Canestrelli et al., 2013; Drenhaus et al., 2014; Xu et al., 2015). However, the fact that the scores on a series of production tasks in French and on the one in Russian are not systematically affected by the same type of coherence relations suggests that teenagers have acquired at least a basic knowledge of various types of relations and can mark them with an appropriate connective.

Chapters 3, 5 and 7 also revealed that the existence of other functions does not necessarily render polyfunctional connectives more challenging to use in a cloze task, at least when alternative meanings are controlled for in the experimental design. This result was observed both for native speakers of French (Chapters 3 and 7) and Russian (Chapter 5). Moreover, this lack of difficulty was evidenced both when the dominant function (e.g., *en effet* 'for', *hotia* 'even if', *no* 'but', *odnako* 'however', and *da i* 'moreover') and non-dominant functions of connectives were assessed (e.g., for *dans la mesure où* 'in so far as' and *or*, a connective for which no direct translation exists in English). Hence, H5 (see Section 2.4.2) was corroborated by the results from these studies, revealing that polyfunctionality does not predict the teenagers' performance with connectives in off-line production tasks.

Chapters 3, 4 and 5 provided evidence that connective frequency (see H3 in Section 2.4.2) and the mode (see H4 in Section 2.4.2) in which they are typically used (oral or written) may explain the lower performance with certain connectives, especially by younger speakers. However, Chapter 7, which examined connectives with particularly low frequencies, showed that frequency alone was not the factor hindering the use of connectives. As a matter of fact, the consequence connective *aussi* 'therefore' and the additive connective *en outre* 'moreover' received the lowest scores on the cloze test even compared to connectives with lower frequencies. Thus, H3 on the role of frequency for the mastery of connectives (Section 2.4.2) was only partially supported.

Taken together, the findings from the different chapters indicate that none of the mentioned factors alone can account for the difficulty to use connectives experienced by teenagers. It is likely a combination of factors and particularities of the contexts of use that make some connectives more challenging than others. The connectives *aussi* ‘therefore’ and *en outre* ‘moreover’ particularly stand out from the point of view of their observed difficulty, as they received significantly lower scores across several experiments. In the case of *aussi*, the challenge to use this connective in cloze tests may have stemmed from a low distribution of the consequence function, which is present in approximately 10% of uses in corpora (27 occurrences per million words), in comparison to the dominant additive function (90% of uses and 1,253.39 occurrences per million words). Moreover, the two functions do not only differ in terms of frequency distributions, but also tend to be used in different modes. The additive function appears both in oral[1] (196.5 occurrences per million words) and written[2] (1191.43 occurrences per million words) modes, whereas the consequence function is present in written language (40.65 occurrences per million words) and is not really used in oral speech (no occurrences found in corpus data). Therefore, the existence of a much more frequent additive function may dominate and hamper the development of the knowledge of the considerably less frequent consequence function. This may be especially the case if

[1] The frequency in oral speech was calculated based on the oral sub-corpus of French Orféo (Benzitoun et al., 2016).
[2] The mean frequency in writing was calculated based on three corpora, namely journalistic (Le Monde corpus), argumentative (the French part of the Europarl corpus, Koehn, 2005), and a corpus of literary texts (the Frantext corpus, ATILF, 1998-2022).

speakers have scarce exposure to the written language, in which the consequence meaning tends to appear. The importance of exposure to text for developing a mastery of connectives is discussed later in this chapter.

As for the connective *en outre*, several potential reasons may account for its difficulty. First of all, this connective tends to be used in writing and rarely appears in oral speech, which is the primary source of linguistic input for the majority of native speakers (1.75 and 91.29 occurrences per million words in oral and written corpora, respectively). Not only is this connective predominantly used in writing, but also the written contexts in which it appears, are quite specific, as *en outre* is mostly used in formal and administrative written contexts. Furthermore, the very relation encoded by this connective may have added a layer of difficulty in addition to the fact that (young) speakers are rarely exposed to the contexts in which it is used. The additive relation tends to be conveyed in discourse by a wide variety of alternative signals, such as reference, semantic, syntactic, graphical, genre, and numerical features; and this relation is only marginally signalled by connectives (Das & Taboada, 2018). Therefore, when participants had to insert a connective between sentences targeting the additive relation, they probably were inclined to search for a different interpretation of the coherence relation, where marking by a connective would be more informative and expected.

Finally, Chapter 3 also highlighted how experimental context may affect the use of connectives. It was shown that teenagers and adults had more difficulties inserting an appropriate connective in a context of longer texts than between pairs of isolated sentences, contrary to the initial hypothesis (H13, Section 2.4.6). It is possible that the process of identifying coherence relations in a text involves a greater cognitive load in comparison to a reduced context of pairs of sentences. In texts, participants had more elements to consider in order to establish the coherence relation and to match it with a connective. The findings from Chapter 3 hint that an enhanced exposure to long written texts with various signals of coherence should be promoted in school curricula to facilitate the development of a more extensive knowledge of connectives.

8.1.2. Sensitivity to alternative signals of coherence relations

Since additive relations tend to be marked by alternative signals of coherence, the study in Chapter 6 examined whether teenagers were responsive to such signals. To be more precise, this study assessed whether the presence of expressions of quantity such as *plusieurs* ‘several’ and *différent* ‘various’ in a sentence generated the inference of a list relation, i.e., a subtype of a more generic additive relation in the following sentence that participants had to produce themselves. Moreover, this study also examined whether the generated inference was modulated by the presence of additive and consequence connectives. The findings corroborated the initial hypothesis (H6) and revealed that teenagers were sensitive to the alternative signals of list relations, meaning that even younger and less experienced speakers are responsive to less salient and less prototypical signals of coherence relations. However, teenagers' sensitivity to alternative signals was inferior to that of adults, which suggests that receptiveness to alternative signals for the inference of a coherence relation develops as linguistic experience increases.

The presence of consequence connectives in the task almost completely overrode the inference of the list relation. This implies that connectives signalling the relation of consequence create a much stronger mental representation of this relation than the alternative list signals for the relation of list. However, after the task sentences containing the additive connectives *en plus* and *en outre*, teenagers did not produce significantly more list continuations when a list signal was present than when it was absent, contrary to what was observed for adults. This result indicates that teenagers are probably less receptive to the alternative list signals when a more salient signal like a connective is also present in the sentence.

Yet, the fact that teenagers still produced a comparable number of list continuations after the sentences that combined a more frequent additive connective *en plus* and list signals strongly suggests that the inference triggered by a less salient, more polyfunctional, and less prototypical alternative signal does not necessarily diminish because of the inference triggered by a more specialised and prototypical signal, such as a connective. This result shows that the inference generation of a particular relation is a complex process, involving various types of signals spread in discourse that coexist and influence each other, and probably even more so in real-life discourse.

8.1.3. The role of age and individual variation in linguistic competence

In addition to the factors related to the linguistic characteristics of connectives, this thesis also examined factors of inter-individual variation between young speakers. The findings from all the studies included in the present thesis confirm the initial hypothesis (H11) and all indicate that age plays a minor role for the developing competence with discourse connectives during teenage years, contrary to what was found for younger children. In Chapters 6 and 7, for instance, age predicted neither a better sensitivity to alternative list signals nor a more accurate use of infrequent connectives in cloze tests. Chapters 3, 4 and 5 revealed that teenagers' age was not the most relevant predictor of the performance on cloze tests. The latter was subject to a much more important inter-individual variation related to participants' academic background, vocabulary level, and exposure to print, which tends to confirm H8, H7, and H10 (see Section 2.4.4).

Taken together, these findings are very insightful about the singular nature of teenage years for the development of the competence with connectives and language development in a more general sense. The fact that chronological age is secondary compared to measures of linguistic competences supports the idea that this period is different from the acquisition of language during childhood at least on two dimensions. First, the development takes place at a slower pace and is less salient (see, e.g., Berman, 2004; Nippold, 1993, 2006, 2008). In other words, the development curve during teenage years is less steep than during childhood, because it involves a more complex and subtle development of all the linguistic domains happening at the same time and on two levels (oral and written). The ability to accurately use a wide variety of connectives with different frequencies and from different modes, therefore, depends on the proficiency of the whole linguistic system, as it lies between several linguistic domains. Second, since the development of a more advanced proficiency does not end with cognitive maturation, it means that it can continue also after the teenage period, throughout the whole lifespan. Finally, the finding on the secondary role of age for the increasing competence to use connectives is also informative of the linguistic nature of discourse connectives, which I discuss in the next section.

8.2. Between declarative and procedural knowledge

The studies carried out in this thesis provided evidence on the linguistic status of connectives. The fact that age becomes less relevant for the mastery of connectives during teenage years and does not predict their mastery in adulthood (Chapters 3, 4 and 5), especially in comparison to the measures of linguistic competence, might support the view that connectives do not only encode concepts, but also procedures. Age plays a less prominent role for the mastery of connectives in teenage years and adulthood because connectives may be part of procedural knowledge, which with years requires more time and effort to develop, as procedural learning ability diminishes with age (Ullman, 2001). However, the potential link between the diminishing role of age and procedural nature of connectives is still speculative and needs a more solid evidential basis.

The findings from Chapters 4, 5 and 7 however provided other evidence pointing towards a double-edged nature of connectives between concepts and procedures. The results revealed that both the vocabulary level and the degree of exposure to print predicted accuracy on the connective insertion task among all groups of participants. These results suggest that connectives are part of general lexical knowledge, but the competence to use them in discourse is implicit and is developed through exposure to written language activities. Indeed, the study from Chapter 7, in which I examined the possibility to enhance the mastery of connectives via a training session, demonstrated that connectives are learnt like processing instructions. This study did not support the initial hypothesis (H12, see Section 2.4.5) and demonstrated that metalinguistic explanation of the functioning of connectives, whether it includes an active engagement or a passive involvement of students, does not improve the use of connectives in cloze tests. It is rather a long-term exposure to written language, as measured by the Author Recognition Tests, that allows speakers to internalise their use in discourse. In other words, it seems that connectives are learnt implicitly from a repetitive exposure to its contexts of use, similar to other proceduralised skills and habits (Ullman, 2001).

To summarise, the reported results support the current research arguing for the double nature of connectives, in that they occupy an intermediate position between conceptual and procedural knowledge (Anscombre & Ducrot, 1997; Ducrot, 1972, 1980;

Fraser, 2006; Hussein, 2008; Mauri & van der Auwera, 2012; Moeschler 2002, 2005; Wilson, 2011). In this sense, this finding is in line with the Wilson's (2011) view of conceptual and procedural meaning. On the one hand, connectives are part of the lexicon, as they are probably stored with other lexical items and their meaning can be made accessible to consciousness through an explicit metalinguistic explanation, as it was done in the training activities reported in Chapter 7. On the other hand, to be able to correctly use connectives in discourse, it is not sufficient just to be aware of the mapping between their form and meaning. It is important to develop an automatised procedure through a repetitive long-term exposure to the written language for using connectives in discourse.

Given the dual nature of connectives, the development of their mastery may potentially be supported by both declarative and procedural memory systems, which are probably interconnected. It is possible that speakers may internalise connectives as single lexical items in the declarative memory, where they are stored together with other lexicon items, and may also develop the procedure for using the connectives to signal coherence relations between segments of discourse. However, whether and how exactly the two systems interact with each other remains an open question.

8.3. Between comprehension and production

The primary goal of the current thesis was to assess how teenagers use connectives in production tasks, whereas comprehension was examined only to a limited extent. Nevertheless, some interesting findings about the two types of competences were observed. The results of the off-line comprehension task compared to those of the off-line production tasks, both reported in Chapter 3, revealed the disparity between the level of comprehension of certain connectives and coherence relations that they encode and the ability to use them in discourse. Both teenagers and adults were successful at selecting the correct statements about the parts of texts, linked by the connectives *aussi* ‘therefore’, *en outre* ‘moreover’, *en effet* ‘for’, *toutefois* ‘however’. However, when they had to insert an appropriate connective in cloze tests, their overall performance was lower, especially in the case of the connectives *aussi* ‘therefore’ and *en outre* ‘moreover’. That is, it seems that speakers could understand the meaning of the sentences related by all four connectives, but had a greater difficulty actively using these connectives.

This finding concurs with previous research on childhood years, suggesting that the mechanisms of production and comprehension in language acquisition in general (see, e.g., Clark & Hecht, 1983) and in the acquisition of connectives in particular (see, e.g., Cain & Nash, 2011) do not develop at the same pace. What is more, our results hint that the disparity between these mechanisms remains over the whole lifespan, as both teenagers and adult speakers had the same pattern of results on the comprehension and production tasks. This finding is in line with the asymmetry between word comprehension and production in 2-year-old children and adults aged 18–22 found by Gershkoff-Stowe and Hahn (2013). The authors exposed participants to novel objects and their names during a training session and afterwards assessed their level of comprehension and the ability to use these new words. The comprehension was assessed with a task, in which participants had to point to an object after they heard its name, while the production task consisted in naming the object that they were presented. The results revealed that the ability to use novel words to name objects lagged behind the comprehension of these words in both groups of participants. That is, children and adults required more time to form an ability to use novel words than to understand them. Our findings complement those of Gershkoff-Stowe and Hahn (2013), by showing that asymmetry between comprehension and production skills in adult and young speakers probably exists not only when learning novel words, but also applies to known words, such as connectives tested in Chapter 3. This finding is intriguing and calls for further examination.

However, this outcome should be considered with caution, as it may have at least two alternative explanations not related to the production and comprehension of connectives. First, the scores on the comprehension task were not different between connectives with low (*aussi* ‘therefore’ and *en outre* ‘moreover’) and high (*en effet* ‘for’ and *toutefois* ‘however’) frequencies, contrary to the results of cloze tests, because *aussi* and *en outre* mark continuous coherence relations of consequence and addition. That is, these relations do not necessarily require explicit marking by connectives and can be conveyed implicitly (see, e.g., Asr & Demberg, 2012). As a result, participants in principle could have inferred appropriate coherence relations and have understood the meaning of the texts without understanding these connectives.

Another reason for the high judgement scores on the comprehension task across all the connectives may be rooted in the design of this task. Given that the statements about the content of the texts included equivalent oral connectives, this may have made the coherence relations easier to understand. Furthermore, the fact that a correct answer was expected for all the experimental items and an incorrect one for all filler items, may also have biased the results of the task. Thus, future research could benefit from addressing these design issues, as well as from complementing such designs with on-line measures of connective comprehension. I provide a more complete discussion of the limitations of the present work in the next section.

8.4. Limitations and future directions

Like all experimental research, the studies conducted in this thesis have a number of limitations. Some of them have already been raised in previous chapters, but are outlined here to provide a more exhaustive overview. This section also discusses avenues for future research based on these limitations.

8.4.1. Combining on-line and off-line measures

All the findings of the current thesis were obtained via off-line experimental measures, such as sentence- and text-cloze tests (Chapters 3, 4, 5 and 7), story continuation task (Chapter 6), and an off-line comprehension task (Chapter 3). These measures allowed us to obtain interesting findings on the level of off-line production and (to a lesser degree) comprehension of discourse connectives by teenagers. However, since the participants could take their time to think about their answers, these findings do not provide information about how these connectives were processed. Assessing on-line processing of discourse connectives during production and comprehension tasks would be an important contribution for research on the teenage period, revealing a more complete picture about the level of connective use and comprehension. On-line measures may be particularly sensitive to the differences in processing coherence relations varying in cognitive complexity by teenagers. As revealed in this thesis, although teenagers reached lower accuracy scores than adults across all the tasks, they had a similar pattern of performance with connectives than adults. Given that adults showed differences in the on-line processing of connectives encoding relations with varying degrees of cognitive complexity (see, e.g., Canestrelli et al., 2013; Koehne & Demberg, 2013; Koornneef & Sanders, 2013), it is possible that teenagers also may demonstrate patterns comparable to those of adults in their performance with on-line measures. This would mean that, despite potential challenges related to processing different coherence relations, younger speakers manage to compensate for these difficulties when using connectives in written language.

Including on-line measures for studying the use and comprehension of connectives may also be informative to assess how the combination of different types of signals of coherence are processed by young readers. It would be particularly revealing to compare the processing of connectives and alterative signals between teenagers and adults using eye-tracking tasks involving reading. Such tasks may provide a more precise vision of the hierarchy in the processing of different types of signals of coherence and how this hierarchy is modulated over years.

Another line of research that may benefit from on-line experimental measures is related to the role of training for improving the mastery of discourse connectives. The study conducted in Chapter 7 did not trigger a facilitating effect of the explicit metalinguistic explanation of the connective functions on their use in cloze tests. Introducing eye-tracking or self-paced reading tasks might reveal whether readers change the way they process written texts and pay more attention to the segments including connectives. If this were the case, it would indicate that training may indeed help focus readers' attention to the elements of texts that are relevant for constructing a coherent mental representation of discourse. As a result, readers might continue using this strategy during other reading activities and, hence, will benefit faster from exposure to print activities.

8.4.2. Future work on individual differences in the mastery of connectives

One of the major contributions of this thesis is the development of several measures of individual variation, such as the teenage versions of the ARTs in French (ART-F-CL in Chapter 4) and in Russian (ART-RU-CL in Chapter 5), the teenage versions of the vocabulary level test, also in both languages (Chapters 4 and 5), and the adult version of the ART for native Russian speakers (ART-RU in Chapter 5). Future work on linguistic development during the teenage period will therefore benefit from these measures of exposure to print and vocabulary width, and will further assess the external validity of these measures.

However, some measures used in this thesis did not account for the variations in the mastery of connectives. In Chapter 3, the test on written grammatical competence (introduced by Zufferey & Gygax, 2020a) did not predict performance with connectives on the cloze tasks, and thus, did not support H9 (see Section 2.4.4). This finding probably indicates that grammatical awareness is not related to the ability to use connectives in discourse. Moreover, since the scores received by teenagers were quite low, it also suggests that an adapted version of the grammatical test should be developed for younger speakers in future research.

The fact that the Title Recognition Test adapted for French-speaking teenagers (TRT-F) did not systematically predict variations in the performance with connectives by teenagers is quite surprising, as previous work suggested that this type of test is particularly suitable for younger populations (Echols et al., 1996). This outcome can be interpreted in several ways. First, it is possible that teenagers who participated in the study were infrequent readers of out-of-school fiction and thus were not at all the target of the developed TRT-F. Alternatively, the selected names of books were not narrowed down sufficiently enough to reflect their level of exposure to out-of-school print. Since different schools from different regions in Switzerland participated in the studies, different communities of speakers with distinct reading circles might have been involved. A more time-consuming, but probably more efficient method to adopt in future work could be to develop a separate version of the TRT-F for each school. More targeted versions of the TRT-F could also be compared with the ART-F-CL, based on classical authors, in order to better grasp which measure is more revealing about teenagers' exposure to written language activities. In the absence of an adapted version of the ART, as it was the case in Chapter 3, a measure of subjective exposure to print can sometimes predict the variation in teenagers' mastery of connectives. However, this measure is less reliable and robust than the ARTs, adapted to the tested population, because it can be subject to exaggeration and guessing (see, e.g., Chateau & Jared, 2000; Echols et al., 1996; Wimmer & Ferguson, 2022). In fact, the subjective exposure to print did not account for the observed variations in the performance with connectives in other studies, reported in Chapters 4 and 5.

Chapter 6 also demonstrated that teenagers' sensitivity to alternative list signals was not predicted by the degree of exposure to print, as measured by the ART-F-CL and ART-F in teenagers and adults, respectively. The fact that these measures revealed variations in the performance with connectives on the cloze tests in Chapters 4 and 7 as well as in the study of Zufferey and Gygax (2020a), but not on the story-continuation task, may be revealing about the nature of the two tasks. The ability to match a connective with a corresponding coherence relation, and that to infer a coherence relation from a lexical cue are two distinct competences, and it is probably more challenging to measure the variation in the inference generation for several reasons. First, the task, measuring the sensitivity to alternative list signals, provided evidence only about one type of coherence relation that a participant chose to write down, while all other relations that might have been inferred as well remain unknown. This suggests that participants might have been sensitive to alternative list signals, but the task did not always reveal this sensitivity.

Second, inferring a non-targeted non-list relation from a sentence with a polyfunctional cue, such as an adjective of quantity, which is not specialised in discourse marking, was not false and was still a valid linguistic possibility. That is, even if a reader inferred another relation but produced a coherent continuation, it does not mean that they have low proficiency with written language, as reflected by the ARTs. Exposure to print tests might simply not be most suited to measure inter-individual variation on the tasks which do not have a true or false component. As a result, the ARTs, which were used in Chapter 6, may have not been sensitive enough to the variation in the list inference generation.

8.5. Conclusion

This thesis examined the development of the mastery of discourse connectives during teenage years from several perspectives. It showed that the teenage period is a crucial transitional phase between the basic knowledge acquired over childhood and a more proficient mastery, which continues to develop during adulthood. The studies reported in this thesis also revealed that none of the endogenous factors considered alone affect the competence with connectives. It is rather a combination of those factors that may render specific connectives more challenging. Lastly, this thesis also contributed to the exploration of individual variation in the use of connectives and highlighted the importance of long-term exposure to print for the development of a greater proficiency with connectives. Relying on complementary methodologies, the continued study of how the comprehension and use of discourse connectives develops from childhood into adulthood is essential for building a better understanding of how coherent discourse is formed. Learning more about how this development can be enhanced will be an important endeavour for future research, in order to help attenuating the inequalities in language mastery that may otherwise persist.

References

Acheson, D., Wells, J., & MacDonald, M. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods, 40*(1), 278–289.

Akbasli, S., Sahin, M., & Yaykiran, Z. (2016). The effect of reading comprehension on the performance in science and mathematics. *Journal of Education and Practice, 7*(16), 108–121.

Allen, L., Cipielewski, J., & Stanovich, K. E. (1992). Multiple indicators of childrens reading habits and attitudes: Construct validity and cognitive correlates. *Journal of Educational Psychology, 84*(4), 489–503.

Andreev, L., & Uccelli, P. (2023). The secret life of connectives: a taxonomy to study individual differences in mid-adolescents' use of connectives in writing to persuade. *Reading and Writing.* https://doi.org/10.1007/s11145-023-10425-3

Anscombre, J.-C., & Ducrot O. (1997). *L'argumentation dans la langue* (3d ed.). Pierre Mardaga.

Ariel, M. (1998). Discourse markers and form-function correlations. In A. Jucker & Y. Ziv (Eds.), *Discourse Markers. Descriptions and Theory* (pp. 223–259). Amsterdam: John Benjamins.

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

Asr, F. T., & Demberg, V. (2012). Implicitness of discourse relations. *Proceedings of the International Conference on Computational Linguistics (COLING)* (pp. 2669–2684). Mumbai, India.

Asr, F. T., & Demberg, V. (2020). Interpretation of discourse connectives is probabilistic: evidence from the study of *But* and *Although*. *Discourse Processes, 57*(4), 376–399.

ATILF. (1998-2022). *Base textuelle Frantext* (En ligne) [Data set]. ATILF-CNRS & Université de Lorraine. https://www.frantext.fr/

Au, T. K. (1986). A verb is worth a thousand words: The causes and consequences of interpersonal events implicit in language. *Journal of Memory and Language, 25*, 104–122.

Ballard, D., Conrad, R., & Longacre, R. (1971). The deep and surface grammar of interclausal relations. *Foundations of Language, 4*, 70–118.

Barr, D., Levy, R., Scheepers, C., & Tily, H. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language, 68* (3), 255–278.

Barr, C.D., Uccelli, P., & Phillips Galloway, E. (2019). Specifying the Academic Language Skills That Support Text Understanding in the Middle Grades: The Design and Validation of the Core Academic Language Skills Construct and Instrument. *Language Learning, 69*, 978–1021.

Bates, T. (2021). *umx: A helper package for Structural Equation Modeling in OpenMx.* University of Edinburgh, Edinburgh, UK.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67* (1), 1–48.

Baayen, R. H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R.* Cambridge University Press, Cambridge, UK.

Beek ter, M. (2020). Supporting reading comprehension in history education: The use and usefulness of a digital learning environment. University of Groningen.

Benzitoun C., Debaisieux, J.-M., & Deulofeu, J. (2016). Le projet ORFÉO: un corpus détude pour le français contemporain. *Corpus, 15*.

Berman, R. A. (Ed.). (2004a). Language development across childhood and adolescence. Trends in language acquisition research: Vol. 3. Amsterdam: John Benjamins.

Berman, R. A. (2004b). Between emergence and mastery: The long developmental route of language acquisition. In Author (Ed.), *Language development across childhood and adolescence* (pp. 9–34). Amsterdam: John Benjamins.

Berman, R. A., and Slobin, D. I. (1994). *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Hillsdale NJ: Lawrence Erlbaum.

Berman, R. A., & Verhoeven, L. (2002). Developing text-production abilities across languages, genre, and modality. *Written Languages and Literacy, 5*(1), 1–44.

Bestgen, Y., & Vonk, W. (1995). The role of temporal segmentation markers in discourse processing, *Discourse Processes, 19*, 385–406.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Philadelphia, PA: John Benjamins.

Black, J. B., & Bern, H. (1981). Causal coherence and memory for events in narratives. *Journal of Verbal Learning and Verbal Behavior, 20*, 267–275.

Blakemore, D. (1987). *Semantic constraints on relevance*. Blackwell, Oxford.

Blass, R. (1990). *Relevance Relations in Discourse: A Study with Special Reference to Sissala* (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press.

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.

Blything, L.P., Davies, R., & Cain, K. (2015). Young children's comprehension of temporal relations in complex sentences: The influence of memory on performance. *Child Dev, 86*, 1922–1934.

Blything, L. P., & Cain, K. (2016). Children's processing and comprehension of complex sentences containing temporal connectives: The influence of memory on the time course of accurate responses.*Developmental Psychology, 52*(10), 1517–1529.

Britton, B. K. (1994). Understanding expository text. Building mental structures to induce insights. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 641–674). San Diego, CA: Academic Press.

Britton, B. K., Glynn, S. M., Meyer, B. J. F., & Penland, M. J. (1982). Effects of text structure on the use of cognitive capacity during reading. *Journal of Educational Psychology, 74*, 51–61.

Brown, R., & Fish, D. (1983). The psychological causality implicit in language. *Cognition, 14*(3), 237–273.

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. R., Mak, W. M., & Sanders, T. J. M. (2013). Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements. *Language and Cognitive Processes, 28*, 1394–1413.

Caroll, J. B., Davies, P., & Richman, B. (1971). *Word frequency book*. New York: Houghton Mifflin.

Carlson, L., Marcu, D., & Okurowski, M. E. (2002). RST discourse treebank, LDC2002T07.

Carlson, L., Marcu, D., & Okurowski, M. E. (2003). Building a discourse-tagged corpus in the framework of Rhetorical Structure Theory. In J. van Kuppevelt & R. Smith (Eds.), *Current and new directions in discourse and dialogue* (pp. 85–112). Springer.

Charolles, M. (1983). Coherence as a principle in the interpretation of discourse. *Text - Interdisciplinary Journal for the Study of Discourse*, *3*(1), 71–98.

Carston. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.

Castellà, J. M. (2004). *Oralitat i Escriptura. Dues Cares de la Complexitat del Llenguatge*. Publicacions de l'Abadia de Montserrat, Barcelona.

Chafe, W. (1982). Integration and involvement in speaking, writing and oral literature. In D. Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 35–53). Norwood, New Jersey: ABLEX Publishing Corporation.

Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. *Memory & Cognition, 28*, 143–153.

Clark, E. V., & Hecht, B. F. (1983). Comprehension, production, and language acquisition. *Annual review of psychology*, *34*(1), 325–349.

Clark, H. H., & Murphy, G. L. (1983). Audience design in meaning and reference. In J. F. LeNy & W. Kintsch (Eds.), *Language and comprehension* (pp. 287–299). Amsterdam: North-Holland.

Cook-Gumperz, J., & Gumperz, J. (1992). Changing views of language in education: The implications for literacy research. In R. Beach, J. Green, M. Kamil, & T. Shanahan (Eds.), *Multidisciplinary perspectives on literacy research*. Urbana, IL: National Council of Teachers of English.

Crible, L. (2020). Weak and strong discourse markers in speech, chat, and writing: Do signals compensate for ambiguity in explicit relations? *Discourse Processes, 57*(9), 793–807.

Crible, L. (2021). Negation cancels discourse-level processing differences: Evidence from reading times in concession and result relations. *J Psycholinguist Res,* 50, 1283–1308.

Crible, L., & Cuenca, M. (2017). Discourse markers in speech: Characteristics and challenges for annotation. *Dialogue and Discourse, 8*(2), 149–166.

Crible, L., & Demberg, V. (2021). The role of non-connective discourse cues and their interaction with connectives. *Pragmatics and Cognition*, *27*(2), 313–338.

Crible, L., & Pickering, M. J. (2020). Compensating for processing difficulty in discourse: Effect of parallelism in contrastive relations. *Discourse Processes, 57*, 862–879.

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.

Crosson, A., & Lesaux, N. (2013). Does knowledge of connectives play a unique role in the reading comprehension of English learners and English-only students? *Journal of Research in Reading, 36*, 241–260.

Crowhurst, M. (1980). Syntactic complexity in narration and argument at three grade levels. *Canadian Journal of Education, 5*, 6–12.

Cunningham, A., & Stanovich, K. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology, 82*(4), 733–740.

Das, D. (2014). Signalling of Coherence Relations in Discourse (PhD dissertation). Simon Fraser University, Burnaby, Canada.

Das, D., Scheffler, T., Bourgonje, P., & Stede, M. (2018). *Constructing a Lexicon of English Discourse Connectives. Proceedings of the SIGDIAL 2018*, Melbourne, Australia.

Das, D., & Taboada, M. (2018). Signalling of coherence relations in discourse, beyond discourse markers. *Discourse Processes, 55*(8), 743–770.

Deaton, J.A., & Gernsbacher, M.A. (2000). *Causal conjunctions: Cue mapping in sentence comprehension*. Manuscript, University of Madison-Wisconsin.

De Beaugrande, R., & Dressler, W. (1996). *Introduction to Text Linguistics*. New York, 84–112.

Degand, L., Lefèvre, N., & Bestgen, Y. (1999). The impact of connectives and anaphoric expressions on expository discourse comprehension. *Document Design*, 1, 39–51

Degand, L., & Sanders, T. (2002). The impact of relational markers on expository text omprehension in L1 and L2. *Reading and Writing, 15*(7–8), 739–758.

Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge: C.U.P.

Drenhaus, H., Demberg, V., Koehne, J., & Delogu, F. (2014). Incremental and predictive discourse markers: ERP studies on German and English. *Proceedings of the 36th Annual Conference of the Cognitive Science Society (CogSci)* (pp. 403–408). Quebec, Canada.

Ducrot, O. (1972). *Dire et ne pas dire*. Hermann, Paris.

Ducrot, O. (1980). Analyses pragmatiques. *Communications, 32*, 11–60.

Duggleby, S. J., Tang, W., & Kuo-Newhouse, A. (2016). Does the use of connective words in written assessments predict high school students' reading and writing achievement?. *Reading Psychology, 37*(4), 511–532.

Echols, L. D., West, R. F., Stanovich, K. E., & Zehr, K. S. (1996). Using children's literacy activities to predict growth in verbal cognitive skills: A longitudinal investigation. *Journal of Educational Psychology, 88*(2), 296–304.

Ellis, N. (Ed.). (1994). *Implicit and explicit learning of languages*. London: Academic Press.

Ellis, R., Loewen, S., Elder, C., Reinders, H., Erlam, R., & Philp, J. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching*. Bristol, Blue Ridge Summit: Multilingual Matters.

Evers-Vermeul, J., Hoek, J., & Scholman, M. (2017). On temporality in discourse annotation: Theoretical and practical considerations. *Dialogue & Discourse, 8*(2), 1–20.

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.

Evers-Vermeul, J., & Sanders, T. (2011). Discovering domains – on the acquisition of causal connectives. *Journal of Pragmatics, 43*(6), 1645–1662.

Flores d'Arcais, G. B. (1978). Levels of semantic knowledge in children's use of connectives. In A. Sinclair et al. (Eds.), *The child's conception of language* (pp. 133–153). New York: Springer.

Fodor, J. (1975). *The language of thought*. New York: Crowell.

Fraser, B. (1990). An approach to discourse markers. *Journal of Pragmatics, 14*, 383–395.

Fraser, B. (2006). Towards a theory of discourse markers. In K. Fischer (Ed.), *Approaches to Discourse Particles* (pp. 189 - 204): Elsevier Press.

Fuentes, P. (1998). Reading comprehension in mathematics. *The Clearing House*, *72*(2), 81–88.

Gaddy, M. L., van den Broek, P., & Sung, Y.-C. (2001). The influence of text cues on the allocation of attention during reading. In T. Sanders, J. Schilperoord & W. Spooren (Eds.), *Text representation: Linguistic and psycholinguistic aspects* (pp. 89–110). Amsterdam/Philadelphia: Benjamins.

Gernsbacher, M. A. (1997). Coherence cues mapping during comprehension. In J. Costermans & M. Fayol (Eds.), *Processing interclausal relationships. Studies in the production and comprehension of text* (pp. 3–21). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Gershkoff-Stowe, L., & Hahn, E. R. (2013). Word comprehension and production asymmetries in children and adults. *Journal of experimental child psychology*, *114*(4), 489–509.

Giora, R. (1985). Notes towards a Theory of Text Coherence. *Poetics Today*, *6*(4), 699–715.

Garnham, A., Oakhill, J. V., & Johnson-Laird, P. N. (1982). Referential continuity and the coherence of discourse. *Cognition, 11*(1), 29–46.

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.

Grice, H. P. (1989). *Studies in the way of words*. Harvard UP, Cambridge, Mass.

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

Grishina, E., & Savchuk, S. (2009). Korpus ustnyh tekstov v NKRJa: Sostav i struktura [Corpus of oral texts in RNC: Content and structure]. In V. A. Plungjan, E. V. Rahilina, & T. I. Reznikova (Eds.), *Nacional'nyj korpus russkogo jazyka: 2006—2008. Novye rezul'taty i perspektivy [National corpus of the Russian language: 2006–2008. New results and perspectives]* (pp. 129—149). SPb.: NestorIstorija.

Grisot, C., & Blochowiak., J. (2021). Temporal relations at the sentence and text genre level: The role of linguistic cueing and non-linguistic biases—an annotation study of a bilingual corpus. *Corpus Pragmatics, 5*, 379–419.

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

Haberlandt, K. (1982). Reader expectations in text comprehension. In J.-F. Le Ny & W. Kintsch (Eds.), *Language and Comprehension* (pp. 239–249). Amsterdam: North-Holland.

Haberlandt, K., & Bingham, G. (1978). Verbs contribute to the coherence of brief narratives: Reading related and unrelated sentence triples. *Journal of Verbal Learning and Verbal Behavior, 17*, 419–425.

Halliday, M. (1987). Spoken and written modes of meaning. In R. Horowitz & S. J. Samuels (Eds), *Comprehending oral and written language* (pp. 55–82). San Diego: Academic Press. Inc.

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

Hart, B., & Risley, T.R. (1995). *Meaningful differences in the everyday experiences of young children*. Baltimore, MD, Brookes

Heim, I. (1982). The semantics of definite and indefinite noun phrases (Ph. D. dissertation). *University of Massachusetts, Amherst*.

Hempel, S., & Degand, L. (2008). Sequencers in different text genres: Academic writing, journalese and fiction. *Journal of Pragmatics, 40*, 676–693.

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

Hobbs, J. (1985). *On the coherence and structure of discourse*. Stanford, CA: CSLI.

Horowitz, R., & Samuels, S. J. (1987). Comprehending oral and written language: Critical contrasts for literacity and schooling. In Authors (Eds.), *Comprehending oral and written language* (pp. 1–52). San Diego: Academic Press. Inc.

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51*, 539–558.

Hussein, M. (2008). The truth-conditional/non-truth-conditional and conceptual/procedural distinctions revisited. *Newcastle Working Papers in Linguistics, 14*.

Huttenlocher, J., Vasilyeva, M., Waterfall. H., Vevea, J., & Hedges, L.V. (2007). Varieties of caregiver speech. *Dev. Psychol. 43*, 1062–1083.

Huttenlocher, J., Vasilyeva, M., Waterfall. H., Vevea, J., & Hedges, L.V. (2010). Sources of variability in children's language growth. *Cogn. Psychol. 61*, 343–365.

Imam, O. A., Mastura, M. A., Jamil, H., & Ismail, Z. (2014). Reading comprehension skills and performance in science among high school students in the Philippines. *Asia Pacific Journal of Educators and Education, 29*, 81–94.

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.

Iten, C. (2000). *'Non-truth-conditional' meaning: relevance and concessives.* PhD dissertation, University of London.

Jakubíček, M., Kilgarriff, A., Kovář, V., Rychlý, P., & Suchomel, V. (2013, July). The TenTen corpus family. 7th International Corpus Linguistics Conference CL, Lancaster, 125–127.

Kamp, H. (1981). A theory of truth and semantic representation. In J.A.G. Groenendijk, T.M.V. Janssen, & M.B.J. Stokhof (Eds.), *Formal methods in the study of language* (pp. 277–322). Amsterdam : Mathematical Centre.

Karlsson, J., Jolles, D., Koornneef, A., van den Broek, P., & Van Leijenhorst, L. (2019). Individual differences in children's comprehension of temporal relations: Dissociable contributions of working memory capacity and working memory updating. *Journal of experimental child psychology*, *185*, 1–18.

Katriel, T., & Dascal, M. (1984). What do indicating devices indicate? *Philosophy and Rhetoric, 171*, 1–15.

Katz, E.W., & Brent, S.D. (1968). Understanding connectives. *Journal of Verbal Learning and Verbal Behavior, 7,* 501–509.

Keenan, J. M., Baillet, S. D., & Brown, P. (1984). The effects of causal cohesion on comprehension and memory. *Journal of Verbal Learning and Verbal Behavior, 23*, 115–126.

Kehler. A. (1994). Temporal relations: reference or discourse coherence? *Proceedings of the 32nd annual meeting on Association for Computational Linguistics (ACL '94).* Association for Computational Linguistics, USA, 319–321.

Kidd, E., & Donnelly, S. (2020). Individual differences in first language acquisition. *Annual Review of Linguistics, 6*, 319–340.

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends Cogn Sci, 22*(2), 154–169.

Kidd, E., Lieven, E. V. M., & Tomasello, M. (2006). Examining the role of lexical frequency in the acquisition and processing of sentential complements. *Cognitive Development, 21*, 93–107.

Kleijn, S., Pander Maat, H. L. W., & Sanders, T. J. M. (2019). Comprehension effects of connectives across texts, readers, and coherence relations. *Discourse Processes, 56*, 447–464.

Knaebel, R., & Stede, M. (2022). Towards identifying alternative-lexicalization signals of discourse relations. *Proceedings of the 29th International Conference on Computational Linguistics*. International Committee on Computational Linguistics, Gyeongju, Republic of Korea, 837–850.

Knjazev, J. P. (2007). Ontogenez značenij obuslovlennosti [Ontogeny of causal meanings]. In S. N. Cejtlin (Ed.), *Semantičeskie kategorii v detskoj reči [Semantic categories in child speech]* (pp. 339–358). St. Petersburg, Russia: Nestor-Istorija.

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.

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

Koehn, P. (2005). Europarl: A Parallel Corpus for Statistical Machine Translation. *Conference Proceedings: the tenth Machine Translation Summit*, Phuket, Thailand, 79–86.

Koehne, J., & Demberg, V. (2013). The time-course of processing discourse connectives. *Proceedings of the 35th Annual Meeting of the Cognitive Science Society (CogSci).* Berlin, Germany.

Koornneef, A. W., & Van Berkum, J. J. A. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from selfpaced reading and eye-tracking. *Journal of Memory and Language, 54*, 445–465.

Korpershoek, H., Kuyper, H., & van der Werf, G. (2015). The relation between students' math and reading ability and their mathematics, physics, and chemistry examination grades in secondary education. *International Journal of Science and Mathematics Education, 13*, 1013–1037.

Kuperberg, G. R., Paczynski, M., & Ditman, T. (2011). Establishing causal coherence across sentences: An ERP study. *Journal of Cognitive Neuroscience, 23*, 1230–1246.

Kuznetsova, A., Bruun Brockhoff, P., & Christensen, R. H. B. (2014). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software, 82* (13), 1–26.

Lascarides, A., & Asher, N. (1993). Temporal interpretation, discourse relations and commonsense entailment. *Linguist Philos 16*, 437–493.

Lascarides, A., & Asher, N. (2007). Segmented discourse representation theory: Dynamic semantics with discourse structure. In H. Bunt & R. Muskens (Eds.), *Computing Meaning* (Vol. 3, pp. 87–124).

Leadholm, B., & Miller, J. F. (1992). *Language sample analysis: The Wisconsin guide*. Madison, WI: Department of Public Instruction.

Lenth, R. (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.5.1*. https://CRAN.R-project.org/package=emmeans

Liu, Y. (2019). Beyond the Wall Street Journal: Anchoring and comparing discourse signals across genres. *Proceedings of the Workshop on Discourse Relation Parsing and Treebanking 2019*. Association for Computational Linguistics, Minneapolis, USA, 72–81.

Liu, S., & Reynolds, B. L. (2022). Empirical support for the Involvement Load Hypothesis (ILH): A systematic review. *Behavioral Sciences*, *12*(10), 354.

Liu, Y., & Zeldes, A. (2019). Discourse relations and signalling information: Anchoring discourse signals in RST-DT. *Proceedings of the Society for Computation in Linguistics*, *2*, Article 35, 314–317.

Loban, W. (1976). *Language development: Kindergarten through grade twelve* (Research Report No. 18). ChampaignIL: National Council of Teachers of English.

Loman, N. L., & Mayer, R. E. (1983). Signaling techniques that increase the understandability of expository prose. *Journal of Educational Psychology, 75*, 402–412.

Lorch, R. F., & Lorch, E. P. (1986). On-line processing of summary and importance signals in reading. *Discourse Processes, 9*, 489–496.

Lyashevskaya, O., & Sharoff, S. (2009). *Chastotny slovar sovremennogo russkogo yazyka [Frequency dictionary of contemporary Russian]*. Azbukovnik, Moscow.

Mak, P., Lomako, J., Gagarina, N., Abrosova, E., & Tribushinina, E. (2020). Keeping two languages apart: Connective processing in both languages of Russian–German bilinguals. *Bilingualism: Language and Cognition, 23*, 532–541.

Mak, W. M., & Sanders, T. J. M. (2013). The role of causality in discourse processing: Effects of expectation and coherence relations. *Language and Cognitive Processes, 28*, 1414–1437.

Mak, W. M., Tribushinina, E., & Andreiushina, E. (2013). Semantics of connectives guides referential expectations in discourse: an eye-tracking study of Dutch and Russian. *Discourse Processes, 50*(8), 557–576.

Mann, W., & Thomson, S. (1986). Relational propositions in discourse. *Discourse Processes, 9*(1), 57–90.

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

Mano, Q. R., & Guerin, J. M. (2018). Direct and indirect effects of print exposure on silent reading fluency. *Reading and Writing, 31*, 483–502.

Martin, J. R. (1983). Conjunction: The logic of English text. In J. S. Petöfi & E. Sozer (Eds.), *Micro and macro connexity of texts* (pp.116–152). Hamburg: Helmut Buske Verlag.

Martin, J. R. (1992). *English Text: System and Structure*. Amsterdam and Philadelphia: John Benjamins.

Mauri, C., & Auwera, J. (2012). Connectives. In K. Allan & K. Jaszczolt (Eds.), *The Cambridge Handbook of Pragmatics* (Cambridge Handbooks in Language and Linguistics, pp. 377–402). Cambridge: Cambridge University Press.

McBride-Chang, C., & Chang, L. (1995). Memory, print exposure, and metacognition: Components of reading in Chinese children. *International Journal of Psychology, 30*, 607– 616

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

McClure, E.F., & Steffensen, M.S. (1985). A study in the use of conjunctions across grades and ethnic groups. *Research in the Teaching of English, 19*, 217–236.

Menzel, K. (2016). Understanding English-German contrasts: a corpus-based comparative analysis of ellipses as cohesive devices. [Doctoral dissertation, Saarland University]. Der Wissenschaftsserver der Universität des Saarlandes.

Meyer, B. J. F. (1975). *The organization of prose and its effects on memory*. Amsterdam: North-Holland.

Meyer, B. J. F., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly, 16*, 72–103.

Millis, K. K., Graesser, A. C., & Haberlandt, K. (1993). The impact of connectives on the memory for expository texts. *Applied Cognitive Psychology, 7*(4), 317–339.

Millis, K. K., & Just, M. A. (1994). The influence of connectives on sentence comprehension. *Journal of Memory and Language, 33*, 128–147.

Mills, K. L., Goddings, A. L., Herting, M. M., Meuwese, R., Blakemore, S. J., Crone, E. A., Dahl, R. E., Güroğlu, B., Raznahan, A., Sowell, E. R., & Tamnes, C. K. (2016). Structural brain development between childhood and adulthood: Convergence across four longitudinal samples. *NeuroImage*, *141*, 273–281.

Moeschler, J. (2002). Connecteurs, encodage conceptuel et encodage procédural. *Cahiers de linguistique française*, *24*, 265–292.

Moeschler, J. (2005). Connecteurs pragmatiques, inférences directionnelles et représentations mentales. *Cahiers Chronos, 12*, 35–50.

Morera, Y., León, J., Escudero, I., & de Vega, M. (2017). Do causal and concessive connectives guide emotional expectancies in comprehension? A double-task paradigm using emotional icons. *Discourse Processes 54*(8), 583–598.

Mullis, I. V. S., von Davier, M., Foy, P., Fishbein, B., Reynolds, K. A., & Wry, E. (2023). *PIRLS 2021 International Results in Reading.*Boston College, TIMSS & PIRLS International Study Center.

Murray, J. D. (1995). Logical connectives and local coherence. In J. R. F. Lorch & E. J. O'Brien (Eds.), *Sources of Coherence in Reading* (pp. 107-125). Hillsdale, NJ: Lawrence Erlbaum.

Murray, J. D. (1997). Connectives and narrative text: The role of continuity. *Memory & Cognition, 25*, 227–236.

Nation, P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher, 31*(7), 9–13.

Nation, I. S. P., & Coxhead, A. (2021). *Measuring native-speaker vocabulary size*. John Benjamins.

New, B., Pallier C., Ferrand L., & Matos R. (2001). Une base de données lexicales du français contemporain sur internet [A lexical database of contemporary French]. *LAnnée Psychologique, 101*, 447–462.

Nippold, M. (1993). Developmental markers in adolescent language: Syntax, semantics, and pragmatics. *Language, Speech, and Hearing Services in Schools, 24*, 21–28.

Nippold, M. (2004). Research on later language development international perspectives. In R. Berman (Ed.), *Language development across childhood and adolescence* (pp. 1–8). John Benjamins Publishing Company.

Nippold, M. A. (2006). Language development in school-age children, adolescents, and adults. In K. Brown (Ed.), *Encyclopedia of Language & Linguistics* (pp. 368–373) (2d ed.). Elsevier.

Nippold, M. (2008). *Later language development: school-age children, adolescents, and young adults* (3rd ed., 2nd printing). PRO-ED.

Nippold, M., Schwartz, I., & Undlin, R. (1992). Use and understanding of adverbial conjunctions: A developmental study of adolescents and young adults. *Journal of Speech and Hearing Research, 35*, 108–118.

Oğuz, E., & Özge, D. (2020). A Developmental Study of Turkish Connectives. *Dilbilim Araştırmaları Dergisi, 31*(2), 339–361.

O'Reilly, T., & McNamara, D. S. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "high-stakes" measures of high school students' science achievement. *American Educational Research Journal, 44*(1), 161–196.

Pander Maat, H., & Degand, L. (2001). Scaling causal relations and connectives in terms of speaker involvement. *Cognitive Linguistics, 12*(3), 211–245.

Péry-Woodley, M., Ho-Dac, L.M., Rebeyrolle, J., Tanguy, L., & Fabre, C. (2017). A corpus-driven approach to discourse organisation: from cues to complex markers. *Dialogue Discourse, 8*, 66–105.

Peterson, C. (1986). Semantic and pragmatic uses of 'but'. *Journal of Child Language, 13*, 583–590.

Pit, M. (2003). How to express yourself with a causal connective: Subjectivity and causal connectives in Dutch, German and French. Amsterdam: Rodopi.

Pons Bordería, S. (2001). Connectives/Discourse Markers. An Overview. *Quaderns de Filologia. Estudis Literaris.*, *VI*, 219–243.

Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A., & Webber, B. (2008). *The penn discourse treebank 2.0.* Paper presented at the 6th International Conference on Language Resources and Evaluation (LREC).

Prasad, R., Joshi, A., & Webber, B. (2010). *Realization of Discourse Relations by Other Means: Alternative Lexicalizations.* Paper presented at the COLING 2010, Beijing.

Prasad, R., Webber, B., Lee, A., & Joshi, A. (2019). *Penn Discourse Treebank Version 3.0*. Philadelphia: Linguistic Data Consortium.

Pyykkönen, P., & Järvikivi, J. (2010). Activation and persistence of implicit causality information in spoken language comprehension. *Experimental psychology*, *57*(1), 5–16.

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

RAND Reading Study Group, & Snow, C. (2002). *Reading for Understanding: Toward an R&D Program in Reading Comprehension*. RAND Corporation.

Ravid, D., & Tolchinsky, L. (2002). Developing linguistic literacy: A comprehensive model. *Journal of Child Language*, *29,* 419–448.

Rebuschat, P. (Ed.). (2015). *Implicit and explicit learning of languages*. Amsterdam: John Benjamins.

Riegel, M., Pellat, J.-C., & Rioul, R. (2021). *Grammaire méthodique du français*. (8th ed.). Paris: Presses Universitaires de France.

Rohde, H., & Horton, W. S. (2014). Anticipatory looks reveal expectations about discourse relations. *Cognition, 133*, 667–691.

Roulet, E., Auchlin, A., Moeschler, J., Rubattel, C., & Schelling, M. (1985). *L'Articulation du Discours en Français Contemporain.* Berne: Peter Lang.

Roze, C., Danlos, L., & Muller, P. (2012). LEXCONN: A French lexicon of discourse connectives. *Discours*, *10*, 1–15.

R Core Team. (2020). *R: A language and environment for statistical computing. [Computer software manual].* Vienna, Austria. http://www.R-project.org/

Salihu, L., Aro, M., & Räsänen, P. (2018). Children with learning difficulties in mathematics: Relating mathematics skills and reading comprehension. *Issues in Educational Research, 28*, 1024–1038.

Sanders, T. (1997). Semantic and pragmatic sources of coherence: On the categorization of coherence relations in context. *Discourse Processes, 24*(1), 119–147.

Sanders, T. (2005). Coherence, causality and cognitive complexity in discourse. *Proceedings/ Actes SEM-05, First International Symposium on the Exploration and Modelling of Meaning*, 105–114.

Sanders, T., Demberg, V., Hoek, J., Scholman, M., Asr, F., Zufferey, S., & Evers-Vermeul, J. (2018). Unifying dimensions in coherence relations: How various annotation frameworks are related. *Corpus Linguistics and Linguistic Theory*, *17*(1), 1–71.

Sanders, T., & Spooren, W. (2007). Discourse and text structure. In D. Geeraerts & H. Cuykens (Eds.), *Handbook of cognitive linguistics* (pp. 37–60). Oxford: Oxford University Press.

Sanders, T., Spooren, W., & Noordman, L. (1992). Towards a taxonomy of coherence relations. *Discourse Processes, 15*, 1–36.

Sanders, T., & Noordman, L. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes, 29*(1), 37–60.

Santrock, J.W. (1996). *Adolescence* (6th ed.). Madison, WI: Brown & Benchmark.

Schiffrin, D. (1987). *Discourse Markers*. Cambridge: Cambridge University Press.

Schleicher, A. (2019). *PISA 2018. Insights and Interpretations*. Paris: OECD Publishing.

Scholman, M.C.J., & Demberg, V. (2017). Crowdsourcing discourse interpretations: On the influence of context and the reliability of a connective insertion task. *Proceedings of the 11th Linguistic Annotation Workshop*, Valencia, Spain.

Scholman, M.C.J., Demberg, V., & Sanders, T. J. M. (2020). Individual differences in expecting coherence relations: Exploring the variability in sensitivity to contextual signals in discourse. *Discourse Processes*, *57*(10), 844–861.

Schreiber-Gregory, D. N. (2018). Ridge regression and multicollinearity: An in-depth review. *Model Assisted Statistics and Applications, 13*, 359–365.

Schwab, J., & Liu, M. (2020). Lexical and contextual cue effects in discourse expectations: Experimenting with German 'zwar…aber' and English 'true/sure…but'. *Dialogue & Discourse, 11*(2), 74–109.

Segal, E. M., Duchan, J. F., & Scott, P. J. (1991). The role of interclausal connectives in narrative structuring: Evidence from adults' interpretations of simple stories. *Discourse Processes, 14*, 27–54.

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson, & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112–133). Cambridge, UK: Cambridge University Press.

Spear-Swerling, L., Brucker, P. O., & Alfano, M. P. (2010). Relationships between sixth-graders reading comprehension and two different measures of print exposure. *Reading and Writing: An Interdisciplinary Journal, 23*(1), 73–96.

Spenader, J. (2018). Children's comprehension of contrastive connectives. *Journal of Child Language*, *45*, 610–640.

Sperber, D. (2001). An evolutionary perspective on testimony and argumentation. *Philosophical Topics,* 29*,* 401–413.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition* (1st ed.). Oxford: Blackwell Publishers Ltd.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Blackwell, Oxford and Harvard UP, Cambridge MA.

Spooren, W., & Sanders, T. (2008). The acquisition order of coherence relations: on cognitive complexity in discourse. *Journal of Pragmatics, 40*, 2003–2026.

Spyridakis, J. H., & Standal, T. C. (1987). Signals in expository prose. *Reading Research Quarterly, 22*, 285–298.

Stainthorp, R. (1997). A childrens author recognition test: A useful tool in reading research. *Journal of Research in Reading, 20*(2), 148–158.

Stanovich, K., & West, R. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24*(4), 402–433.

Stanovich, K., West, R., & Harrison, R. (1995). Knowledge growth and maintenance across the life span. The role of print exposure. *Developmental Psychology, 31*(5), 811–826.

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological methods, 14*(4), 323–348.

Stukker, N., & Sanders, T. (2012). Subjectivity and prototype structure in causal connectives: A cross-linguistic perspective. *Journal of Pragmatics, 44*, 169–190.

Sweetser, E. (1990). *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge University Press, Cambridge.

Taylor, K.S., Lawrence, J.F., Connor, C.M., & Snow, C. (2019). Cognitive and linguistic features of adolescent argumentative writing: Do connectives signal more complex reasoning?. *Reading and Wriing,* 32, 983–1007.

Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics, 11*, 61–82.

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.

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

Traxler, M. J., Sanford, A. J., Aked, J. P., & Moxey, L. M. (1997). Processing causal and diagnostic statements in discourse. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 88–101.

Tribushinina, E., Dubinkina, E., & Sanders, T. (2015). Can connective use differentiate between children with and without specific language impairment? *First Language, 35*(1), 3–26.

Tribushinina, E., Mak, W.M., Andreiushina, E., Dubinkina, E., & Sanders, T. (2017). Connective use by bilinguals and monolinguals with SLI. *Bilingualism: Language and Cognition 20*(1), 98−113.

Tribushinina, E., Valcheva, E., & Gagarina, N. (2017). Acquisition of additive connectives by Russian-German bilinguals: A usage-based approach. In J. Evers-Vermeul & E. Tribushinina (Eds.), *Usage-based approaches to language acquisition and language teaching* (pp. 207–232). Berlin & Boston: Mouton.

Tskhovrebova, E., Wetzel, M., Gygax, P., & Zufferey, S. (2023, submitted). The role of training and exposure to print for the mastery of connectives in French.

Tskhovrebova, E., Zufferey, S., & Gygax, P. (2022). Individual variations in the mastery of discourse connectives from teenage years to adulthood. *Language Learning, 72*(2), 412–455.

Tskhovrebova, E., Zufferey, S., & Gygax, P. (2023, submitted). Exploring the sensitivity to non-connective signals of coherence relations: The case of French speaking teenagers.

Tskhovrebova, E., Zufferey, S., & Tribushinina, E. (2022). French-speaking teenagers' mastery of connectives: The role of vocabulary size and exposure to print. *Applied Psycholinguistics, 43*(5), 1141–1163.

Tskhovrebova, E., Zufferey, S., & Tribushinina, E. (2023, submitted). The mastery of mono- and polyfunctional connectives in Russian: The role of vocabulary size and exposure to print.

Uccelli, P., Dobbs, C. L., & Scott, J. (2013). Persuasive writing of high school students. *Written Communication, 27*, 66.

Ullman, M.T. (2001). The declarative/procedural model of Lexicon and Grammar. *Journal of Psycholinguistic Research, 30*, 37–69.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP). (2022). *Reading Assessments*.

Van Dijk, T. (1977). Connectives in text grammar and text logic. In T. Dijk & J. Petöfi (Eds.), *Grammars and descriptions: Study in text theory and text analysis* (pp. 11–63). Berlin, Boston: De Gruyter.

Van Silfhout. (2014). Fun to read or easy to understand? Establishing effective text features for educational texts on the basis of processing and comprehension research [Doctoral dissertation, Unniversity of Utrecht]. Utrecht University Repository.

Van Silfhout, G., Evers-Vermeul, J., & Sanders, T. J. M. (2015). Connectives as processing signals: how students benefit in processing narrative and expository texts. *Discourse Processes, 52*(1), 47–76.

Van Veen, R.V. (2011). *The acquisition of causal connectives: The role of parental input and cognitive complexity* (Doctoral dissertation). Retrieved from https://www.lotpublications.nl.

Van Veen, R., Evers-Vermeul, J., Sanders, T., & Van den Bergh, H. (2009). Parental input and connective acquisition in German: A growth-curve analysis. *First Language, 29*, 267–89.

Van Veen, R., Evers-Vermeul, J., Sanders, T., & Van Den Bergh, H. (2013). The influence of input on connective acquisition: A growth curve analysis of English because and German weil. *Journal of Child Language, 40*(5), 1003–1031.

VanPatten, B., & Smith, M. (2022). *Explicit and implicit learning in second language acquisition* (Elements in Second Language Acquisition). Cambridge: Cambridge University Press.

Volodina, A., Heppt, B., & Weinert, S. (2021). Relations between the comprehension of connectives and school performance in primary school. *Learning and Instruction*, *74*, 101430.

Volodina, A., & Weinert, S. (2020). Comprehension of Connectives: Development Across Primary School Age and Influencing Factors. *Front. Psychol. 11*, 814.

Webber, B., Prasad, R., Lee, A., & Joshi, A. (2019). The penn discourse treebank 3.0 annotation manual. *Philadelphia, University of Pennsylvania*. 35, 108.

Welie, C., Schoonen, R., Kuiken, F., & van den Bergh, H. (2017). Expository text comprehension in secondary school: For which readers does knowledge of connectives contribute the most*? Journal of Research in Reading, 40*, S42–S65.

Wetzel, M., Crible, L., & Zufferey, S. (2022). Processing clause-internal discourse relations in a second language: a case study of specifications in German and French. *J. Second Lang. Stud. 5* (2), 206e234.

Wetzel, M., Zufferey, S., & Gygax, P. (2020). Second language acquisition and the mastery of discourse connectives: Assessing the factors that hinder L2-learners from mastering French connectives. *Languages*, *5*(3), 35.

Wilson, D. (2011). The conceptual-procedural distinction: past, present and future. In V. Escandell-Vidal, M. Leonetti & A. Ahern (Eds.), *Procedural Meaning: Problems and Perspectives* (pp. 3–31). Bingley: Emerald Group Publishing.

Wilson, D., & Sperber, D. (1993). Linguistic form and relevance. *Lingua,* 90, 1–25.

Wimmer, L., & Ferguson, H.J. (2022). Testing the validity of a self-report scale, author recognition test, and book counting as measures of lifetime exposure to print fiction. *Behavior Research Methods*.

Wong, W., & Simard, D. (2015). *Focusing on form in language instruction* (1st ed.). New York: Routledge.

Xiang, M., & Kuperberg, G. (2015). Reversing expectations during discourse comprehension. *Language, Cognition and Neuroscience, 30*, 648–672.

Xu, X., Jiang, X., & Zhou, X. (2015). When a causal assumption is not satisfied by reality: Differential brain responses to concessive and causal relations during sentence comprehension. *Language, Cognition and Neuroscience, 30*(6), 704–715.

Xu, X., Chen, Q., Panther, K.-U., & Wu, Y. (2018). Influence of causal and concessive conjunctions of pragmatic processing: online measures from eye movements and self-paced reading. *Discourse Processes 55*(4), 387–409.

Yefremova, T. F. (Ed.). (2000). Novyy slovar' russkogo yazyka. Tolkovo-slovoobrazovatel'nyy [New dictionary of the Russian language. With definitions and derivations]. Moscow, Russia: Russkiy yazyk.

Yevgen'eva, A. P. (1999). Slovar' russkogo yazyka: V 4-kh t., RAN, In-t lingvistich. issledovaniy (4th ed.) [Dictionary of the Russian language: In 4 volumes, Russian academy of sciences, Institute of linguistics]. Moscow, Russia: Russkiy yazyk, Poligrafresursy.

Yung, F., Demberg, V., & Scholman, M. (2019). Crowdsourcing discourse relation annotations by a two-step connective insertion task. *Proceedings of the 13th Linguistic Annotation Workshop*, Florence, Italy, 16–25.

Zeldes, A. (2017). The GUM corpus: Creating multilayer resources in the classroom. *Language Resources and Evaluation, 51*(3), 581–612.

Zeldes, A., & Liu, Y. (2020). A neural approach to discourse relation signal detection. *Dialogue & Discourse, 11*(2), 1–33.

Zufferey, S. (2010). *Lexical pragmatics and theory of mind: The acquisition of connectives.* John Benjamins Publishing Company.

Zufferey, S., & Gygax, P. (2016). The role of perspective shifts for processing and translating discourse relations. *Discourse Processes, 53*(7), 532–555.

Zufferey, S., & Gygax, P. (2020a). “Roger broke his tooth. However, he went to the dentist”: Why some readers struggle to evaluate wrong (and right) uses of connectives. *Discourse Processes, 57*(2), 184–200.

Zufferey, S., & Gygax, P. (2020b). Do teenagers know how to use connectives from the written mode? *Lingua, 234,* 102779, 1–12.

Zufferey, S., Mak, W.M., & Sanders, T. (2015). A cross-linguistic perspective on the acquisition of causal connectives and relations. *International Review of Pragmatics, 7*(1), 22–39.