6 Processing and Understanding Discourse Relations and Connectives

6.1 INTRODUCTION

In this chapter, we discuss the way people read, remember and understand discourse, depending on the type of relations that link discourse segments together. We also illustrate the role of connectives and other discourse signals as elements guiding readers' interpretation. Throughout the chapter, we review empirical evidence from experiments that involve various methodologies. Some of them are simple enough, as they only consist of asking people to remember elements from a discourse, or to answer targeted comprehension questions about its content. However, other studies involve a more complex methodology, especially those involving online reading. It is important to have a general idea of the way they work in order to be able to critically assess their findings. We therefore present them briefly here.

The simplest method to study online reading is called self-paced reading. In a self-paced reading task, participants read sentences appearing segment by segment on a computer screen. Usually, segments disappear once they have been read, in order to prevent participants from displaying the whole sentence first before actually reading it, as illustrated in Figure 6.1, in which each word would appear one at a time on the same line.

But this also means that reading is not entirely natural, as people cannot go back to reread previous segments once they have moved on, as they usually do. Participants can pace their reading by pressing a computer key (typically the spacebar) to move on to the next segment when they have finished reading the current segment. By comparing the time it takes readers to move on to the next segment in each experimental condition, researchers can infer the processing complexity of a given segment. In these experiments, a discourse can be divided into sentences, groups of words or even single words. The more fine-grained the division, the better it allows researchers to locate the source of an effect, but the less natural the reading becomes. For this

This .. …..….. ……..
…. is …..….. ……..
…. .. self-paced …….
…. .. ……. reading.

Figure 6.1 Illustration of a self-paced reading task

reason, most studies divide sentences into two or three reading segments corresponding to its meaning components, as illustrated in (1) in which the six regions are indicated between slashes, and represent a division between subject, verb and complement.

(1) Lisa / enjoys being / on her own / and Kurt / enjoys being / with other people.

[from Crible & Pickering, 2020: 871]

The rationale behind self-paced reading is that longer processing times reflect greater processing difficulty. An important aspect of these reading studies is that the critical discourse segment, for which reading times are recorded, must be strictly identical across all conditions so that reading times are comparable. The difference between them is therefore created by manipulating the content of the text preceding this critical segment.

Another more sophisticated measure of online reading can be obtained by using eye-tracking. This method consists of making people read sentences on a computer screen while an infrared camera detects their corneal reflection and infers with great precision their gaze direction on the screen. Eye-tracking is more informative than self-paced reading as it enables participants to read sentences in a normal way rather than segment by segment, and regressions to previous portions of a text for rereading are also recorded. Thus, eye-tracking enables researchers to compute the time that people spend on a region when they first read it, but also when they come back to it later on (as well as many other fine-grained processing measures). Other methodological aspects of these experiments are very similar to self-paced reading, as critical segments must also be identical between conditions to make them comparable. An important practical difference between them however is that eye-tracking still mostly requires that participants come to the lab for testing whereas self-paced reading can be conducted remotely. Online recruiting represents a big advantage to reach a larger and more diverse sample of participants.

Finally, we will discuss studies that measure brain activity in the form of Event Related Potentials (ERPs). ERP studies measure neural

processes with a very good temporal resolution. Since this method was developed, several electric waves have been associated with language processing. One of them is the N400 effect, a negative-going deflection that peaks at around 400 milliseconds after the stimulus onset, hence its name. In the literature, its amplitude has often been found to be related with a revision of an expectation, and this signal is often relevant when comparing coherent and incoherent discourse. Another relevant signal is the P600, a late positive component often associated with syntactic reanalysis and repair, but also with semantic or pragmatic violations, and with the updating of mental information in a discourse (Brouwer, Fitz & Hoeks, 2012). Throughout the chapter, we will see how all these methods complement each other and allow us to reach a global picture of discourse processing and comprehension.

To begin this chapter, we will discuss the processing and comprehension of different types of discourse relations, without paying too much attention to the way they are conveyed. In a second part of the chapter, we will address the question of the way these relations are signaled, and underline the many ways in which the presence of connectives influences the processing, recall and understanding of discourse. We will also address the role of alternative cues such as the repetitive argument structure typically found across the two segments of contrastive relations. We will see that most studies testing the role of connectives have focused on the handful of lexical items that are often used in spoken language, but we will mention studies that have specifically focused on connectives bound to the written mode, such as *therefore* and *nevertheless* in English. We will also address the ways in which ambiguous connectives, in other words, connectives that can convey different relations depending on context, are processed during reading. In a third part of the chapter, we will address the question of individual differences between adult native speakers, and argue that even though studies are still scarce on this topic, there is good evidence that these differences exist, and affect the way people use and understand connectives in their native language. A final part of the chapter will be dedicated to reviewing studies that have underlined cross-linguistic similarities and differences in the way discourse relations are processed across languages, depending on the way these relations are marked by connectives. This last part will launch the discussion on crosslinguistic comparisons, that will be dealt with in more detail in Chapter 7.

6.2 PROCESSING DISCOURSE RELATIONS

The notion of discourse relations plays a central role in explaining the local coherence of a discourse across all major theoretical models of discourse representation (see Chapter 2). It has therefore been an important empirical question at least since the 1980s to determine whether this notion is merely an analytic one, useful for the linguistic description of discourse structure, as has been suggested by Grosz and Sidner (1986), or whether it is a cognitive notion that plays a role in the way readers process and understand discourse, as many other authors have argued (e.g., Hobbs, 1983; Mann & Thompson, 1986; Sanders, Spooren & Noordman, 1992). In this respect, among the numerous studies that have assessed the processing of discourse relations, almost all of them have clearly demonstrated that they are of paramount importance for the way adult native speakers process discourse.

To begin with, evidence for the importance of easily inferable discourse relations for text comprehension comes from experiments in which the local coherence of real academic textbooks has been improved by adding links such as causal explanations in short passages, and have compared the ability of readers to understand the content presented in the various versions (e.g., Beck et al., 1991; McNamara et al., 1996). Results consistently indicate that texts with improved local discourse relations were understood better. Conversely, other studies have shown that when discourse relations cannot be inferred because the signal given by a connective is incompatible with the linguistic content of the segments, as in (2), processing is heavily disrupted (Murray, 1997; Ferstl & von Cramon, 2001; Xu et al., 2018). We will come back to these studies in more detail in the next section, when discussing the role of connectives for discourse processing.

(2) Manny needed to publicize the garage sale. However, he arranged for flyers to be made.

[from Murray, 1997: 231]

The cognitive account of discourse relations also implies that different relations should have a different effect for discourse processing and comprehension, as each of them create different coherence links between discourse segments. For example, in the Cognitive model of Coherence Relations put forward by Sanders et al. (1992), relations vary in their degree of cognitive complexity, and this complexity can be precisely determined by the values that each relation takes on four primitives, namely: basic operation, order of the segments, source of coherence and polarity (see Chapter 2). There is indeed a lot of evidence

in the literature that discourse relations do not all play the same role for discourse processing.

Early studies focused mostly on the relation of causality, and compared its effects to that of other relations, or even to the effect of unrelated segments. For example, Haberlandt and Bingham (1978) found that causally related sentences were read more quickly compared to noncausally related ones. Similarly, Keenen, Baillet and Brown (1984) found that reading times decreased as the level of causality increased between sentences. Other studies have also found that in short narratives, causally related sentences were remembered better (e.g., Trabasso, Secco & van den Broeck, 1984). These early studies therefore underlined the importance of causality for discourse processing and comprehension. They did not, however, systematically compare causal relations to other relation types, as the noncausal relations in these studies often encompassed a variety of different links.

In this respect, Sanders and Noordman (2000) were the first to systematically compare the processing and recall of two types of discourse relations, conveyed with and without connectives. We will focus on the difference between relations in this section. The two relations they analyzed were the problem–solution relation, illustrated in (3), that can be considered as a subtype of causal relations, and the list relation, illustrated in (4), that represents a subtype of additive relations.

(3) It has been dangerous to cross that street for years. The city council has now decided to build a subway for pedestrians.

(4) The city council has decided to build a subway for pedestrians. New traffic lights will be installed nearby.

[from Sanders & Noordman, 2000: 38]

These two types of relations differ in the degree of coherence they create in discourse, as causal relations between events typically create a higher degree of coherence by providing an explanation, compared to mere added facts. The authors therefore hypothesized that causal relations could have a higher effect on discourse processing and recall compared to the less informative list relations. However, causal relations are also more cognitively complex than additive relations in Sanders, Spooren and Noordman's taxonomy (1992). It could therefore also be expected that having to deal with this additional complexity might slow down the processing of causal relations compared to list relations.

In order to assess these hypotheses, Sanders and Noordman (2000) designed a reading experiment in the form of a self-paced reading task. Results revealed that problem–solution and list relations were not processed in the same way, as critical segments involving causal relations were read more quickly compared to segments involving a list relation. This result therefore means that cognitively more complex relations like causal relations do not necessarily involve greater processing difficulty. A possible explanation put forward by Sanders and Noordman is that causal relations have a special status in discourse processing, as readers expect to find explanations for the described problems. Taking this observation one step further, Sanders (2005) later suggested a 'causality-by-default' hypothesis, stating that causality represents a default interpretation in discourse. This would explain why readers access other interpretations only when a causal meaning cannot be obtained in context, and why causal relations are accessed more quickly than simpler but less informative ones.

In addition to reading time differences, the two relations also differed in terms of recall, as participants remembered elements from causal relations better than elements from list relations. This was apparent both in a task occurring immediately after reading, in which participants had to decide if statements had been presented in what they just read or not, and in a free recall task in which participants had to remember as many elements from the discourse as possible. These results provide further proof of the special status of causal relations in the mental models that readers create after reading a discourse.

More recently, a number of studies have focused on the comparison between causal and concessive relations. This comparison is particularly revealing, as both relations involve a causal link between the segments, but in the case of concessive relations, an expectation created in one of the segments is denied in the other (e.g., Louwerse, 2001) as illustrated in (5). In this sentence, Mary's state of happiness leads the reader to infer a positive outcome in the competition, but this expectation is explicitly denied in the second segment. For this reason, Sanders, Spooren and Noordman (1992) classify concessions as negative causal relations.

(5) Mary is very happy although she lost the competition she hoped to win.

[constructed example]

Studies that have compared the way readers process causal and concessive relations have all found important differences between them. First, causal relations are processed more quickly than concessive

relations (Köhne & Demberg, 2013; Xu et al., 2018). Differences were also found between tasks involving offline comprehension and recall. For instance, people found it more difficult to remember sentences that contain the concessive connective *but* compared to the causal connective *because* (Caron, Micko & Thüring, 1988) and have more difficulties filling in blank slots between sentences when a concessive relation is involved (Goldman & Murray, 1992). Note, in addition, that these studies have been conducted in different languages, namely English, German, Dutch and Chinese, which indicates that the observed differences are not due to specific features of the connectives used to convey them in a given language. We will discuss cross-linguistic similarities and differences in the way connectives encode discourse relations in more detail in Chapter 7 and mention cases in which these cross-linguistic differences have an impact on discourse processing in Section 6.6.

Finally, differences between causal and concessive relations have also been underlined in studies that have looked at the processing of incoherent relations, that is, relations for which the signal given by the connective is incompatible with the linguistic content of the segments. While an effect of incoherence is always found for causal relations, this effect is not always apparent for concessive relations in studies relying on reading times, possibly because processing cost is already high for congruent concessive relations and this might obscure differences between conditions in some experimental contexts (see Lyu, Tu & Lin, 2020 for a discussion). The overarching conclusion from all these studies is that the processing cost of concessive relations is higher compared to that of causal relations. We will come back to this issue in the next section, when discussing concessive connectives.

Taken together, all the studies we have discussed so far consistently demonstrate that not all discourse relations are processed in the same way. It seems that causal relations have a privileged status, as they create a high degree of coherence between discourse segments. However, the category of causal relations itself can be decomposed into several subcategories (e.g., Sweetser, 1990). An important difference, which has been tested in several experiments, is the one separating objective causal relations, such as (6), that relate objective facts or events occurring in the world, to subjective causal relations, such as (7), relating claims and conclusions derived in the mind of the speaker (see Chapter 3).

(6) Heidi was proud and happy, because she won first prize at the art show.

(7) Heidi could imagine and create things, because she won first prize at the art show.

[from Traxler, Bybee & Pickering, 1997: 485]

Traxler, Bybee and Pickering (1997) have compared the way readers process sentences like (6) and (7) using eye-tracking. In this experiment, the region of interest is the one following the connective ('she won first prize at the art show'). For this reason, the words used in this segment were identical in objective and subjective relations. The different interpretation of the relation comes from the indications given in the first segment. In (6) the fact of winning first prize is the reason for Heidi's objective state of happiness, while in (7) the fact of winning first prize merely leads the speaker to conclude in their mind that she is probably good at imagining and creating things. This experiment uncovered an important processing difference between objective and subjective causality, as readers slowed down when they processed subjective relations compared to objective ones. This effect occurred at the words 'first prize', because it is the point in the sentence when an objective interpretation can be ruled out. The authors therefore conclude that having to infer a subjective meaning is what slowed readers down compared to the default objective interpretation. Further indications that this explanation is on the right track comes from another set of experiments (Traxler et al., 1997) in which the authors inserted a marker of subjectivity already in the first segment, as in (8).

(8) Eugene thinks that Heidi could imagine and create things because she won first prize at the art show.

[from Traxler et al., 1997: 99]

In this experimental condition, the additional reading time needed to process subjective relations in the second segment disappeared, as readers already had the right relation in mind, thanks to the cue given in the first segment. Canestrelli, Mak and Sanders (2013) found a similar asymmetry in Dutch, both with backward causal connectives like *because* but also with forward causal connectives like *so* in English. Similarly, the addition of a subjectivity marker also removed the extra processing cost in these experiments. We discuss them in more detail in Section 6.6. More recent experiments have shown, however, that adding a subjectivity marker does not always eliminate the extra processing cost of subjective relations (Kleijn, Mak & Sanders, 2021), and more research is still needed to determine why some markers orient readers towards subjectivity while others don't.

The difference between objective and subjective relations does not only involve a processing difference. These relations also play a different role in persuasive texts. Indeed, subjective relations provide readers with indications that the text contains subjective claims rather than objective facts, and this clue forewarns readers, leading them to be more critical towards its content (Kamalski, Sanders & Lentz, 2008). This effect was measured by comparing texts including objective causal connectives, subjective causal connectives and no connectives. Results showed both types of connectives had a reverse effect on persuasion compared to implicit relations. While objective connectives improved the integration of information and therefore increased persuasion, subjective connectives alerted the readers to the persuasive intent of the author, which caused resistance and led to a less persuasive outcome. We can therefore conclude that the use of connectives is not neutral as a way to mark discourse relations, but the effect created by specific connectives should be analyzed separately. We discuss the role of connectives further in the next section.

6.3 THE ROLE OF CONNECTIVES FOR DISCOURSE PROCESSING

In pragmatics and discourse analysis, connectives are often described as lexical items encoding procedural instructions guiding discourse interpretation (see Chapter 3). Processing studies can shed further light on their role as guides for interpretation. As we will see in this section, connectives do indeed have an immediate effect on discourse processing, by speeding-up the processing of the words immediately following them, but also by reversing expectations. Concessive connectives in particular immediately reverse readers' real-world expectations about the content of the following segment (e.g., Xiang & Kuperberg, 2015). In addition, some connectives that are often used in a pair, like *on the one hand*, raise expectations about a following contrast that readers sustain even with intervening linguistic materials between the two related segments (Scholman, Rohde & Demberg, 2017). Other linguistic signals, for example, the adverb *zwar* in German ('yet, however'), similarly raise expectations about an upcoming concession (Schwab & Liu, 2020). We will review these various effects of connectives in detail in this section, starting with the difference between relations containing and not containing connectives.

6.3.1 Processing Explicit versus Implicit Relations

Since the earliest attempts to measure discourse processing, connectives have repeatedly been found to speed up processing of the segment

they introduce compared to implicit relations (e.g., Britton et al., 1982; Haberlandt, 1982). In their experiment comparing causal and additive relations, Sanders and Noordman (2000) confirmed that the use of a connective speeded up processing for both problem–solution relations and list relations compared to implicit relations. In that sense, there does not seem to be a processing advantage for causal connectives compared to connectives indicating other relations. An important difference does exist, however, between relations. This difference is linked to the necessity of using a connective to convey the relation. While some relations like causality (9) are easily inferable even when they are conveyed implicitly in the absence of a connective as in (10), others like concession (11) are more difficult to interpret when communicated implicitly (12).

(9) Mary was very happy because she won the competition.

(10) Mary was very happy. She won the competition.

(11) Mary was very happy although she lost the competition.

(12) Mary was very happy. She lost the competition.

[constructed examples]

This difference between causal and concessive connectives has been tested empirically in a series of experiments. Murray (1995) compared the online processing of concessive relations compared to relations of cause and addition in self-paced reading experiments. All relation types were conveyed both explicitly with a connective in half of the sentences and implicitly in the other half. Overall, Murray found that the implicit version did not trigger longer reading times compared to explicit relations. However, a significant difference was observed specifically in the case of concessive relations, for which the implicit version triggered longer reading times compared to the explicit one. This difference led Murray to conclude that concessive connectives have a greater impact on creating intercausal coherence compared to causal and additive connectives. This interpretation is in line with the continuity hypothesis put forward by Segal, Duchan and Scott (1991). Following this hypothesis, a continuous relation introduces new information that can be integrated in the ongoing construction of the narrative's meaning, whereas a discontinuous relation introduces a rupture in the narration. This rupture must be indicated explicitly with a connective. As observed in his experiment, failure to indicate this rupture disrupts processing, while the removal of a causal or additive connective does not create the same problem.

In another set of experiments, Murray (1997) assessed the impact of an incoherence created by the improper use of a concessive, causal or

additive connective in contexts that were not congruent with their meaning. For example, the use of a concessive connective between two segments that typically indicate a consequence–cause relation as in (9). In this experiment, Murray only compared the inappropriate use of connectives with their absence (i.e., implicit relations). Results indicate that reading times globally increase for sentences conveyed by an inappropriate connective compared to no connective. This effect was however stronger for inappropriately used concessive connectives compared to the other two relations. Murray concludes once again that concessive connectives are more indispensable than causal and additive connectives for discourse processing because they introduce a discontinuous relation that must be explicitly marked to be processed adequately. Conversely, continuous relations such as causality and addition need not be marked explicitly by a connective to be smoothly processed.

Since then, other studies have assessed the role of connectives for causal and concessive relations, using other experimental methods, contexts and in different languages. In one of them, Köhne and Demberg (2013) found that causal relations conveyed by German connectives raise an expectation of congruence with the preceding context in a visual world experiment using eye-tracking in which participants look at images while hearing linguistic stimuli, whereas concessive connectives raise expectations of incongruence leading people to look for alternative referents. This means that causal and concessive connectives are both rapidly integrated during sentence processing but that concessive connectives involve a more complex mental operation, which could explain why they trigger longer reading times in reading experiments. In a study involving Event Related Potentials (ERP), Xiang and Kuperberg (2015) found that reading a connective with a negative polarity like *even so* in English, immediately reversed people's expectation about coherence, as sentences like (13) became more coherent than sentences like (14).

(13) Elizabeth had a history exam on Monday. She took the test and failed it. Even so, she went home and celebrated wildly.

(14) Elizabeth had a history exam on Monday. She took the test and aced it. Even so, she went home and celebrated wildly.

[from Xiang & Kuperberg, 2015: 649]

In this experiment, the N400, a signal often associated with the revision of expectations, was smaller when participants evaluated the coherence or even simply read sentences like (13) than (14), which

means that the information provided by the connective quickly overrode world knowledge about the consequences of a failing a test. This experiment also provided further confirmation that the processing of negative relations comes at a cost, as a sustained negativity effect was found even at the end of the sentence. A similar negativity effect (N400) linked to incongruence was also evidenced in another ERP study comparing the roles of causal and concessive connectives in German and English (Drenhaus et al., 2014). In this study, a later positivity effect (P600) was also evidenced for concessive relations. This later effect can be associated with the cost of predicting and revising a causal relation. These brain studies therefore reveal the presence of neural correlates for the added complexity involved in the processing of concessive relations evidenced with behavioral measures such as reading.

In another line of enquiry, Zufferey and Gygax (2016) extended the notion of continuity to include not only temporally and causally continuous relations, but also relations that are continuous or discontinuous with respect to the narrative perspective presented in the first segment. Continuous relations from a narrative perspective involve the same narrative voice across the two segments, as in the relation of elaboration exemplified in (15). By contrast, relations that are discontinuous from a narrative perspective involve a change in narrative voice between the two segments, as in the relation of correction in (16) and confirmation in (17). While the narration always takes the external perspective of Emma in the first segment, the voice changes to that of the narrator in the second segment of (16) and (17), who intervenes to either correct Emma's hypothesis (16) or confirm its veracity (17).

(15) Emma thought that Bill would Marry Ann. She hoped that they would invite her to their wedding.

(16) Emma thought that Bill would marry Ann, but the marriage did not take place.

(17) Emma thought that Bill would marry Ann. And indeed, the marriage took place last June.

[from Zufferey & Gygax, 2016: 535]

In a self-paced reading experiment, Zufferey and Gygax (2016) compared the processing of causal and confirmation relations either conveyed explicitly by the polyfunctional French connective *en effet* that can be used in both causal and confirmation contexts, or implicitly. They found that using the connective immediately speeds-up processing at the beginning of the second clause, independently of

the relation involved. They therefore replicated previous findings on the role of connectives to speed-up discourse processing discussed at the beginning of this section. But they also found another effect, located this time at the end of the sentence, which revealed a difference between the two relations during the phase of meaning integration. At this point, removing the connective had an impact on reading times but the effect was much larger for confirmation relations. This effect is therefore in line with Murray's findings, namely that discontinuous relations involve an increase in processing times when conveyed implicitly, but not continuous relations. It also shows that continuity should be thought of as a broad notion including temporal, causal but also narrative continuity.

6.3.2 The Role of Alternative Signals

Taken together, the studies summarized so far underline the role of discourse connectives for discourse processing compared to relations that are conveyed implicitly on the basis of the linguistic content of the segments and world-knowledge inferences. However, the explicit communication of discourse relations by means of a connective or its implicit communication do not exhaust the many ways in which a relation can be communicated. In fact, many of them can also be conveyed by the use of alternative signals (e.g., Das & Taboada, 2018; Hoek et al., 2019) such as lexical indications of contrast (18), of enumerations (19), or even for causal relations verb forms (20), and punctuation signs (21).

(18) Tom is a bit *overweight*; John is really *skinny*.

(19) Martha did *several* things to prepare for her holidays. She renewed her passport, went shopping and packed her suitcase.

(20) *Feeling* ready to leave, Martha was quite relaxed.

(21) Martha was quite relaxed: she was ready to leave.

[constructed examples]

Across three self-paced reading experiments, Crible and Pickering (2020) assessed the role of structural parallelism, in other words the repetition of the same argument structure across the two segments as in (22) as a cue to help readers process contrastive relations, and pitched the usefulness of this cue against that of a connective by comparing contrastive relations conveyed by the connective *but* to relations conveyed by the semantically underspecified connective *and* (see Spooren, 1997; Blackmore & Carston, 1999; Crible et al., 2019 for analyses of *and* in terms of underspecification).

(22) Nick always eats in low-budget restaurants but Grace always eats in fancy places.

[from Crible & Pickering, 2020: 869]

In the first two experiments, the task simply involved reading and answering simple verification questions, in other words questions that either contained or did not contain an element of the sentence that was just read, that readers had to evaluate as being 'true' or 'false'. They found that readers used parallelism as a cue for sentence processing both when it was conveyed by the adequate contrastive connective *but* and also when it was conveyed by the underspecified *and*. However, in a third experiment involving a harder comprehension question that followed the online reading, in which participants were asked to identify the discourse relation linking the segments, they found that the effect of parallelism was stronger when the relation was conveyed by the underspecified connective *and* compared to the contrastive connective *but*. They attributed this difference to the depth with which participants processed the sentences in both types of experiments. For the simple verification task, shallow processing was sufficient, whereas the more challenging task involving relation identification required deeper processing. This additional difficulty revealed that the use of alternative signals is all the more important when the connective itself does not fully serve to identify the intended relation. This study thus reveals the important role of alternative signals for discourse processing. This role will need to be further examined in future work for a broader array of relations and contexts before we fully understand the complex interplay of various signals for discourse processing.

6.3.3 Processing Connectives from the Written Mode

So far, the experiments we have reported involved connectives that are very frequently used in spoken language such as *because* and *but* in English. However, Indo-European languages possess a vast repertoire of connectives (see Chapter 3), among which a large number are mostly used in the written mode, such as *therefore* and *nevertheless* in English. Only a few studies so far have assessed the role of connectives from the written mode for discourse processing. In one such study, Crible, Wetzel and Zufferey (2021) tested the different roles played for the processing of contrastive relations by a connective frequently used to express contrast in spoken French (*par contre*) and a French connective bound to the written mode (*en revanche*) in a self-paced reading task. They found that reading the connective from the written mode immediately delayed processing compared to the frequent connective in the

spoken mode. However, this additional complexity was quickly resolved, as no differences in reading times remained by the time readers reached the end of the sentence.

In another study, Wetzel, Zufferey and Gygax (2022) tested the perception of incoherence in causal and concessive relations created by misusing either a frequent connective in spoken French (*donc* 'so' and *mais* 'but') compared to a connective bound to the written mode (*ainsi* 'therefore' and *néanmoins* 'nevertheless') expressing the same relations. Results indicate that readers react to the incoherence earlier when reading the second segment when a frequent connective from the spoken mode is used compared to a less frequent one from the written mode. Importantly, however, the effect of incoherence is also apparent with less frequent connectives from the written mode, which indicates that their meaning is integrated by readers, albeit more slowly.

Taken together, these two experiments provide some initial indication that adult native speakers also integrate the meaning of connectives from the written mode during discourse processing, even though their intuitions about their correct usage is vaguer (see Section 6.5). However, integrating their meaning takes more time, leading to delayed reactions compared to sentences conveyed by frequent connectives. The difference between frequent and infrequent connectives bound to the written mode can be compared to the ease of access to the mental lexicon between frequent and less frequent content words like nouns. In this case as well, less frequent words were found to take more time to be accessed across various types of experimental paradigms (e.g., McRae, Jared & Seidenberg, 1990).

6.3.4 Processing Ambiguous Connectives

In order to complete our survey of the role of connectives for discourse processing, we need to address one more important issue. So far, we have illustrated the usefulness of connectives for discourse processing by focusing mainly on cases where connectives conveyed one relevant meaning. However, many connectives are ambiguous and can convey different discourse relations depending on context (see Chapter 3). In some cases, these relations belong to different categories of relations in many taxonomies (see Chapter 2) such as temporality and causality, in the case of *since* in English. In other cases, the meanings are closely related, for example, concession and contrast, or even objective and subjective causality. In this section, we will discuss the way in which disambiguation takes place when readers have to process ambiguous connectives.

As part of a study aiming at assessing processing differences between causal and concessive relations, Wetzel, Zufferey and Gygax (2022) tested the way readers integrate the meaning conveyed by an ambiguous connective in a self-paced reading experiment. They tested the French connective *aussi* in causal relations ('in addition'), a connective that also has an additive meaning (similar to the English 'also'). For concessive relations, they tested the connective *or* (similar to 'however') that also has a background meaning (this meaning cannot be fully translated into English but it resembles some uses of 'in fact'). In both cases, they found that readers integrate these connectives at the same pace as monofunctional connectives from the written mode, indicating that ambiguity as such does not trigger interpretation problems for adult native speakers. The important difference seems to be whether connectives are frequently used, especially in the spoken mode, or whether they are bound to the written mode.

This experiment illustrates the quick disambiguation process that takes place when readers have to integrate the correct meaning of a connective that can be used to convey two very different relations. Asr and Demberg (2020) have tested the ways in which people understand and process relations of concession and contrast encoded by the two closely related connectives *but* and *although* in English. The authors stress that according to the theoretical accounts of discourse connectives that describe them as lexical items encoding a procedural instruction guiding interpretation (see Chapter 3), these connectives should be considered as ambiguous between the two relations and therefore interchangeable. Yet, they found that these connectives do not express the two relations with a similar frequency in corpus data. In fact, when used in the sentence-medial position, *although* conveys a relation of contrast and a relation of concession roughly equally frequently. In contrast, *but* conveys a contrast three times more frequently than a concession. In an offline sentence continuation task, they found that English speakers are sensitive to this difference, as they chose *but* to indicate contrastive relations in 75 percent of the sentences, and equally used *although* for concession and contrast. In addition, they judged contrastive sentences with *but* as more coherent than contrastive sentences with *although* in a coherence judgment task. In an online reading experiment using eye-tracking, they also read sentences with a less expected combination (i.e., concessive sentences with *but*) more slowly compared to the same sentences with *although*. The eye-movement data indicated that these less expected combinations triggered more regression to previously read regions compared to the more frequent combinations in corpus data. All

together, these results indicate that readers have probabilistic expectations about the meaning conveyed by a connective depending on the frequency with which a connective expresses a given relation in language use. This means that native speakers do not only integrate the meaning of connectives but also their most frequent patterns of use, and use both types of information during discourse processing.

In their experiments, Asr and Demberg (2020) also compared the same connective *although* when used in two different syntactic positions: sentence initial and sentence medial. This comparison was based on the observation that this connective typically conveys distinct relations in the two positions in corpus data. While *although* is equally used for concession and contrast relations in the sentence-medial position, in the sentence-initial position this connective overwhelmingly conveys a concessive relation. Again, the authors found that readers are sensitive to this syntactic difference, and adjust their expectations about the meaning of *although* depending on its syntactic placement. This result thus provides some empirical validation of the intricate interrelations between connectives' meaning and syntax discussed in Chapter 4. Similar research is still needed to determine whether this sensitivity can also be observed for other connectives and discourse contexts, as research is still scarce on that topic.

6.4 THE ROLE OF CONNECTIVES FOR RECALL AND COMPREHENSION

While the role of connectives for speeding up discourse processing has been consistently found across studies and is now firmly established, their impact on the way readers recall and understand information in a discourse is much more debated. While some early studies found an effect on connectives for recall, these findings have not always been replicated and several methodological limitations have been evidenced. For example, Meyer, Brandt and Bluth (1980) found an effect of connectives on recall, as readers who had read the explicit version of a text containing connectives most often reproduced the original text structure in a recall task. In contrast, Meyer (1975) found no effect of either connectives or other signaling devices (e.g., expressions such as 'more importantly') on either immediate or delayed recall. Later on, Millis and Just (1994) did find an effect of connectives on both recall and comprehension. In their experiment, participants had to read segments that were either linked by a connective or not, and had to judge whether a probe word had been included in the segments they had just read. They found that recall was faster when the sentences were linked

by a connective, and that readers were also faster and responded more accurately to comprehension questions about the segments they had just read. Caron, Micko and Thüring (1988) found that semantically unconnected sentences linked by the causal connective *because* led to better recall than the same sentences simply connected by the additive connective *and*. It is not certain, however, that this result would generalize to other discourse relations, as Caron et al. (1988) found that people remember with more difficulty sentences that contain the concessive connective *but* compared to the causal connective *because*.

The effect of connectives on recall could not be clearly established either in the study by Sanders and Noordman (2000). The authors measured both immediate recall in the form of statements that participants had to verify (i.e., decide if the statement was correct or incorrect based on what they had just read) and delayed recall based on free recall tasks in which people were instructed to mention as many elements as they could remember reading. Results indicated a very marginal effect of connectives on immediate recall, as reaction times were faster for explicit relations, but only in the case of problem-solution relations. In addition, there was no effect of explicit marking on accuracy in this task. Similarly, the explicit marking of coherence relations by connectives did not have any effect on free recall. They concluded that connectives play a role for the online processing of coherence relations but do not play a lasting role on the mental model of discourse once it has been constructed.

Yet, even though their importance for recall is not attested, it does not mean that connectives do not help readers understand the content of a discourse. In fact, Degand, Lefèvre and Bestgen (1999) found evidence of the opposite. In an offline comprehension task, they assessed the role of causal connectives for discourse comprehension by asking participants to give short answers to questions targeting both the parts of the texts in which connectives were used in the explicit version and removed in the implicit version, and questions targeting their understanding of other elements of the text that were not part of a causal relation. Results from this experiment clearly showed that connectives are useful for text comprehension, as participants reached higher scores after reading the explicit than the implicit version of the texts. In addition, they also reached higher scores for questions not targeting the causal relations after reading the explicit version, which indicates that the presence of connectives increased their overall comprehension of textual content, and helped them form a coherent mental model. These findings were moreover replicated in a similar experiment aiming at comparing the role of connectives in

participants' native language and in a foreign language (Degand & Sanders, 2002). It seems therefore that connectives, and causal connectives in particular, can be useful for comprehension, at least in order to perform some tasks going beyond mere recall, and tapping on discourse comprehension. We will come back to these issues in Chapters 8 and 9, when discussing the importance of connectives for young readers and second language learners.

Going one step beyond comprehension, speakers' integration of the meaning conveyed by specific connectives also has an impact on their evaluation of the validity of an argument. In a series of experiments, Schumann, Zufferey and Oswald (2020) assessed the different roles of French causal connectives that can all be used to convey an attributive meaning for the acceptability of discourse segments containing a fallacious argument, such as *since*, *as*, and *given that* in English. In these experiments, participants read short dialogues like (23), and then had to evaluate the acceptability of Alexander's reply based on different factors.

(23) Barbara: Il est crucial de mieux soutenir les jeunes parents parce qu'avoir un enfant signifie beaucoup de charges financières.
Alexandre: Augmentons les allocations familiales, puisqu'on ne pense qu'à l'argent.
'Barbara: It is crucial to better support young parents because having a child means having a lot of financial responsibilities.
Alexandre: Let's raise the family allowance, since it's only about the money.'
[from Schumann, Zufferey & Oswald, 2021: 5]

In part of the dialogues, the answer contained a fallacious argument in the form of a straw man fallacy, as in (22), in other words a distortion of the opponent's argument to make it appear more extreme and therefore less acceptable (e.g., van Eemeren & Grootendorst, 1992). The authors manipulated the connective used to introduce the argument by alternating between four different French connectives (*puisque, comme, étant donné que, vu que*). Results indicated that participants reacted differently to these connectives, even though all of them convey a causal meaning and indicate an attribution of content to another source. However, only the connective *puisque* used in this experiment is often used in ironic contexts to express a caricature or an exaggeration (Franken, 1996; Zufferey, 2014). Participants reacted strongly to this cue, as *puisque* acted as a potent forewarner, alerting readers to the distorted nature of the argument. This experiment thus demonstrates that readers specifically integrate the fine-grained

nuances of connectives' meaning and use this information when processing, understanding, and evaluating the content of a discourse.

6.5 INDIVIDUAL DIFFERENCES BETWEEN READERS

So far, we have discussed the processing of discourse relations and connectives by adult native speakers, thereby implicitly treating them as a homogeneous group, who should not vary in their comprehension, recall and reading of connectives and discourse relations. This assumption of homogeneity is quite widespread in experimental studies in both linguistics and psychology involving between-subject designs, such as studies comparing differences between native vs. non-native speakers, teenagers with a high or low academic background, etc. In these studies, variations within groups are often discarded, because they are perceived as noise or even measurement errors. This reliance on group comparison has led researchers to overlook the important individual differences that exist in the way even adult native speakers process and understand their mother tongue. Yet, such differences have now been clearly established for many aspects of lexical and syntactic processing and comprehension (see Kidd, Donnelly & Christiansen, 2018 for a review). Given that connectives are at the interface between the lexicon, syntax and discourse, there are reasons to believe that individual variations in the way people process and understand them should also exist. Empirical evidence to this claim remains, however, few and far between.

Some evidence of individual variations was already hinted at in the early studies of discourse processing, as a way to explain some contradictory findings between them. For example, Meyer, Young and Barlett (1989) state that not all readers are sensitive to differences between discourse relations. More specifically, readers with average verbal ability do not seem to be sensitive to these differences. The effect of verbal ability was also mentioned by Rickards et al. (1997) who did not find an effect of signaling (i.e., connectives) for high-ability readers, which again points to the existence of variabilities among speakers depending on their level of linguistic competence.

Another source of variability was found in studies who compared people with high prior knowledge of the topic of discourse with people with low prior knowledge (McNamara & Kintsch, 1996; Kamalski, Sanders & Lentz, 2008). These studies found that connectives benefit more to readers with low prior knowledge compared to readers with high prior knowledge. Kamalski et al. (2008) even found a negative

effect of connectives for readers with high prior knowledge in their understanding of informative texts. The authors explain this effect by the deeper processing required to understand a more implicit discourse, which in turn increased participants' level of understanding.

Yet, these early studies did not explore in more detail the various speaker characteristics that lie at the source of these individual variations, beyond the role of verbal ability and prior knowledge. This topic has been dealt with again in two recent studies, underlying the link between people's sensitivity to connectives and other discourse signals and their degree of exposure to print.

Zufferey and Gygax (2020a) focused on the ability of adult native French speakers to evaluate the correct and incorrect uses of connectives from the written mode, by comparing the roles of cognitive complexity and connective frequency in corpus data in adults' ability to use them. They focused on four French connectives, each conveying a different discourse relation, with a variable degree of cognitive complexity: an additive connective [*en outre*], a consequence connective [*aussi*], a causal connective [*en effet*] and a concessive connective [*toutefois*]. Two of these connectives were highly frequent in written language as measured in corpus data (*en effet* and *toutefois*) and the other two were less frequent (*en outre* and *aussi*). Results demonstrated that *en outre* and *aussi*, the two less frequent connectives, yielded a lower performance compared to *en effet* and *toutefois*. It seems therefore that what causes difficulties in the mastery of connectives from the written mode for adult speakers is their low frequency.

The degree of exposure to the written mode that people have is also quite variable. The authors therefore expected that the ability to handle connectives from the written mode should be linked to the degree of exposure to print that people have. This was indeed the case, but importantly, individual variations were significant only for the two less frequent connectives (*en outre* and *aussi*). This means that all adult native speakers master the frequent connectives, but greater individual differences are found for less frequent connectives. For them, a greater exposure to print is necessary, and not all adults have had it, even in their native language.

Scholman, Demberg and Sanders (2020) also assessed individual variations among adult native speakers, but this time on their ability to use alternative signals indicating list relations. In their experiment, participants had to complete short discourses of two sentences with one or two sentences. Crucially, half of the stimuli contained an indication that a list was expected to follow while the other half did not. These indicators were the expressions *a few*, *a couple*, *multiple* and *several*.

If participants are sensitive to these cues, it is expected that they should produce more continuations corresponding to a list relation than when sentences do not contain a signal. This was indeed the case, as the proportion of list relations was significantly higher when a signal was present. It is interesting to note, however, that in both conditions, the most frequent continuation was a causal relation. This result thus provides some further confirmation for the importance of causality in discourse. More importantly in the context of individual variations, the authors also found that participants who had a higher degree of exposure to print also provided the most list continuations. This study thus confirms the importance of print exposure to develop a sensitivity to the marking of discourse relations, be it with connectives or alternative signals.

In a nutshell, the studies reviewed in this section indicate that the analysis of individual differences represents an important line of enquiry to assess the way people produce, understand and interpret discourse relations and connectives. We will discuss the notion of individual variations again in Chapter 8 in the context of language acquisition.

6.6 SIMILARITIES AND DIFFERENCES OF PROCESSING ACROSS LANGUAGES

The processing studies that we have reviewed in this chapter have been conducted on a handful of different languages, mostly Indo-European (English, Dutch, German, French) but also in Chinese. Even though comparisons are limited to these specific cases, it is noticeable that most of the results seem to reflect great convergence between the languages in the way readers process and understand discourse relations. The difference of cognitive complexity between causal and concessive relations is a case in point, as was underlined in studies conducted in English (Murray, 1995; Morera et al., 2017), German (Köhne & Demberg, 2013) and Chinese (Xu et al., 2018). Despite these global similarities, fine-grained differences can also be observed in the way readers process and understand discourse relations across languages, even when they are typologically closely related. We review these differences in this section.

Blumenthal-Dramé (2021) conducted a fine-grained comparison of the way readers process causal and concessive relations when marked with an appropriate connective or by the underspecified connective *and* in English and *aber* in German. The rationale for this comparison is that English is a less synthetic language compared to German, with

little inflectional morphology and a tendency to have invariable word forms that are highly general both grammatically and semantically. As a result, Hawkins (2019) argued that German speakers, who speak a more synthetic language with richer morphology, should allocate more resources to the treatment of individual words compared to English speakers, who should focus more on units beyond words. Blumenthal-Dramé (2021) therefore hypothesized that this difference between the two languages should be reflected in differences in the ways readers from both languages process discourse relations. More specifically, it is expected that German speakers will pay more attention to connectives and will therefore benefit more from their presence. This hypothesis was tested by means of very similar self-paced reading experiments conducted in the two languages with adult native speakers. Results first provide some further confirmation of the causality-by-default hypothesis (Sanders, 2005) discussed above, to the extent that readers globally benefited more from the presence of concessive than causal connectives. An important difference was however also observed between the languages, as causal connectives provided a significant processing advantage over implicit relations for German readers but not for English readers. Besides, the use of connectives overall provided a greater advantage in German than in English, for both relation types. However, in another set of experiments comparing English and German causal and concessive connectives across eye-tracking and ERP experiments, Köhne-Fuetterer et al. (2013) did not find any differences between the two languages. It seems therefore that the exact nature and extent of these differences may depend on the connectives included in the experiment and the experimental paradigms used.

Another difference between the loose form–function mappings of English and the more precise connectives found in other Indo-European languages has also been discussed in the context of causal relations, and more specifically in the difference that separates objective from subjective causality. We already illustrated this difference in (6) and (7), repeated in (24) and (25) for convenience.

(24) Heidi was proud and happy, because she won first prize at the art show.

(25) Heidi could imagine and create things, because she won first prize at the art show.

[from Traxler, Bybee & Pickering, 1997: 485]

These sentences come from the online processing experiment conducted in English by Traxler et al. (1997). As these sentences illustrate,

in English both types of causal relations are typically conveyed by the same connective (*because*). However, in other languages like Dutch, specific connectives are used in both cases (see also Chapter 7). Canestrelli, Mak and Sanders (2013) replicated the English experiment in Dutch, using two different connectives: *omdat* for objective relations and *want* for subjective relations. This experiment revealed important differences between the two languages. While Dutch-speaking readers also slowed down when processing subjective causal relations compared to objective ones, they did so much earlier than in English. Recall that in Traxler et al. (1997), readers slowed down at the words 'first prize' in the second segment. In Dutch, readers slowed down earlier, at the words immediately following the connective *want*. This gap can be explained by the difference of marking between the two languages. In English, the connective *because* does not provide information about the type of causal relation that will follow. The readers have to infer it based on the linguistic content of the segments. As a result, they slow down as soon as an objective interpretation could be ruled out, at the words 'first prize'. In Dutch however, the connective provides an indication of the intended type of causal relation. The connective thus provokes an immediate slow down, but later on in the sentence, readers do not slow down again at the words 'first prize' because they already have a subjective interpretation in mind.

French represents an intermediate case between English and Dutch, as two specific connectives exist to express objective and subjective causality, but the subjective connective *car* is used only in writing, as in speech the objective causal connective *parce que* is used for both types of causal relations (Simon & Degand, 2007).[1] This situation therefore raises the question of whether French speakers should behave more like English speakers, due to the situation in spoken French, or like Dutch speakers, even though their subjective causal connective is restricted to the written mode. In a reading experiment, Zufferey et al. (2018) compared the reading of objective and subjective causal relations with *parce que* and *car* using a translated version of the sentences from Canestrelli, Mak and Sanders (2013). Results from this experiment indicate that French readers share some of their behavior with Dutch readers, and some with English readers. Like Dutch readers, they slow down at the words immediately following *car*, but they also slow down again later on in the sentence like English readers.

[1] Note that in more recent studies a rise of the connective *car* is observed in more informal "spoken-like" registers like texting in which the connective is used in a more objective way (Blochowiak, Grisot & Degand, 2020).

This indicates that the information given by the connective did not allow them to infer subjectivity. The authors argue that the early slow down found at the words immediately following *car* in both objective and subjective contexts can be assimilated to a register effect. Indeed, due to its boundedness to the written mode, *car* is assimilated to a formal register by many French speakers, and it therefore represented an unexpected lexical choice in the informal register used in the experimental items. According to this interpretation, the slow-down observed after *car* reflects its unexpectedness rather than a link with subjectivity.

In sum, the studies summarized in this section testify to the fact that cross-linguistic differences exist even between closely related languages. These differences will need to be assessed further between other language pairs and discourse relations before firm conclusions can be reached regarding their magnitude, and their implications for discourse processing and understanding.

6.7 SUMMARY

This chapter started with an overview of the way discourse relations are processed, recalled and understood. One of the major findings is that not all relations are processed and remembered in the same way. It seems that causal relations play a special role for creating coherence in discourse, as they are processed more quickly and remembered better. Conversely, because they are highly expected, causal relations benefit less from the presence of connectives compared to discontinuous relations like concession and confirmation. While connectives play an immediate role for discourse processing, readers also integrate the information provided by alternative signals such as syntactic structures indicating parallelism, and content words like adjectives (e.g., to infer a lexical contrast) and adverbs (e.g., *zwar* in German). In their native language, speakers are able to take advantage of all sorts of connectives for discourse processing, even those restricted to the written mode, and those that are ambiguous. We will see that the situation is somewhat different for children and learners in Chapters 8 and 9. However, variability also exists among adult native speakers, who should not be treated by default as a highly competent and homogeneous group. Finally, we saw a glimpse of the way crosslinguistic differences in the signaling of relations by connectives impacts their online processing. We will dig deeper into these crosslinguistic differences in Chapter 7.

DISCUSSION POINTS

- What are the main conclusions from studies that have compared the processing of different coherence relations?
- What are the functions of connectives during online processing?
- Can you think of experimental designs that could be used to analyze the processing of discourse relations other than the ones illustrated in this chapter, for example, temporal relations? (See also the way temporal relations were tested with children in the next chapter.)

FURTHER READING

Sanders & Noordman (2000) as well as Murray (1997) remain excellent examples of processing experiments that should be read to get familiarized with the methodology behind online processing for the comparison of discourse relations and connectives. Scholman, Rohde and Demberg (2017) provides an in-depth discussion of the role of connectives in creating expectations about discourse structure. Asr and Demberg (2020) show how corpus data can be combined with experiments to provide a detailed account of the factors affecting connectives' use and processing. Mak, Tribushinina and Andreiushina (2013) illustrate the way processing studies can be used for crosslinguistic comparisons.