How Default is Causality-by-Default?

Marja H. Oudega

Student number 3441636

M.H.Oudega@students.uu.nl

MA Linguistics Thesis

Supervisor: prof dr. T.J.M. Sanders

Cosupervisor: dr. R.W.F. Nouwen

Utrecht Institute of Linguistics OTS

Utrecht University

August 2011

Abstract

Causal coherence relations seem to have special cognitive status in discourse processing. Even though they take less time to process than other types of coherence relations, recall and representation are better. The Causality-by-Default hypothesis explains this paradox by stating that readers maintain a causal coherence relation between two segments as a default assumption, and only reconsider the coherence relation when causal coherence turns out to be impossible. In this thesis I investigate the viability of this default assumption of causal coherence as a cognitive mechanism, and look into how specific characteristics of the segments and the surrounding discourse influence the assumption and the processing of causal coherence. The results of an eye tracking experiment support to the Causality-by-Default hypothesis, although coherence processing is affected by the characteristics in the segments, resulting in different processing patterns.

Acknowledgements

Writing this thesis has been one of the most stressful, demanding, but also satisfying tasks I have done. And even though it is hard to be entirely content with something that has turned out to be so personal, I have enjoyed it very much.

I would like to thank everyone that made it possible for this thesis to happen. Firstly my supervisor Ted Sanders with whom I have worked together over the last year, and who kept having faith that my ideas would turn out into something more substantial than merely ideas. I am very grateful for all his enthusiasm and support. I also would like to thank my second supervisor, Rick Nouwen, for asking the right questions and helping me to stay critical. Many thanks to Arnout Koornneef for all his help and advice on the experiment, the items and the statistical work.

In addition, I would like to thank all my class mates and friends whose company has made the last couple of years go by much too quickly. A very special thanks to Matteo Cavaliere, Saskia de Haan and Maartje Schulpen for their invaluable friendship. My family, without whom I could not have started, continued or finished studying, I wish to thank for their love and support. And last but not least, a big hug to my boyfriend Nol van Meegeren, for always being supportive of me, for making me dinner after a long day, and for not running away screaming on the many occasions where stress made me insufferable.

1. Introduction

To be able to understand a discourse, a reader must construct a coherent representation of the discourse segments and the relations that connect these segments. Examples of such coherence relations are additive, temporal and causal relations. Sometimes this coherence is linguistically marked by a connective or a signaling phrase, but this is often not the case. Readers are usually very good at inferring unmarked coherence relations from the context, even when these are more complex relations such as causality.

When presented as a short discourse such as in example (1) below, the content of the clauses allow the reader to make inferences about likely or necessary other facts that are used as (causal) links between the segments, in order to make the discourse coherent.

1. Jan kwam binnen. Max ging weg. (Sanders 2005)
   John entered the room. Max left

When coherence is established for the example in (1), a fact that is likely to be inferred is for example that Max does not particularly like John. This one fact links the first sentence in (1) logically to the second sentence. A reader makes this type of inferences that act as links between discourse segments because a reader needs to assume that there is coherence between the segments, in order to be able to construct a cognitive representation of the discourse (Hobbs 1979; Sanders, Spooren & Noordman 1992; Kehler 2002). This assumption of coherence causes world knowledge and situational knowledge to play a crucial role in constructing a representation of a discourse.

Causal coherence relations can be seen as cognitively more complex than other coherence relations, since they entail temporal and additive relations: causal ordering is temporally strict in that a cause must occur before the start of the consequence, and when an implication can be deduced, a conjunction can be deduced also. Interestingly it has been found in reading experiments that causally related discourse is read faster than discourse that is not causally related (Haberlandt and Bingham 1978; Sanders and Noordman 2000; Mulder 2008), and that reading time decreases when causality increases (Keenan et al. 1984;

Myers, Shinjo and Duffy 1987; Wolfe et al. 2005). At the same time, causally related discourse is argued to lead to better recall and representation (Trabasso & van den Broek 1985; Noordman & Vonk 1998; Sanders & Noordman 2000). For example, Sanders & Noordman (2000) observed that causal relations lead to faster processing of the connected information, but also lead to better and faster verification of statements and better recall of information; both findings can be seen as evidence for a more integrated representation.
These seemingly contradicting facts, higher cognitive complexity yet shorter reading times and better representation, gives rise to what Sanders (2005) calls 'the paradox of causal complexity'. The paradox lies in the idea that more complex cognitive processes should take more processing effort (and therefore result in longer reading times), and that shorter reading times are unlikely to lead to better representation and recall, because less effort is spent on processing. According to Sanders (2005) the explanation of this 'paradox' can be found in the special cognitive status of causality and that of coherence in discourse processing. The idea that causality is an omnipresent force in the world around us, and therefore in how we interpret the world, is not new. In different domains, such as the visual paradigm, more evidence of our want to infer causal relations can be found (see for example Michotte 1963). Similarly, coherence is not merely a property of a text, but of how the reader puts together the information in a text while constructing a cognitive representation (Sanders et al. 1993). These two cognitive processes, inferring causality and assuming coherence, lead to the Causality-by-Default hypothesis, proposed by Sanders (2005).

The Causality-by-Default (CbD) hypothesis states: "Because readers aim at building the most informative representation, they start out assuming the relation between two consecutive sentences is a causal relation (given certain characteristics of two discourse segments)." (Sanders 2005:9). This hypothesis suggests that a readers' default assumption about coherence in a discourse is one of causal coherence, and that temporal or additive coherence is only inferred when a causal relation is not possible. The CbD hypothesis explains the shorter reading times for causally related discourse when compared to temporal or additively related discourse from a cognitive point of view: Causality is expected, and only when this is not the case in the discourse, other possibilities are investigated. The reason why readers start out building the more complex representation

might be found in the cognitive status of causality itself. Since causal coherence is the most informative relation (containing a strict temporal ordering since the start of a cause must precede the effect), relegation to a less informative structure such as temporal or additive coherence can be expected to take less processing effort than the other way around, where additional information must be added and fit into the already constructed representation. An explanation for better representation and recall from causally related discourse could be related to this same level of informativeness and complexity: causality is highly structured and the segments are linked not only based on the text itself, but also through inferences made by the reader.

The CbD hypothesis provides an apparent simple cognitive solution to the paradox by stating a general assumption of causal coherence in discourse processing. But there are several issues that need to be addressed. Firstly, it is not the case that readers always infer causal coherence in a discourse. The CbD hypothesis does take this into account by the added condition: "given certain characteristics of two discourse segments". This indicates that there are possible characteristics of the discourse that act as restrictions on CbD and either cause the reader not to assume causal coherence, or drop the assumption of causal coherence. Although this condition is placed between parentheses, it raises questions that are vital to the validation of the hypothesis: Empirical research shows the existence of processing and recall differences due to the degree of causal relatedness between segments (Keenan et al. 1984; Myers et al. 1987; Wolfe et al. 2005). In what way is the CbD hypothesis affected by the strength of the causal coherence relation? Secondly, discourse segments are almost always embedded in a larger context. This context seems a likely influence on processing. Is CbD affected by factors such as the context of the two segments? And if so, in what way does context affect the assumption of causality?

Thirdly, so be able to experimentally test these questions presented above, it needs to be possible to turn the hypothesis into a processing theory and to test empirically; specific predictions need to be made about how and when CbD is affected by the factors mentioned above. When during the reading process do readers assume causal coherence? At what point do the characteristics influence the assumption of coherence? And at what point is the assumption dropped when causal coherence is not a possibility?

The issues above can be summarized into three main questions: How does the causal relatedness between the segments affect the CbD hypothesis? How does the context around the segments affect the CbD hypothesis? And can the CbD hypothesis be made more specific when it comes to the moment in processing where causal coherence is assumed or dropped? In this thesis I will attend to these questions which the CbD hypothesis brings about by trying to separate the influences from the text based characteristics of the segments and the preceding discourse on coherence processing from possible default cognitive assumptions in an online reading experiment.

I will discuss in further detail whether there is a theoretical basis fro the assumption of causal coherence, where I focus on the Segmented Discourse Representation Theory (Asher & Lascarides 2003). I will also discuss experimental processing results that are associated with the two text based sources of information that are relevant for this thesis, context and degree of causal relatedness. Lastly I will look into the possibilities for turning the CbD hypothesis into a processing theory and make corresponding predictions with respect to inferring and processing causal coherence in discourse. In the sections that follow this overview, I will provide a detailed explanation for the specific design choices that I made and present the research questions. But first I want to look more closely at the CbD hypothesis and try to find theoretical support for the existence of an initial assumption of causal coherence in discourse processing.

2. Causality-by-default: Theoretical background

The idea of causality as a preferred interpretation is not new. For example Asher and Lascarides (1998) propose *the principle of Maximize discourse coherence*, which states that the informational-richer causal relation (i.e. *explanation* or *result*) is preferred over non-causal relations. "Intuitively, one prefers an interpretation of a discourse that offers explanations of intentional behaviour […] to an interpretation of the discourse where such behaviour is left unexplained […]. We can model this via the partial order of rhetorical relations: Explanation > Background in this case." (Asher & Lascarides 1998:107). In other words, when a causal relation is possible, this relation is preferred over a relation where the information is interpreted as merely background information. But how does a reader know to connect two events in such a way that makes a causal coherence relation possible?

Asher and Lascarides' Segmented Discourse Representation Theory (SDRT, Asher and Lascarides 2003) integrates pragmatics and semantics into a theory of discourse interpretation and representation. A discourse relation is established on the basis of the interaction between World knowledge and Lexical knowledge. According to SDRT a reader has put together a certain amount of assumptions about discourse in general, a Knowledge Base that contains at least the expectation that a text is coherent and that the propositions are true. Besides this, the Knowledge Base contains World knowledge (knowledge about situations and experiences) and Lexical knowledge (knowledge about semantic content), and all indefeasible knowledge such as the order of cause and effect and logical laws. When no overt marker of the coherence relation is present in a discourse, the speaker infers the relation based on these knowledge resources.

Asher and Lascarides state that the default discourse relation interpretation can be described by the axiom of Narration, where "(…) by default, the order in which eventualities are described matches their temporal order in interpretation." (Asher and Lascarides 2003: 200). An example of a *Narration* interpretation is given in (2a) where the third clause contains a cue phrase *then* that allows the reader to infer a similar temporal order between the first two clauses. Additionally, Asher & Lascarides state that "(…) we assume that the 'extra' information one needs for inferring *Narration* is that the first mentioned event e_a 'occasions' the second mentioned event e_b. That is, there's a plan or a 'natural event-

sequence' such that events of the sort described by *a* lead to events of the sort described by *b*." In (2a) it shows that this 'extra' information that is needed to infer *Narration* is the cue phrase *then*. In the case of example (2b) no occasion relation exists between the two clauses; there no natural event sequence between the two study-events. This results in an interpretation where no *Narration* is inferred, and no temporal ordering is applied.

2 a Mary watched TV. She studied. Then she went out.

b Kim studied. Sandy studied too. (Asher & Lascarides 2003)

It is clear that establishing the event order (narrative or not), and therefore inferring this 'natural event-sequence' is crucial to inferring causal coherence. But how does a reader infer this natural event-sequence? According to Asher and Lascarides, script-like knowledge about specific situations helps the reader to connect two events. The lexical knowledge then allows the reader to infer *Narration*, as shown in (3a), where the relation between *push* and *fall* makes it possible to connect the two events and make forward inference. Here the narrative order matches the event-sequence order, since our knowledge about the relation between *push* and *fall* is that *push* is a verb that describes a force that can cause movement to a patient, and that *fall* is a verb that describes movement of its 'subject' (Asher and Lascarides 1997) and therefore the former happens before the latter.

3 a Max pushed John. He fell.

b John fell. Max pushed him.

In (3b), the scriptal knowledge lets the reader infer an event-sequence that does not correspond with the narrative order. Narrative order based in the lexical knowledge is then overruled by the scriptal information, based in world knowledge. The examples in (3a) and (3b) show that the knowledge about specific situations is vital for inferring coherence, and can overrule default narrative order. When the event-sequence is established, a causal interpretation is inferred based on the principle of *Maximize discourse coherence*. This 'preference for causal coherence' combined with Asher and Lascarides' SDRT theory for

discourse interpretation can explain why, without the help of overt marking or correct event order, a causal interpretation can be made.

Both Maximize discourse coherence, as well as the function of world knowledge in inferring event-sequence provides a theoretical basis for the CbD hypothesis. Additionally, SDRT motivates that a reader uses all the information available (lexical, and world knowledge) to be able to construct a causal coherence relation. Only when both lexical and world knowledge come up short, a different coherence relation will be inferred. This conforms to the explanation that the CbD hypothesis gives for the reading time differences between causal and other coherence relations.

3 Text based influences on discourse processing

As discussed in the introduction, a theory for processing (causal) discourse relations must somehow take into account text-based factors. Even though it is likely that coherence resolution is guided by cognitive principles, causal links are inferred based on the textual content, both from the two clauses that are to be related, and from the structure and the expectations about the structure of the preceding discourse.

3.1 Context and Causality-by-Default

Two segments that together form a coherent text are very often embedded in a larger discourse. The mental representation of the two segments and the coherence relation that connects them is necessarily a component of the overall representation of the discourse. Preceding discourse can for example set up a background for inferencing coherence, as the causal links in example (4). Sometimes, it is the background information that makes it possible to infer a narrative link and construct a coherent representation that might not have been possible otherwise, as in (5a) and (5b) taken from Lascarides and Asher (1993). Here the third sentence in (5b) introduces a topic that links the two preceding sentences and creates the coherence. In (5a) this topic is absent and coherence is lacking.

4 Max disliked John very much, even though they had been friends in childhood. At the birthday of his sister, Max saw John arrive at the door.
John entered the room. Max left.

5 a ? My car broke down. The sun set.

b My car broke down. The sun set.
Then I knew I was in trouble. (Lascarides and Asher 1993)

The reader uses the knowledge he has about similar situations to create a coherent representation. Together, a car breaking down and the sun setting are sufficient premises to conclude that one is in trouble. But what causes the apparent lack of coherence when the

background topic in (5b) does not follow the discourse, as in (5a)? It does not seem difficult to conclude that the situation described in (5a) is problematic for the author, but still it is very hard for a reader to infer a common topic in order to create coherence.

In my opinion this could be explained by the content of the first segment: *My car broke down* introduces a Subject of Consciousness (the author) that is faced with a problem, a broken down car. Based on Asher and Lascarides' *Maximize discourse coherence*, the reader would prefer a causal coherence relation between this first segment and the next. For example a solution such as *I called the road patrol*, or a cause such as *The engine jammed*. The next sentence is the example, *The sun set*, is not a good candidate for causal coherence, and coherence breaks down. But why does coherence break down, when clearly additive or temporal coherence are both possible? I assume that an explanation for this lack of coherence needs to be found in the informativeness of the discourse segment. A reader has a goal when reading a discourse. Minimally, the goal is to understand the information presented in the discourse. In order to that, readers need to construct a coherent representation, and therefore attempt to find explanations for why certain actions, events and states are mentioned in a text (Graesser et al. 1994, Asher & Lascarides 1998). In the discourse that is given in (5a), even when the segments are temporally or additively related, the explanation for why the segments are relevant is not clear without the addition of a third (topic) sentence as in (5b). This can be illustrated by adding an overt temporal marker to the discourse, as in *My car broke down and then the sun set.* Intuitively, even a causal relation, although yielding a very unlikely situation in this example, could be preferable over the uninformative temporal one.

The example by Lascarides and Asher in (5b) can be expected to behave in a similar way when the topic is introduced preceding the target sentences, as in (6)[1].

6 I was in trouble. My car had broken down. The sun was setting.

Here the topic is presented as a statement, and the following sentences are connected additively to each other, but together they present support for the statement, and an explanation for why the author would make this statement.

[1] Because the temporal order is no longer equal to the order in the discourse, tense marking is necessary to ensure an interpretation similar to (5b).

A topic does not always need to be explicitly present in the discourse. Asher and Lascarides (1993) describe that *Narration* between *two segments* only holds if there is a distinct, common (and perhaps implicit) topic. This rule is used to explain the incoherence of (5a): as long as World Knowledge about cars breaking down and the sun setting is represented as expected, one cannot find a common distinct topic, and so *Narration* between the clauses cannot be inferred.

With respect to the CbD hypothesis, the above discussion and example (5a) illustrate that even though a causal structure, being the most informative structure, might be preferred, there is a limit to the additional information that a reader is willing to infer in order to construct a causal relation. Example (5a) shows that when not enough information is given, the assumption of causal coherence can be dropped.

We have seen that a context can provide cues that aid the reader in inferring a causal link, even when the link itself is not specifically mentioned in the context. There are some questions that can be asked about the interaction between context and CbD in coherence processing. In what way does the presence of a context affect the assumption of causal coherence? Can a preceding context always assure that the assumption of causal coherence is not dropped, even when without context no causal coherence is inferred? Can a context also prevent the assumption of causal coherence to be in place? And when it comes to testing these effects in processing, do the effects of context show up as processing differences? An answer to these questions would shed more light on the assumption of causal coherence.

3.2 Causal relatedness and Causality-by-Default

Even though a context can play a part in creating a causal link, the segments themselves have to make the link possible. In what way does the content of the segments, and the causal strength that is present in the segments, affect the assumption of causal coherence? First I will return to SDRT as a possible theoretical basis for the existence of a degree in causal strength or causal relatedness, and then I will present some experiments and results that focussed on the effect of causal relatedness on processing discourse.

To recap, with SDRT Asher and Lascarides (2003) present the opportunity to construct an event-sequence between two discourse segments as being a fundamental step towards interpreting a causal relation. This event-sequence is constructed based on knowledge about specific situations. In the example given in (3a) and (3b) (here repeated as (7a) and (7b)) the knowledge about the relation between *push* and *fall* creates a strictly ordered event-sequence and overrules narrative order in favour of a backward *explanation* relation (consequence- cause). An important question is: Is this knowledge that a reader has about situations always so strong that an event order can be deduced? Or, to turn the question around: Does it always require such a strong cause-effect relation such as between *push* and *fall* to deduce event order (and overrule the narrative order)?

Asher and Lascarides' answer to this can be interpreted as there being two levels at which situational knowledge overrules narrative order: On the one hand there is what they call the *push*-causal law, which is applicable to example (7a) and (7b). On the other hand there are causal relations based on weaker causal or situational knowledge such as in (7c) and (7d).

7 a Max pushed John. He fell.

b John fell. Max pushed him.

c John fell. Max helped him up.

d ? Max helped John up. He fell.

e Max stood up. John greeted him.

f John greeted Max. He stood up.

(Adapted from Asher and Lascarides 2003)

Asher and Lascarides argue that in the case of example (7a) based on the first segment, world knowledge about possible outcomes of the event is entailed. The *push*-causal law incorporates necessary outcomes of *push* since these outcomes are so strongly related to the cause (Asher & Lascarides 2003). A causal law governs the relation, and cause necessarily precedes affect, so that is (7b) narrative order is overrules.

This differs from the examples in (7c) and (7d). In (7c) through forward inferencing, a cause-consequence interpretation is available, but the example in (7d) shows that it is harder to overrule narrative order than in the (7b) example. When tense marking would be added to

the second sentence, as in *He had fallen*, the temporal order would be clear, and therefore the causal coherence would be clear: an explanation relation (consequence-cause). Yet without tense marking, the temporal order is unclear, and the interpretation of the discourse (who falls, and does this happen before or after John is helped up) is at least ambiguous.

Asher and Lascarides name the type of knowledge that helps to infer a causal link in (7c) and (7d) as 'scriptal' (indicating a weaker causal relation). In (7c) the knowledge about the event in the 'cause' segment does not entail information about a consequence since only possible but not necessary outcomes are incorporated. Instead processing the events in both clauses (fall, and help up) is necessary before the event-sequence can be inferred.

The last two examples (7e) and (7f) show that situational knowledge can lead to the interpretation of an event-sequence, but is not ordered with a strict cause-effect ordering. Both *greeting* can cause someone to stand up, as well as *standing up* can cause someone to be greeted. According to Asher and Lascarides, narrative order is kept for both examples, leading to an opposite cause-consequence ordering of the events.

Although the theoretical distinction between causal laws and scriptal knowledge cannot be directly translated into degrees of causal relatedness, Asher and Lascarides demonstrate that there are different classes of causal relations. It is a possibility that in processing causal coherence, these classes affect the construction of a mental representation in a different way. Several researchers have looked into the effect of the degree of causal relatedness on processing and recall of discourse, mostly as self paced reading experiments. Causal relations in discourse are found to improve recall (e.a. Black and Bern, 1981; Myers et al. 1994), and a higher degree of causal relatedness leads to a decrease in reading times of the second clause (Keenan et al. 1984, Myers et al. 1987). Keenan et al. (1984) as well as Myers et al. (1987), who used the same narratives as stimulus items, found that the percentage of correct recall varied systematically with the degree of causal relatedness.

Keenan et al. had created four different levels of causal relatedness by varying the cause sentence with respect to the number of steps between cause and consequence. Although they have not specified per level what specific steps are necessary, they have verified the degree of causal relatedness for the levels by asking participants to score the four cause-

sentences with respect to the consequence sentence on a 7-point scale for causal connectedness. According to Keenan et al, the dissimilarity in causal strength arises from a difference in the probability of the consequence. The researchers' intuitions that the relations differ in terms of causal strength was confirmed by the participants in the rating task. The sentence pairs below are examples of high, intermediate (2 levels) and low causal relatedness as used by Myers et al. The cause-consequence relation was varied over four levels.

Highly related:	Joey's big brother punched him again and again. The next day his body was covered with bruises.
Moderately related 1:	Racing down the hill, Joey fell off his bike. The next day his body was covered with bruises.
Moderately related 2:	Joey's crazy mother became furiously angry with him. The next day his body was covered with bruises.
Distantly related:	Joey went to a neighbours' house to play. The next day his body was covered with bruises.

Keenan et al as well as Myers et al. found that the differences of the recall between the causal strength levels did not vary in a linear way: Although causally related sentences were recalled better than sentences that were not causally related, highly related sentences were recalled less well than intermediately related sentences.

Both Keenan et al. and Myers et al. assume the fact that the highly related sentences take less processing effort to be the cause for poorer recall that was observed for this category. Figure 1 from Myers and Duffy (1990) gives a possible representation of the readers 'internal network, corresponding to three types of causally related sentence pairs. Starting with the middle figure, in order to create coherence in moderately (causally) related discourse, the reader needs to elaborate (E) by inferring additional information that is necessary to connect the two sentences S1 and S2, and the elaboration can logically link (B1 and B2) the sentences

together and a representation of the discourse (P) can be constructed. For highly related discourse, displayed in the first figure, no elaborative inferences are needed in order to infer the causal link (B), and to construct the full representation (P). In case of distantly (causally) related discourse, represented in the third figure, according to Myers and Duffy, no logical connection is made between S1 and S2.

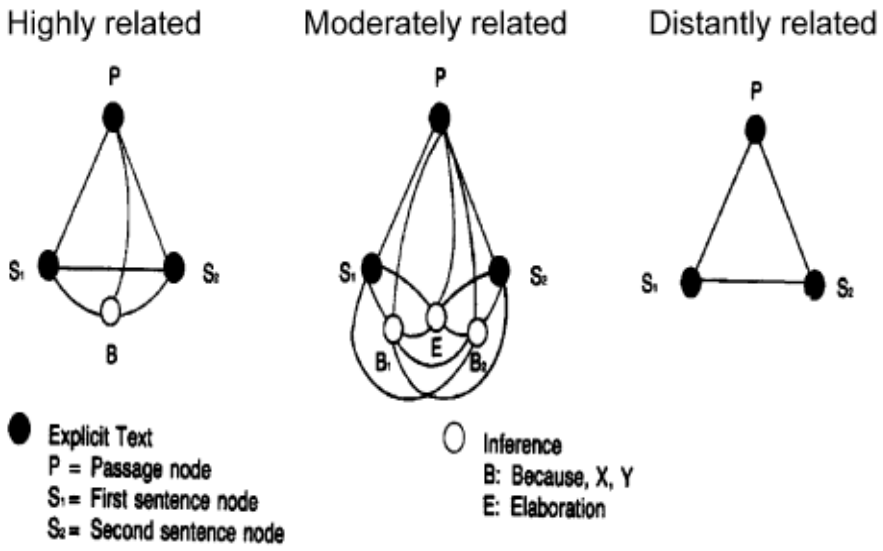

Figure 1: Possible representations of the readers'internal network, corresponding to three types of causally related sentence pairs (Myers and Duffy 1990).

Based on figure 1 the difference in recall and reading time between highly and moderately related discourse can be explained: the additional elaboration in the moderately related discourse causes more processing effort, which is reflected in the longer reading time, yet at the same time, more links are inferred, resulting in better recall.

A problem for this account might be that for lower levels of relatedness, processing time is significantly longer than for moderately related levels, and still recall is poorer. Myers et al (1987) explain the longer processing time as being an attempt to generate causal inferences and the deliberation of possible causal links. The lower recall for this level would then be caused by the lack of success in integrating any inferences in the representation since the sentence are minimally (perhaps only temporally or additively) connected (Myers et al. 1987). If this last assumption is true, in the case of the example for distantly related

discourse above, the resulting coherence would be only be a temporal one (since the discourse contains an explicit temporal marker) and the representation would be comparable to the third representation in figure 1.

This interpretation of the observed longer reading times and worse recall results in distantly related discourse can at least be questioned. It is clear that the processing and resulting representation in this category differs from the other two categories, but how can we be sure that no causal link is made? Intuitively, as a reader I would still want to find an explanation for the state in the S2 *The next day his body was covered with bruises*. And it is not at all hard to imagine a logical relation between the S1 *Joey went to a neighbours' house to play* and that S2. For example, you can imagine that the neighbours have very rowdy children and that playing with them left Joey with bruises, or that like in the moderately related 1 example, they went biking together and fell of their bikes. If it really is true that no causal link is made, a both Keenan et al and Myers at al suggest, the reason is that the elaboration turns out to be too complex, and the reader settles for a less complex and less informative temporal coherence relation. Could this be seen as a counterexample for the CbD hypothesis, which states that readers assume causal coherence when possible? At minimal, it seems to place a restriction on the level of complexity of the amount of inferences that are needed to establish a causal relation.

An alternative explanation for the longer reading times and lower recall, one that fits in the assumption of and preference for causal coherence better, is also possible: Readers connect the sentences causally, yet since the choice for an elaborative link that connects the S1 and S2 is so much more divers and complex (containing more steps at least) than those in the moderately related category, the reader does not infer one specific elaborative link but assumes only a vague, underspecified elaboration. Like Mulder (2008) I believe that the distantly related category requires the most elaborative processes. The elaborative link is possibly only additively of temporally linked to the S1 sentence, but serves as a causal explanation for the S2, for example in the form: *Joey went to a neighbours' house to play* + [something happened there] → *The next day his body was covered with bruises* (where + indicates an additive link, and → indicates a logical, causal link). This schema S1 + E → S2, differs from the schema that would represent the moderately related discourse: S1 → E →

S2. The latter supports a complete causal chain, with causal links between both segments and the elaborative link, whereas the former only supports a causal link between the elaboration and the S2. The longer reading times can then be explained by the lack of a logical (causal) relation between the S1 and the elaboration, and the lower recall to the vagueness of the link and the lack of a complete causal chain to represent the entire discourse. This interpretation would still fit the observed effects and at the same time follow the CbD hypothesis in that a reader maximally tries to find causal connections.

The above discussion demonstrates that there is a chance that possible causal links in the low causal relatedness condition are missed in the Myers et al experiment, The use of a task, such as recall questions, is perhaps not a precise enough instrument to test the representation of coherence relations. To automatically state that the representation of discourse in the distantly related category does not contain a causal link because recall is lowest seems a bit coarse. There are several options to solve this problem. For example, either an entirely temporal category where no causal relation is expected or possible, could be added, to which low causal strength processing could be compared.

Besides self paced reading experiments into causal strength where conducted within various different methods, such as fMRI (Mason and Just 2004) and ERP (Kuperberg et al. 2010). An advantage of examining neural activity as a measure for processing discourse is that it is a more direct measure than self paced reading. This directness can be used to examine processes on a word-by-word basis, and test for instance the increase of causal coherence across sentences (Kuperberg et al. 2010) and make clear whether the creation of coherence influences processing on a word-by-word basis and truly is an incremental process (Traxler et al. 1997).

Kuperberg et al. (2010) conducted an ERP study into establishing causal coherence relationships over sentences. The study examines the neural activity during online processing and comprehension of cause-consequence relations in narratives that are manipulated for causal relatedness on three levels: high, intermediate and unrelated.

The level of relatedness in the experiment was based on the complexity of the (causal) inferences that are necessary to establish causal coherence, and are similar to the levels in the experiments by Myers et al. In the class of high causal relatedness, the activation of

relevant world knowledge can be directly entailed. The level of intermediate causal relatedness, additional inferences need to be made that are specific to the situation described in the discourse. In the unrelated discourse, no clear inferences can be made to link the sentences causally, and the final sentence did not logically follow from the first two sentences. Examples of the narratives, which contain an introductory sentence, the cause-sentence and the consequence-sentence in the three levels of relatedness, are given below.

Highly related	Jill had very fair skin. She forgot to put sunscreen on. She had sunburn on Monday.
Intermediately related	Jill had very fair skin. She usually remembered to wear sunscreen. She had sunburn on Monday.
Causally unrelated	Jill's skin is always tanned well. She always puts on sunscreen. She had sunburn on Monday.

At least two factors are of importance in the design of this experiment: how are the levels of causal relatedness established and reflected in the discourse items, and what additional factors could be confounding the results. Kuperberg et al. did take these two factors into account. They conducted several offline pretests in which they asked participants to rate the discourse in the three levels of causal relatedness. Additionally, they controlled for semantic relatedness within the items by conducting a latent semantic analysis (LSA, Landauer and Dumais 1997) for each content word. Controlling for lexico-semantic relatedness made it possible to interpret the ERP results on a very local (word-) level instead of relying on an overall representation.

Kuperberg et al. looked at the effect of causal relatedness on representation as well as on processing. They found that causal relatedness, and more specifically the causal inferences, affects lexico-semantic processing at the earliest stage and on a word by word basis.

This means that the cause presented in the narratives is interpreted as such (a possible cause) as soon as possible, and the likeliness of the cause affects processing of the

consequence. This finding could be interpreted as support for the CbD hypothesis, since the causal interpretation happens as soon as possible.

There are some issues with the experiment by Kuperberg et al. that may have confounded the results. Firstly, the cause in the highly related category is presented as a problem (as least for this stimulus example) and is likely to generate some expectations about a consequence, whereas the other two categories do not do this. So, even though the semantic relatedness is kept the same by using LSA, the categories still differ on a pragmatic level. This is not necessarily a problem, since high causal relatedness is likely to generate expectations, yet the effect of the levels differs in this respect from that in the Myers et al. experiment.

Secondly, causal coherence was explicitly targeted by the task the participants were doing while reading the narratives: they were asked to rate how easy or difficult it was to connect the sentences causally by pushing a button (easy, in-between, difficult). The addition of this task is likely to have affected the reading process, making the participants focus more closely on the causal coherence relation. This makes it a little harder to interpret the results with respect to the CbD hypothesis: since the task caused the reader already to expect some kind of causal relation, an initial assumption of causality is already present due to this task.

In the sections above I have tried to provide some theoretical and experimental background to inferring and processing causal coherence that is relevant to the CbD hypothesis and the two factors that are investigated in this thesis. Even though SDRT (Asher and Lascarides 2003) is a theoretical approach to coherence interpretation, principles like *Maximize discourse coherence*, *Narration*, *push-causal law* and the use of situational scripts in coherence inferencing are compatible with the CbD hypothesis, and help us get a better insight into why causal coherence is a crucial factor in discourse processing.

The experimental results with respect to the degree of causal relatedness between two segments support the CbD hypothesis when it comes to the effect of low causal relatedness on processing. How can the results be related to the CbD hypothesis? Based on the CbD hypothesis the initial assumption of causality predicts longer reading times in the absence of

causality. Yet in the case of an intermediate degree of causality, the CbD hypothesis is not compatible with the longer reading times that are found for the second segment by for example Myers et al. (1987). After all, if causality is a default assumption, discourse that is causally related in a similar way (cause-consequence) should be processed in a similar way when the possibility of causality is present, since there is no need to drop the assumption of causal coherence. On the other hand, when causality is based solely on the content of the clauses, reading times are expected to differ since highly causally related clauses activate causal expectations, whereas less causally related clauses might not do this to the same extend, and in the case of no causal relation, no expectations are generated at all. In order to test the effect of causal relatedness and context on the CbD hypothesis in discourse processing the CbD hypothesis must be transformed into a processing theory.

In the next chapter I will look at an online experiment that tested the CbD hypothesis, and I will create a working hypothesis for CbD as processing theory.

4. Processing Causality-by-Default

In this paragraph I will look more closely at CbD as a possible processing theory. But before that I first have to go over some general matters to do with processing coherence relations in discourse. The literature reports two main processes in the reading and interpretation of (causally connected) discourse: integration and inferencing. The first process takes place when the information of the current segments is integrated with the previous information in order to construct a coherent discourse, as modeled in the c*onnective integration model* (Millis and Just 1994). Research shows that connectives play an important role in this process; they mark the type of relation between the two segments and help the reader to integrate the two segments (Noordman and Vonk 1997). When no connective is present, according to the model, the reader is not obligated to integrate the segments in the same way (or at all) as when the connective is present. Although Traxler, Bybee and Pickering (1997) proposed that the process of integration happens incrementally during the reading process, processing differences have been found between the integration of marked and unmarked discourse (Haberlandt 1982; Cozijn 2000).

The second process is inferencing the links necessary to construct a causal relation between two segments. Broadly speaking there are two theoretical approaches to how inferences are made: Albrecht and Myers (1995) propose a bottom-up approach where inference are constructed based on the input of the discourse and the representation. On the other hand, Graesser, Singer, and Trabasso (1994) propose a more top-down approach, where the search of the reader to infer coherence causes inferences to be made.

An interesting question is in what way these two processes interact with each other. Can inferences be made before the integration of the two segments? Can the integration of causally related segments happen without inferencing?

Li (2010) tested the CbD hypothesis experimentally by comparing reading times for both explicitly marked causal and temporal coherence relations as well as implicit versions in an online reading experiment. The results of the experiment by Li provided at least two observations that are relevant to the research in this thesis. Firstly she found that readers try to integrate clauses as soon as possible. The presence of an overt coherence marker caused

immediate integration at the start of the second clause, after the coherence relation was revealed by the overt marker. Implicit coherence results in later integration, when more information about the second clause became available. These observations can be taken as evidence for the *incremental integration model* by Traxler, Bybee and Pickering (1997), and show that a reader does not wait until the end of the second clause to integrate the clauses and construct a representation, but tries to do this as soon as possible.

Secondly, Li concluded that her results could be interpreted in favour of the CbD hypothesis. She found that causal coherence relations marked with a temporal marker, and therefore underspecified for causal coherence, did not differ in reading times from the same relations marked with a causal marker. Additionally, she observed that causally coherent implicitly marked discourse was read faster than temporally coherent implicitly marked discourse. These two observations, inferring a causal relation when the discourse is temporally marked, and longer reading times for unmarked temporally connected discourse, can be interpreted as that readers assume an initial causal interpretation, and do not drop this assumption even when a temporal marker is presented.

But why do readers not try to integrate implicitly marked clauses at the same moment in the reading process as explicitly marked clauses? If causality is assumed in both cases, implicit marking should not lead to slower processing.

Li states that this difference in integration does not pose a problem for the CbD hypothesis: "(...) This does not run into conflict with the causality-by-default hypothesis, since that hypothesis claims only that readers start out assuming that the relation between two consecutive sentences is a causal relation (given certain characteristics of two discourse segments) but does not specify when readers begin to make that 'causal assumption'." (Li 2010: 30). Here Li seems to suggest that a reader only assumes a causal relation at the moment that a coherence relation between the sentences is expected, at the connective marking. This implies that there is a possibility that the reader might not expect coherence at forehand. I do not agree with this: a reader assumes a text to be coherent, and therefore assume coherence relation to exist between the segments. Furthermore, if causality is assumed by default, these coherence relations are causal ones. Therefore, this difference in processing between marked and unmarked causal coherence could still be problematic for the CbD hypothesis.

It must be said that the experimental setup of Li's experiment was constructed in such a way that made it impossible to compare marked and unmarked coherence directly at the beginning of the second clause, since that specific region contained a connective in the marked conditions where there was none in the unmarked condition, causing longer reading times due to the region containing more words. This means that the question how the CbD hypothesis affects coherence processing without overt coherence marking remains open. Similarly, even though Li compared temporal with causal coherence, the way the information in the first segment (the position where the type of coherence was determined) affected coherence processing or the assumption of causality is not clear.

To be able to test the CbD hypothesis in implicitly marked discourse, I have constructed a working hypothesis for CbD as a processing theory. When CbD is truly interpreted as a cognitive mechanism of interpretation, one that is not limited to text but affects the interpretation of events in the world around us, the default expectation of causal coherence must be in place before the reader starts processing the discourse, since the existence of the assumption is independent of the content of the specific discourse segments. This then inevitably results in the possibility that anywhere during processing of the two segments the assumption of causality may be dropped. In other words, both the first and second segment may influence the assumption of causal coherence: the influence of first segment depends on the degree of causal strength of that segment and whether any causal expectation is generated, while the influence of the second segment depends of the degree of causal relatedness between the two segments.

Also, CbD predicts that a processing difference exists between causal and non-causal coherence, since in the latter case the assumption must be dropped and the coherence relation has to be re-analysed, leading to longer reading times. In case of a different degree of causal strength or causal relatedness, the CbD hypothesis does not a predict processing difference, since both in case of high and intermediate causal strength or relatedness, causal coherence is present in the discourse and the assumption of causal coherence does not need to be dropped.

A working hypothesis for Causality-by-Default as a processing theory can be summarized as follows:

1. The assumption of causal coherence is in place at the start of processing the first segment and can be affected by characteristics in both segments.
2. A difference of causal vs. non-causal coherence results in reading time difference, while no differences are predicted when causal coherence is inferred.

With the above working hypothesis as a basis I have designed an eye tracking experiment to test the validity of the CbD hypothesis. The experimental design incorporates *causal strength* and *context* as experimental factors, and aims to test processing (inferring) of coherence relations, as well as the representation after the relation is inferred and the clauses integrated. With the factor *context*, I will try to find out how the preceding discourse affects processing and the representation of the coherence relation. With the factor *causal strength* I will try to find out how the degree of causal strength of the first segments and the causal relatedness between the segments affects processing and interacts with the assumption of causal coherence.

Short narratives that contain cause-consequence relations are presented as stimuli. The target sentences in the eye tracking experiment are S1, S2 and S3, where S1 is the cause sentence and is manipulated for causal strength, and S2 is the consequence sentence. In these sentences the reading times are interpreted as an interaction between the experimental factors and CbD with respect to processing the coherence relation.

The S3 is the segment where the representation of the coherence relation is tested. This sentence makes the reader reflect on the causal relation between S1 and S2. The experimental factor *context* is presented as a sentence preceding the S1. In the next section I will go further in the choices I made with respect to the experimental design and the experimental factors.

5. Experimental choices

I have proposed that the CbD hypothesis needs more fine tuning when it comes to processing implicitly marked discourse coherence. How can CbD be tested when we do not know in what way textual aspects influence this cognitive assumption? I have indicated two types of influences: the more global, such as expectations build up by the reader based on knowledge about text structure and contextual influence, and local, such as lexical-semantic influences on coherence inferencing that might affect causal relatedness.
In the experimental set up I want to focus on these factors: context as a more global factor and causal strength as a local factor. Both are likely to affect the processing of causal coherence. Although it is not my intention to state that these are the only or the most important factors that are of influence. Factors like referential coherence, type of causal relation (for example forward or backward and subjective or objective), and many more play a role in processing causal coherence, but since they have already been extensively researched and they are not in the scope of this paper, I have tried to keep these factors out of the experimental set up by keeping them similar over conditions and items.

Research into discourse processing and representation is mostly conducted with help of the reading time measures, either in self paced reading or eye tracking (e.a. Traxler et al. 1997; Sanders and Noordman 2000; Li 2010). In the interpretation of eye tracking results reading time is seen to reflect processing effort. Even though there no one to one mapping between the reading time of a certain part of a discourse and the time it takes to process the information in that specific component of the discourse, since later processes such as integration of text with the preceding discourse must also taken into account, experimental eye tracking data can give a deeper understanding of what processes are part of discourse interpretation. I have conducted an eye tracking experiment that focussed on processing causal coherence in short narrative discourses. The discourses are constructed in such a way that a cause-consequence relation is possible between two segments: the S1 (possible cause) and the S2 (possible consequence).

Implicit coherence

It is important to point out that the experimental items are created in such a way that the reader is not obliged to create a causal link between the events; the causal coherence relation remains implicit and is not overtly marked by a causal connective. In this way, when a causal link is inferred and causal coherence is constructed, the reader has inferred this link only based on the semantic information in the S1 and S2 segments. For the stimuli items, causal coherence is likely to be preferred over additive or temporal coherence since an important characteristic of the event in S2 is that it needs a volitional cause. In other words, the event in S2 can only happen through an (animate) agent. A referential coherence link between the S1 and S2 ensures that the reader is prevented from interpreting the two sentences as being part of different event sequences.

Cause-consequence order

I have chosen to use *cause-consequence* relations (result relation) to test causal coherence and the effect of causal strength. There are two reasons for this. Firstly, other structures, such as *problem-solution* generate expectations about the structure of the relations (Mulder 2008). Reading the *problem* part leads the reader to activate the structure from memory and anticipate a solution. This makes it easier to process the solution-part (cf. Van den Broek, 1990). Secondly, backward causal relations, such as *consequence -cause* (explanation relations) are claimed to be 'canonical' (Moeschler 2001). Readers generally prefer explanation relations over result relations, since often a cause is easier to infer than a consequence. This preference could be useful when a reader does not construct the cause-consequence relation in the stimuli items due to the low causal strength in the cause sentence. Since a consequence-cause relation is preferred, the reader is expected to anticipate the cause to follow the consequence (instead of precede it) and that the coherence can still be constructed when the following discourse does provide a cause. This could prevent the reading time effects in the target regions from getting muddled by reader behavior when a reader was not able to construct any coherence between S1 and S2 and starts rereading these specific clauses. To prevent readers from reading the S1 as a consequence and the S2 as a cause, which might be a possible link for some of the stimuli items, both clauses are in present tense. Present tense, together with the referential

coherence ensures that the only interpretation is a *Narration* interpretation (Asher and Lascarides 2003), where the order of the events in the coherence relation is the same as the order of the segments in the discourse.

Representation: recall task vs. S3

The causal interpretation of short discourses such as *John entered the room. Max left.* demonstrates our preference for causal coherence. When *John entered the room* is read, hypothetically, two interpretations of the function of this sentence in the discourse are possible: it is a mere description of ongoing events, and is for example followed by *He opened a window*, or it is a set up for what happens next, i.e. a cause. Because the event in the second clause, that of someone leaving, needs to be instigated by something, the second interpretation is likely to be made, resulting in a cause-consequence relation. Yet in a reading experiment where I want to test whether these sentences are causally interpreted under the assumption of causal coherence, how can I be really sure that the reader actually constructs this specific representation?

In discourse coherence research it has often been claimed that causally related discourse is represented better than additively related discourse. These claims are based on the results of recall and verification tasks conducted after the participant read the causally or otherwise related discourse segments. The tasks can be used to test to the coherence relation in the text, for example a verification tasks that ask after the cause for an event. Even though recall and verification tasks yield good results, it is not a very useful tool for this experiment. There are at least four disadvantages when using a verification task that targets causal coherence: Firstly the tasks are carried out after the complete discourse is read. This way, only the overall representation of the discourse item can be used as a source for the task, and not a specific local representation of the relation between two segments. Secondly, when the reader had not yet created a causal representation, the task itself, when asking the reader to confirm or deny a causal link, can be a cue for inferencing that specific causal link. Thirdly, sloppy reading or answering of the task can cause unwanted variation. Lastly, the task can trigger the participant to actively search for causal coherence in the stimuli texts.

I have tried to find another method that tests representation, without this representation being affected by a task of some sort. A causal interpretation of the relation between two clauses can be supported by adding additional information to the discourse with stronger causal relatedness to the second clause, as in: *John entered the room. Max left. He disliked John very much*. The added information in the third clause supports a causal connection between the first and the second clause. When a reader has inferred a causal coherence relation, the third sentence supports this relation. When a reader has not inferred a causal relation, the third sentence will cause a reading time effect when compared with cases where a causal relation is inferred.

This set up can be a useful way to test the representation of local coherence without interference of an explicit task. It is more direct than a task, and no offline reasoning is necessary. Unlike with an offline task, the effect of *causal strength* on the representation of causal coherence is not confounded by memory effects.

Yet there are some drawbacks of this method to test representation too. Firstly, the content of the third sentence must ensure the causal link, even in experimental conditions where a causal link in not very likely. This can make the discourse seem unnatural, even though the effect of this unnaturalness is only placed after the processing of the third sentence. Secondly, a baseline condition is needed to test whether processing in other conditions divers from the expected. This can be for example a condition in which it is certain the reader constructs a causal link. When reading times differ from the reading times in this condition, coherence representation is likely to differ also, since the content of the third sentence is the same over conditions. Because no explicit markers of coherence are presented in the stimuli texts in this experiment, it is not immediately clear whether or not a reader constructs a causal coherence relation. As such a baseline the strong causal coherence relation in the context condition is taken. In this condition, with both high causal strength and a context that aids inferencing, it is almost certain that a reader constructs a causal coherence relation.

Based on the experimental design features as described above I have composed research question that target the CbD hypothesis and the assumption of causal coherence in discourse.

6. Research questions and predictions

With an eye tracking experiment I set out to further refine Causality-by-Default as a hypothesis for a cognitive processing strategy for implicit coherence in narrative discourse, by focussing on the two factors identified before: *context* and *causal strength*.
The research questions with which I want to investigate the viability of the Causality-by-Default hypothesis and the possible effects of the textual factors *context* and *causal strength* on the hypothesis are:

1. a. Does *context* affect the assumption of causal coherence? If so, at what point in processing?
   b. What is the effect of *context* on inferring and processing causal coherence?
   c. What is the effect of *context* on the representation of an implicitly marked coherence relation?

2. a. Does *causal strength* affect the assumption of causal coherence? If so, at what point in processing?
   b. What is the effect of *causal strength* on inferring and processing causal coherence?
   c. What is the effect of *causal strength* on the representation of an implicitly marked causal coherence relation?

Predictions for the effect of Causal Strength on processing

The CbD hypothesis predicts that when it is possible, a reader assumes causal coherence and tries to infer causal links. This means that, depending on certain characteristics, a reader might drop the assumption of a causal relation between the segments. Since the assumption is in place initial to processing, the characteristics in the S1, the S2, and how the S1 and S2 relate to each other, can be possible reasons to drop the assumption. It is this possibility that I want to investigate further with the factor *causal strength*, which contains three levels: high, intermediate and low causal strength. The factor *causal strength* can be divided into two separate effects: the effect of causal strength on the assumption of causal coherence, and the effect of causal relatedness on processing and the construction of a coherence

representation. The first effect is due to the verb phrase in the S1, and is applicable to the reading times in the S1, when the reader cannot yet construct a complete representation. The second effect is due to the relation between the event in S1 and in S2, and is applicable to the reading times of the S2, when S1 is processed and the S2 can be integrated with S1, and a coherence relation can be inferred.

When we take a two sentence discourse, such as *John pushed Max. He fell.* it seems likely that the causal strength of the first segment can be influential when it comes to expecting a causal coherence relation or dropping the assumption of causal coherence. But will a reader drop the assumption of causal coherence based on low causal strength in the first segment? This would go against the idea that a reader prefers the most informative coherence relation (Graesser et al. 1994; Asher & Lascarides 2003; Sanders 2005), since even in case of low causal strength; a causal coherence relation might follow. Yet on the other hand, keeping an expectation of causal coherence when this is not at all supported by the content of the first segment could be very un-economical when it comes to processing. If it is the case that an initial assumption of causal coherence is in place, and this can be dropped based on the lack of causal strength of the S1, reading time difference are expected in S1 between low and high causal strength conditions. In case the CbD hypothesis does not hold and expectations are based merely on the causal strength of S1, a reading times difference between high, intermediate and low causal strength conditions is expected. It could also be that the assumption of causal coherence is not affected by the causal strength of one segment, but only by the strength of the causal relation between the two segments. If this is the case, reading time difference between high and low causal strength are expected in S2, when the representation of the coherence relation can be constructed. If the CbD hypothesis does not hold, we would again expect to observe a reading time difference between high, intermediate and low causal strength.

High vs. intermediate causal strength

According to Myers et al (1987) one of the differences between high and intermediate causal strength is due to a difference in degree of causal relatedness is that highly causally related discourse simply entails the world kowledge that is necessary to create the causal link, while

intermediately related discourse relies on more situation specific inferencing. For highly causally related discourse, of the type *push-fall*, Asher and Lascarides (2003) have proposed that the relations might be governed by laws based in world knowledge, while for intermediate causally related discourse, from the type *greet-stand up*, scripts build up patterns based experiences that are closely linked to specific situations. Based upon the difference that is found in recall and processing time (Keenan et al 1984, Myers et al. 1987), and the predictions of how these difference come about, it could be concluded that the different degrees of causal strength result in two different types of processes of coherence construction. Does this possibility of two different processes pose a problem for the Causality by Default hypothesis? Not immediately: the assumption of causality is distinct from inferencing causality, since the latter relies on specific information that is presented in the discourse, where as the former is a starting point from which the reader uses the information in the discourse to seek confirmation for the assumption.

Now we make predictions about at what point in the discourse the assumption of causal coherence is confirmed or denied. The theoretical distinction together with the recall difference leads to the prediction that processing difference between the two conditions might be found at the point where the inferences are made in the intermediate causal relatedness condition, since these will take more processing effort than those in the highly causally related condition. This is at the part of the sentence where the information about the two events is fully accessible: at the end of the second segment.

When the assumption of causal coherence is not present in discourse processing and the CbD hypothesis does not hold, a processing difference is predicted to come up in the first segment. The *push-causal law* will immediately activate expectations about the causal continuation of the discourse after the encounter of the verb phrase in the first segment, and therefore speed up processing, while in the intermediate condition causal inferences can only be made when a complete script is retrieved, at the end of the second segment.

Low causal strength

When causal coherence is assumed, and Causality-by-Default holds, there are two options for coherence construction in case of low causal strength: either the information in the two

segments is sufficient to infer causal coherence, or the information is insufficient and the assumption of causal coherence must be dropped in favour of a temporal of additive coherence relation. In both cases reading times in the S1 will not differ from those in the other conditions because at this point, the CbD assumption does not need to be dropped. The difference will arise at the end of the S2. In case of the first option, enough information to infer causal coherence, reading times are predicted to be similar to those in the intermediate condition, since in both conditions additional information needs to be inferred. When no causal relation can be constructed, as in the second option, reading times are predicted to be longer than in the intermediate condition, since the discourse structure needs to be reviewed.

If the CbD hypothesis does not hold, a reading time difference is predicted for end of the first segment between high and low causal strength conditions. Because no expectations of causal coherence are activated in the latter condition, the reading time is not predicted to speed up in this condition, as opposed to the high causal condition since in the absence of causal expectation, there are many more options for possible continuations than when the first segments can be identified as a cause.

I would not predict a reading time difference between intermediate and low causal strength in the first part of the second segment; in both conditions causal inferences are only likely to be made at the end of the second segment, and therefore no assumptions of causality are present before this. Predictions for a possible difference between the intermediate and low causal strength condition at the end of the second segment are tricky. At this point in processing there is enough information to infer causality in the intermediate condition, and when possible also in the low condition. But what happens when in the low condition no causal coherence is constructed: Does constructing a temporal or additive representation without there having been the assumption of causal coherence takes less processing effort than needing in to infer causal coherence when no previous assumption was made? In this case it could be very likely that without the assumption of causal coherence, simple temporal or additive relations are inferred faster than causal relations without pre-existing causal expectation. This is somewhat precarious though, because at this point the story becomes a little more complex: When a cause-consequence relation is not expected, in case of low causal strength, it means that the reader does not find the event in the first segment

S1 a likely enough cause for the event in the second segment S2. Yet in the stimuli items, the consequence in the S2 is constructed in such a way that it needs a cause. And besides the above mentioned temporal or additive coherence relation the reader has another option to create coherence: to postpone the construction of coherence in favour of a backward explanation relation (consequence-cause), where the event that is to be causally related to the S2 expected in the discourse that follows the S2. A reader can wait for later information before the resolving the coherence relation of the current segment (Caenepeel 1991, Asher and Lascarides 2003). How this postponing of inferring coherence affect the reading times, I cannot say.

The predictions for the effect of *causal strength* on processing can be summarized as follows:

Prediction 1: Degree of causal strength (high vs. intermediate)

a) If Causality by Default: An effect of causal strength is expected to be observed at S2.

b) If no Causality-by-Default: An effect of causal strength is expected to be observed at S1.

Prediction 2: Causal or no causal processing (high vs. low)

a) If Causality by Default: An effect of high vs. low strength is expected to be observed at S2.

b) If no Causality-by-Default: An effect of high vs. low causal strength is expected to be observed at S1.

c) If no Causality-by-Default: No effect of intermediate causal vs. low causal strength is expected to be observed at S2.

Predictions for the effect of Causal strength on representation

Because no marker of coherence is explicitly present in the narratives, there is no certainty about the type of coherence relation the reader constructs. In section 5 I have discussed the issues to do with testing representation with recall questions. Instead of using recall, the items contain a third sentence, S3, where the representation that is constructed for the coherence of the two previous sentences S1 and S2 is tested by comparing the reading times.

In order to be able to find representational differences, a baseline is constructed: a condition in which there is the most certainty about what kind of coherence representation (causal or not causal) the reader constructs. As this baseline I have taken the reading times in the condition in which the reader seems most likely to create a causal coherence relation: the high causal strength condition. When reading times in other conditions are longer than those in the high causal strength condition, it will be interpreted as evidence that no causal representation was constructed. When no differences are found in the reading times for the representation sentence S3, I assume that causal coherence is inferred.

Based on the CbD hypothesis, it is predicted that both in the high and intermediate causal strength conditions, a causal coherence relation will be constructed, and that therefore no reading time difference between these conditions will be found in the S3. In the case of low causal strength, it depends on whether the reader has enough information to construct a causal link between the S1 and S2. When this is not the case, the reading times in the S3 are predicted to be longer than in the other two conditions.

Prediction 3: Representation

a) If *causal strength* affects the construction of a causal representation, an effect is predicted at S3.

Predictions for the effect of context

The factor *context* has two levels: context or no-context (null context). The function of context as an experimental factor is to allow for inferencing additional information in order to make it easier to create a causal link. A reader builds expectations about an explanation or backup for the statement that is presented in the context sentence. Also the establishment of the agent in the discourse is likely to aid inferencing: more information about the agent in the situation makes it easier to access relevant world knowledge about the situation.

These two factors together predict that contextual information aids inferencing, and therefore speeds up processing, especially at the point where the necessary inference are to be made, in the S2.

From a causality-by-default point of view, no effect of context is predicted in the reading times of the two segments that are to be linked causally. When the assumption of causality is already present, a context that enhances causality is not necessary to generate the assumption of causality. If the CbD hypothesis does not hold, context is likely to make inferencing easier in the intermediate and low causal strength conditions, but not in the high causal strength condition, since here inferencing is already very easy.

Prediction 4: Context

a) If Causality by Default: No effect of *context* is expected on inferencing and constructing causal coherence in S1 and S2.

b) If no Causality by Default: during coherence processing in S1 and S2, *context* is expected to interact with *causal strength.*

7. Experiment

7.1 Method

Materials

The materials for the offline experiment consisted of two sentence S1 and S2 in Dutch. S1 described an action conducted by an agent towards a patient (either an object or a person) and S2 described the patient of S1 in a non-volitional event, in such a way that the S2 could be a possible consequence of the action in S1. Two experimental factors were introduced. Firstly, for the factor *causal strength*, three variants of the transitive verb phrase in the S1: high, intermediate and low *causal strength*. Secondly, for the factor *context*, two levels were created, one without any context and one with a subjective context sentence that preceded the S1 sentence. This sentence contained the agent of the S1 as subject combined with a predicate that conveyed an opinion of some sort.

The factor Context

The factor context is added to the experiment as a factor that influences the processing of the cause-consequence relation on a global level. The context itself is not part of the cause-consequence relation (not part of S1 and S2) but precedes this. The factor contains two levels: either the S1 and S2 are preceded by a context sentence, or there is no context.

The context sentence consisted of a statement about the agent in the discourse. This statement revealed a subjective aspect of the agent as a person, for example that the agent was happy, or a bad driver, etc.

The function of this context sentence is twofold: Firstly the presence of an introductory sentence establishes the agent as a prime character in the discourse, making it easier for the reader to reason about him or her. And secondly the statement reveals a certain aspect of the agent that is likely to lead to inferences that are similar as those necessary to infer the causal link that will follow in the discourse. For example when the context states that the character is a bad driver, this information can act as a topic for the following discourse.

The factor Causal strength

The factor *causal strength* is comparable to what Meyer et al. (1987) and Kuperberg et al. (2010) call causal relatedness. The factor is added as a characteristic that influences processing on a local level. The verb phrase in the S1 is manipulated in order to create three levels of causal strength: high, intermediate and low. High causal strength means that a relation between the cause in S1 and the consequence in S2 can be constructed without any inferences of additional events in the discourse. For the intermediate level the event in S1 is very likely to be interpreted as a cause for the event in S2, but in order to construct this causal link, some additional, situation specific inferences are needed. For the third level, low causal strength, the S1 consists of an event that is very unlikely to be causally linked to the S2, although this link is still possible when the reader is prepared to add more and complex inferences to the final representation. In the S1, the verbs in the level low causal strength are mainly verbs of perception, where substantial background information is necessary to infer construct causal link. The majority of the items are based on a causal relation for physical change, such as *push-fall*, *drop-break* or *hit-harm*. A small percentage (approximately 7 percent) is based on causal links to do with emotional reactions, such as *shout at- get scared*, and *swear at- get angry*.

The levels of causal strength were created based on intuitions about the amount and complexity of the inferences necessarily to create a causal link. Besides that, two discourse researchers criticized and reviewed the stimuli materials during the construction process. In order to test the effect of the factors *causal strength* and *context* on how easy or difficult it was to create a causal link, I conducted an offline pre-test. The purpose of pre-testing the relations between the verbs in the three conditions is two-fold: Firstly I wanted to be certain that the three categories I created for the *causal strength* factor truly differed with respect to the complexity of creating a causal relation between the two events. Secondly I wanted to test the effect of the context: was a causal relation easier to create when a context sentence preceded the S1-S2 sentences.

The pre-test was presented as a pen-and-paper experiment in a classroom setting. The participants, 66 first-year students of Utrecht University (av. age 20 years), were presented with 24 different discourses. The experiment was a 2 by 3 design, where each participant

only got to read and rate one version of each text. Participants were given three stimuli texts per page, with the S1 and S2 on different lines. For each text they were asked to rate the 'difficulty' of combining the events in S1 and S2 in such a way that the first event caused the second event, on a response scale from 1 (very difficult) to 7 (very easy)[2]. The participants were not compensated for their efforts.

The acquired dataset consisted of a set of ratings per participant, per item. I have analyzed the ratings with a Linear Mixed Effects Analysis. An ANOVA comparison of the models was made to determine the significance of the effects. The ratings showed a significant main effect for *causal strength*. Items in the high causal strength condition were rated significantly higher (average score: 6.4) than those in the intermediate condition (average score: 5.0; β = 1.4, SE = 0.092, $p < .001$), and the items in the intermediate condition were rated significantly higher than those in the low causal strength condition (average score: 3,2; β = 1.78, SE = 0.091, $p < .001$). No main effect for *context* was found but there was a significant interaction effect: the scores in the low:context condition were significantly higher than in the low:no-context condition (β = .5, SE = 0.134, $p < .001$) whereas no significant effects of context were found for the other two *causal strength* conditions (β = .13, SE = 0.085, p = .104; β = .07, SE = 0.122, p = .582).

In other words, the three levels of causal strength were rated with a significant difference for the difficulty of creating a causal coherence relation. The effect of context was only apparent at the most difficult level, where the items with context were seen as less difficult when compared to the no-context items. The absence of an effect for *context* in the high causal strength level was expected; the causal relations were already very easy to create: the

[2] Additionally, the participants were asked to write down any characteristics of objects, persons or of the situation in general that would be necessary to connect the two events in a causal way and create a coherent story. It was allowed to write that no additional information was necessary. These remarks can be expected to be comparable to the causal links a reader might infer when connecting the two segments causally. The remarks were categorized depending on what type of additional information they introduced. The categories were: additional characteristic of the agent, additional characteristic of the object/patient, and additional event, agent, or object/ patient.

A quick analysis of the remarks showed that the complexity of the causal links became higher in proportion as the difficulty rating increased. Although in some cases, an item in low causal strength condition was scored as very easy (7), yet the causal links in the remarks were very complex (contained an additional agents, events, etc).

The resulting scores for the three conditions were not controlled for the complexity of the causal links in the remarks, due to time constraints.

participants rated then at the top end of the scale, both with or without context[3]. The absence of an effect for *context* in the ratings for the intermediate level was not expected. An explanation for this lack of effect could be that perhaps the participants did not rate the difficulty of coming up with inferences, but the complexity of the inferences themselves. Although the inferences might be easier to make in the context condition as opposed to the no-context condition, the participants rate the difficulty the same since the inferences that are necessary to construct the causal link are the same in both conditions.

The significant effect of *context* in the low causal strength condition could then be explained by assuming that the context helped the participants to come up with a suitable causal link, since the context provided more information about the situation at hand.

Materials in the online experiment

The stimuli items in the online experiment were similar to those in the offline experiment. A third sentence that consisted of two short coordinated clauses, was added to the stimuli items. The first clause (the S3) referred back to the agent in the S1 sentence (the cause) and described an emotional reaction of this agent. The second clause described an event that ended the discourse. In (8) an example of the S3 and the closing sentence is given. The function of this third sentence was two-fold: Firstly, the first clause (S3) in this sentence created a link between the agent in the S1 and the event the S2, suggesting that the agent was responsible for what happened in the S2. It marks the agent in S1 as being the same agent that has caused the event in S2 by assigning an emotion to the agent in such a way that it can only be interpreted as an emotional reaction to the event in (S2), which will therefore be interpreted as the consequence of an action by this agent. Secondly, the second clause acts as a spill-over region for the first clause and at the same time concludes the discourse and wraps up the short narrative.

Readers that have inferred a causal relation between S1 and S2 will find this inference supported by the S3 sentence. When a causal link in the preceding discourse was not inferred, the S3 sentence will force the reader to reconstruct the mental representation of the discourse when possible.

[3] The rating of the high causal strength condition at ceiling level rectifies the use of this condition as a baseline for which can be assumed that a reader will construct a causal coherence relation.

Stimulus presentation and task

The layout on the computer screen was carefully constructed in such a way that for each item, the sentences were place in the same position, on the same line. Even though it was not possible to prevent target regions to be placed at the beginning or the end of a line, which can affect fixations due to line sweeps, over items and conditions the positions were kept similar. The layout on screen was presented as in the examples below.

A verification task was added in order to ensure that the subjects would read the texts carefully and not scan or skip parts of the text. This task consisted of a statement that had to be confirmed (answered with 'yes') or denied (answered with 'no'). The statements were the same over conditions and carefully constructed in such a way that they did not refer to a causal coherence relation in the discourse. This was of importance because I wanted to prevent that the participants would search for a causal relation between segments in expectation of the task.

Half of the texts in the experiment where followed by such a verification task.

8: Example of the manipulations in a stimulus item (containing the context sentence, S1, S2, S3 and the closing sentence).

High:context

Johan is erg agressief.
Hij slaat een portier. De portier raakt gewond.
Johan probeert het goed te praten en gaat dan vlug naar huis.
Johan is very aggressive.
He hits a doorman. The doorman gets hurt.
Johan tries to make up and then goes home quickly.

Intermediate:context

Johan is erg agressief.
Hij bedreigt een portier. De portier raakt gewond.
Johan probeert het goed te praten en gaat dan vlug naar huis.
Johan is very aggressive.
He threatens a doorman. The doorman gets hurt.
Johan tries to make up and then goes home quickly.

Low:context

Johan is erg agressief.
Hij praat met een portier. De portier raakt gewond.
Johan probeert het goed te praten en gaat dan vlug naar huis.

Johan is very aggressive.
He talks to a doorman. The doorman gets hurt.
Johan tries to make up and then goes home quickly.

Verification

Johan raakte gewond.
Johan gets hurt.

Twice as much fillers as stimuli items were added to the setup. The fillers consisted of similar narratives that both contained cause-consequence and consequence-cause order causal coherence relations. This variation ensured that participants would not anticipate the type of causal relation in the narrative.

Stimuli lists were created in order to make sure that every participant would be shown each of the 24 items and all 6 conditions, but with only one version of each item per list. This resulted in 6 lists. The items and fillers were put in pseudo-random order: no more that 2 test items, and no more that 2 times the same condition followed each other in the experiment. A second set of lists was created in the opposite order and the 12 resulting lists were divided into 4 blocks.

7.2 Participants

The participants were 47 students and employees of Utrecht University, 6 males and 41 females. The age ranged from 19 to 50, with an average of 24. They were all native speakers of Dutch with normal or corrected-to-normal vision.

7.3 Procedure

The eye tracking experiment was conducted with an Eye Link 1000 Remote. The participants were seated in front of the Eye Link camera and a computer screen. A written instruction was given to each participant, explaining the procedure and the task. A calibration of the eye movements was conducted at the start of the experiment, and at the start of each

subsequent block. When the calibration was successful, the participants were presented with a short trial of two stimuli and a task, after which the actual experiment began. After reading a text, the participant could go forward to the next text or respond to the verification task by using the buttons on a response box. In total the experiment including instructions and calibration took around 20 minutes per participant.

7. 4 Analysis

Each text item was divided into 5 regions (see (9). Region 0 consisted of the context sentence, and was therefore absent in half of the experimental items (those in the no-context condition) and not part of the analysis. Region 1 consisted of the S1 sentence that contained the experimental manipulation of the verb. Region 2 consisted of the S2 sentence, and regions 3 and 4 consisted of the two clauses in the final sentence, with the connective being part of the fourth region. The regions of interest were regions 1 to 4.

9 Johan is erg agressief.
/$_1$ Hij slaat een portier. /$_2$ De portier raakt gewond./
/$_3$ Johan probeert het goed te praten /$_4$ en gaat dan vlug naar huis./

Johan is very aggressive.
/$_1$*He hits a doorman.* /$_2$*The doorman gets hurt*./
/$_3$*Johan tries to make up* /$_4$*and then goes home quickly.*/

For each of these regions I have analyzed the reading times in four different reading time measures with a linear mixed effects regression analysis. The reading time measures are *First pass reading time*, that consists of the first fixations on a region before it is left (either progressively or regressively) including the out-saccade from that region into the next region. *Right bounded reading time*, that consists of all fixation on the region before it is left in a progressive manner. *Regression path duration* consists of the total reading time between the first fixation on a region and the last. *A*nd finally *total reading time* that includes all fixations in a region.

8. Results

Table 1 gives the mean reading times per condition for each region of interest and in each reading time measure. I have analyzed the reading times in order to find main effects and interaction effects for the two conditions *causal strength* (3 levels: high, intermediate, low) and *context* (2 levels: context, no-context) with a *linear mixed models regression* analysis on the log of the reading times. The model contained the factors *item* and *participant* a crossed random factors.

Table 1: Mean reading time (ms) for condition, region and reading time measure (example in high causal strength condition).

	Johan slaat een portier. *Johan hits a doorman.* **Region 1**	De portier raakt gewond. *The doorman gets hurt*. **Region 2**	Johan probeert het goed te praten *Johan tries to make up* **Region 3**	en gaat dan vlug naar huis. *and then goes home quickly.* **Region 4**
First pass reading time				
High- context	863	695	656	672
Interm. - context	824	713	656	686
Low -context	856	762	678	657
High – no context	1340	693	664	698
Interm. no-context	1300	744	715	731
Low- no context	1326	757	739	716
Right bounded reading time				
High- context	930	709	660	842
Interm. - context	847	749	680	871
Low -context	873	787	719	880
High – no context	1355	752	692	861
Interm. no-context	1307	780	759	908
Low- no context	1337	820	746	985
Regression path duration				
High- context	997	747	675	1688
Interm. - context	866	825	707	1805
Low -context	898	845	809	2065
High – no context	1355	800	729	1518
Interm. -no context	1307	821	809	1632
Low- no context	1337	876	766	2073
Total reading time				
High- context	1159	824	830	842
Interm. - context	1115	903	823	871
Low -context	1261	1012	935	879
High – no context	1706	857	840	861
Interm. no-context	1661	932	925	908
Low- no context	1895	1057	1002	985

At region 1 a main effect of *context* was found: the reading times in the context condition are significantly shorter than on the no-context condition, for all reading time measures (*First Pass*: χ2 [1]= 399,48; p< .001; *Right Bounded*: χ2 [1]= 406,01; p< .001; *Regression Path*: χ2 [1]= 344,43; p< .001; *Total Time*: χ2 [1]= 341,32; p< .001).

For *causal strength* also a main effect was found in all reading times measures, minus *First pass reading time* (*Right Bounded*: χ2 [2]= 9,55; p< .01; *Regression Path*: χ2 [2]= 10,69; p< .01; *Total Time*: χ2 [2]= 42,21; p< .001). For *right bounded reading time* and *regression path duration,* the reading times in the intermediate condition were significantly shorter than those in the high causal strength condition (intermediate< high: β= .079, SE= .026, p< .01; intermediate< high: β= .088, SE= .027, p< .01). In these two reading time measures, no significant difference was found between the reading times in the high causal strength and low causal strength conditions, or between intermediate and low causal strength.

For *total reading time*, a significant difference in reading times was found between all three conditions: the reading times in the intermediate causal strength condition were significantly shorter than those in the high causal strength condition (β= .071, SE= .028, p< .05) and those in the low condition (β= .140, SE= .028, p< .01), and the reading times in the high causal strength condition were shorter than in the low causal strength condition (β= .069, SE= .028, p< .05).

Although the number of words in this region was not similar between conditions for each item since in this region the verb was manipulated for causal strength, the number of words did not significantly differ between conditions (F (2,69)= 2,78, p= .07). Therefore the effect that is observed can still be ascribed to the semantic difference of the verbs between the three causal strength conditions. The most striking effect in this region is the significant difference between the intermediate and the high causal strength condition: intermediate is read faster than high causal strength. The presence of an effect of *causal strength* in region 1 could be seen as problematic for the CbD hypothesis since is complies with prediction 1b, yet the direction of the effect is not as expected and this indicates that further interpretation is necessary. In *total reading time* a significant difference was found between high and low causal strength, complying with prediction 2b. Yet in prediction 2b is to be affirmed, the absence of an effect in the other reading time measures must be explained.

The main effect for *context* in this region is as expected since in the no-context condition, region 1 was the first sentence the participants read, and it is known that readers are a little bit slower in reading the first sentence of a text (see amongst others, Rayner et al. 1995).

At region 2 a main effect for *causal strength* was found in all of the four reading time measures: For *first pass reading time* ($\chi 2$ [2]= 6.1121; p< .05) the reading times in the high condition were significantly shorter than those in the low condition (β= .069, SE= .029, p< .001). No significant difference was found between the reading times in the high and intermediate, or low and intermediate condition (β= .053, SE= .029, p= .071; β= .016, SE= .029, p= .589). For *right bounded reading time* ($\chi 2$ [2]= 13.86; p< .001) the reading times differed significantly between the high and low conditions (high < low: β=.083, SE= .022, p< .001). For *regression path duration* ($\chi 2$ [2]= 13.412, p< .01) the picture is similar with a significant difference between these two conditions (high< low: β= .095, SE=.026, p< .001). Finally, for *total reading time* ($\chi 2$ [2]= 37.294, p< .001) significant differences were found between all three conditions (high< low: β= .169, SE= .028, p< .001; high < intermediate: β= .070, SE= .027, p< .05; intermediate < low: β= .099, SE= .028, p< .001). At region 2 no effect for *context* was found is any of the reading time measures, also interaction effects between the two factors were absent.

The predictions for *causal strength* in 1a are not entirely borne out for the measures *first pass reading time, right bounded reading time* and *regression path duration*. The processing difference between high and low causal strength shows that readers do take into account causal strength as a factor, but the lack of a processing difference between intermediate and low causal strength, and between high and intermediate causal strength, suggest that readers might apply different strategies and have different expectations based on the semantic information given by the verb in the S1. On the other hand, in *total reading time*, the processing difference between the three conditions does show that causal strength affects integrating the S2 with the S1 in a different way. The lack of main effect of *context*, or an interaction effect between *context* and *causal strength* is apparent, and can be interpreted entirely in favour of the CbD hypothesis, since it supports prediction 4a and goes against prediction 4b.

The most striking outcome for region 2 is that the intermediate condition sits in between high and low causal strength, but does not differ significantly from either. This result is very uninformative and in order to be able to look more closely at the effect of *causal strength* in this region, I have split the region in two sub regions: region 2a and region 2b. Region 2a consists of the determiner phrase that is the beginning of S2. Region 2b consists of the verb phrase at the end of S2 (see example (10)).

10 Johan is erg agressief.
/$_1$ Hij slaat een portier. /$_{2a}$ De portier /$_{2b}$ raakt gewond./
/$_3$ Johan probeert het goed te praten /$_4$ en gaat dan vlug naar huis./

This division is relevant since at region 2a, the reader cannot yet infer the causal coherence relation because not all information is available. This means that processing differences that are found here are due to the manipulation of *causal strength* in the S1, and reflect the degree of causal strength. In region 2b, all information about the coherence relation is available and the full causal link can be constructed. Processing differences observed in this region are due to the degree of causal relatedness between S1 and S2.

At region 2a, a main effect for *causal strength* is found for all four reading time measures (*first pass*: $\chi 2$ [1]= 9,70; p< .01 ; *right bounded*: $\chi 2$ [1]= 15,08; p< .001 ; *regression path*: $\chi 2$ [1]= 12,50; p< .01 ; *total time*: $\chi 2$ [1]= 18,54; p< .001). The reading times in the high and intermediate condition are significantly shorter than those in the low condition, while no significant difference is observed between high and intermediate (*first pass reading time*: high < low: β= .073, SE= .027, p< .01; intermediate < low: β= .071, SE= .027, p< .01; high = intermediate: β= .002, SE= .022, p> .1; *right bounded reading time*: high < low: β= .086, SE= .026, p< .01; intermediate < low: β= .091, SE= .026, p< .001; high = intermediate: β= .005, SE= .057, p> .1; *regression path duration*: high < low: β= .094, SE= .031, p< .01; intermediate < low: β= .094, SE= .031, p< .01; high = intermediate: β= .0001, SE= .060, p> .1; *total time*: high < low: β= .195, SE= .033, p< .001; intermediate < low: β= .154, SE= .033, p< .001; high = intermediate: β= .041, SE= .033, p> .1)

No main or interaction effect for context was observed. The results observed at the first part of the S2 clearly show no processing difference between high and intermediate causal strength, and a significant difference between the above and low causal strength. If this effect is interpreted as a spill-over effect of region 1, the observations go strongly against prediction 1b and can be interpreted in favour of the CbD hypothesis since predictions 1A and 2a are borne out.

At region 2b, a main effect for *causal strength* is observed (*first pass*: $\chi 2$ [1]= 12,66; p< .01 ; *Right bounded*: $\chi 2$ [1]= 9,20; p< .05 ; *regression path*: $\chi 2$ [1]= 7,34; p< .05 ; *total time*: $\chi 2$ [1]= 24,22; p< .001).

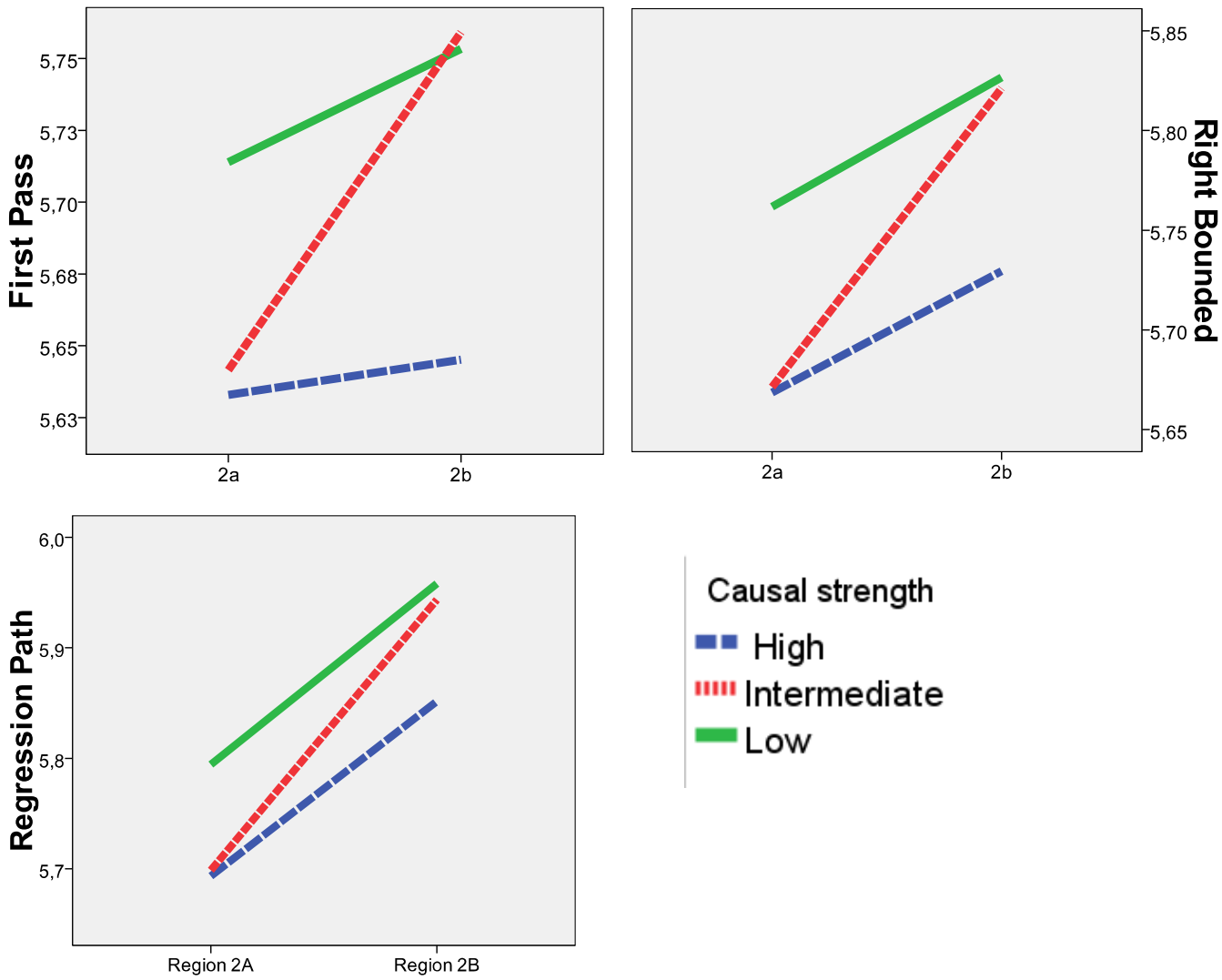

Figure 2: Mean (log) reading times (ms) for region 2a and 2b, per causal strength condition, for first pass, right bounded, and regression path.

This effect differs from that at region 2a, in that the reading times in the intermediate and low condition are grouped together, and are significantly longer than those in the high condition (*first pass reading time*: high < low: β= .105, SE= .035, p< .01; high < intermediate: β= .112, SE= .046, p> .01; intermediate = low: β= .008, SE= .065, p> .1; *right bounded reading time*: high < low: β= .094, SE= .035, p< .01; high < intermediate: β= .091, SE= .035, p< .01; intermediate = low: β= .003, SE= .070, p> .1; *regression path duration*: high < low: β= .103, SE= .042, p< .05; high < intermediate: β= .090, SE= .041, p< .05; intermediate = low: β= .013, SE= .094, p> .1; *total time*: high < low: β= .162, SE= .038, p< .001; high < intermediate: β= .104, SE= .038, p< .01; intermediate = low: β= .059, SE= .038, p> .1).

Again no effects for the factor *context* are found. These results indicate that high causal strength is processed significantly faster than both intermediate and low causal strength at the point in processing where causal inferences are expected to be made. Figure 2 gives the reading times for region 2a and 2b, per condition for three reading time measures: *first pass reading time, right bounded reading time,* and *regression path duration*. The pattern is the same for all three reading time measures.

At region 3, a main effect for *context* is found in all four reading time measures: the reading times in the no-context condition were significantly longer than those in the context condition (*first pass reading time*: χ2 [1]= 9.477, p< .01; *right bounded reading time*: χ2 [1]= 11.479, p< .001; *regression path duration*: χ2 [1]= 6.81, p< .01; *total reading time*: χ2 [1]= 11.757, p<.001). This main effect is modulated by an interaction effect between *context* and *causal strength* in *first pass reading time* (χ2 [2]= 6,82; p< .05): no effect of *context* is observed in the high causal strength condition (β= .029, SE= .024, p< .01; p> .1) while in intermediate and low causal strength the reading times in the context condition are significantly shorter than in the no-context condition (intermediate: β= .095, SE= .048, p< .05; low: β= .119, SE= .048, p< .05). The result on this interaction is that the reading times for high causal strength in the no-context condition are significantly shorter than those in the other two causal strength conditions (χ2 [2]= 9,122; p< .01: high:no-context < intermediate:no-context: β= .088, SE= .034, p< .01; high:no-context < low:no-context: β= .099, SE= .034, p< .01), while no difference is observed for degree of causal strength in the

context condition (χ2 [2]= 0,414; p> .1). Figure 3 presents the interaction effects found in this region.

An main effect of *causal strength* is present in the reading time measures that reflect later processes: For *Right bounded reading time* (χ2 [2]= 6.918, p< .05) the reading times in the low condition were significantly longer than in the high condition (high < low: β= .061, SE= .024, p< .01) while no difference was observed between high and intermediate, and intermediate and low causal strength. Also, for *Regression path duration* (χ2 [2]= 8.683, p< .05) the reading times in the high condition were significantly shorter that those in the low condition (β= .081, SE= .028, p< .01). For *total reading time* no difference between the high and intermediate condition was found, but they both differed from the low condition (high < low: β= .130, SE= .0298, p< .001; intermediate < low: β= .094, SE= .028 p< .001).

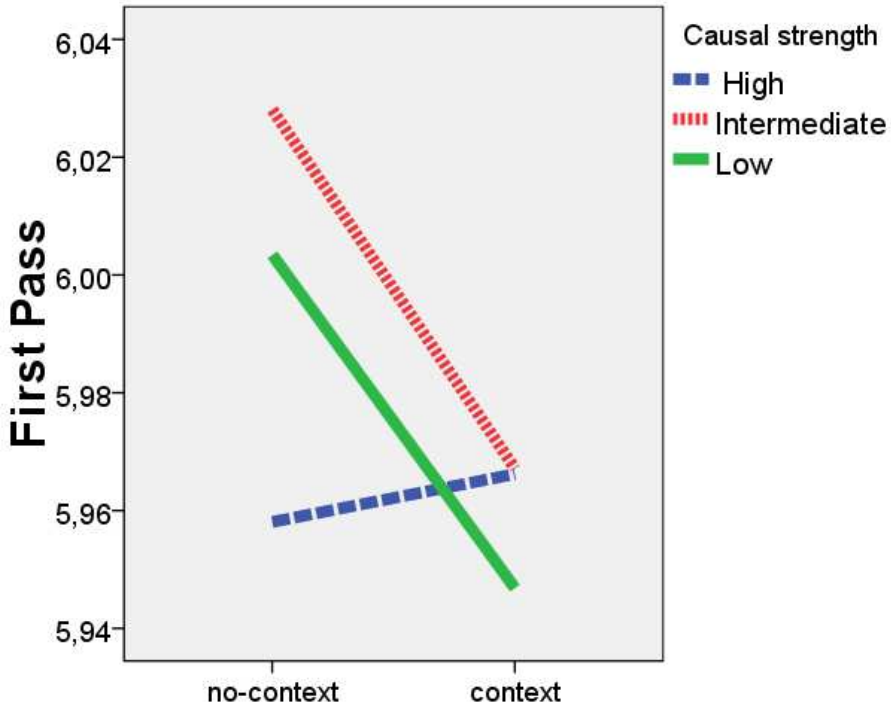

Figure 3: Mean (log) reading times (ms) for region 3, in First pass: the effect for context for each level of causal strength.

In stark contrast to region 2, region 3 shows processing differences due to *context* in all four reading time measures[4]. Since region 3 is where the reader integrates the S3 with the S1 and S2, it seems likely that context helps this integration process.

The other important observation in this region is the division between high and intermediate causal strength on the one side, and low causal strength on the other. Since the function of this region was to test the representation of the two preceding segments, this result can be interpreted as a difference in representation between the causal strength levels: high = intermediate ≠ low, as predicted in 3a.

At region 4 no effects were found in *first pass reading time*. In *right bounded reading time* as well as in *total reading time*, main effects for both *causal strength* ($\chi 2$ [2]= 6.918, p< .05; $\chi 2$ [2]= 23.538, p< .001) and *context* (context< no-context: $\chi 2$ [1]= 11.479, p< .001; $\chi 2$ [1]= 11.757, p< .001) were found. Both in *right bounded reading time* and *total reading time,* the reading times in the high causal strength condition are shorter than those in the low condition (high < low: β= .084, SE= .030p< .01; β= .082, SE= .029, p< .01 resp.).

In *regression path duration* no main effect for *context* was found, but for *causal strength* ($\chi 2$ [2]= 40.384, p< .001) both the high and intermediate condition were read faster than those in low condition (high < low: β= .239, SE= .039, p< .001; intermediate < low: β= .175, SE= .039, p< . 001). No interaction effects between *context* and *causal strength* were found in any of the reading time measures. Region 4 contains the last clause of the text, and wrap up effects are expected here. The absence of effects in *first pass reading time* shows that at the end of the text, the direct processing effects of the factors introduced at the beginning of the text have faded. In the reading time measures that include reread times and show later processes, differences in processing are still observed.

[4] Note that in *regression path duration*, despite the increased chance of longer reading times in the context condition because the presence of more discourse to regress to, the reading times in the context condition are still significantly shorter.

9. General discussion

In the this paragraphs, I will discuss the effects of the two factors *context* and *causal strength* on inferencing causal coherence, integrating the segments and the representation directly after the segments are processed, and how these processes interact with the possibility of the default assumption of causal coherence. Table 2 gives an overview of the most important results that I observed. Below I have repeated the research questions from section 6 and answered them based on the presented results. In the sections that follow I will discuss the results and the possible interpretations more extensively.

Table 2: Summary of the results (< indicates significant shorter reading time, = indicates no significant difference).

Effect	Emergence	Spill over to	Process
context < no-context	region 1 (FP, RB, RP, TT)		
context < no-context	region 3 (RB, RP, TT)	region 4	integration S3
no-context: high< interm.=low	region 3 (FP)		integration S3
intermediate < high	region 1 (RB, RP, TT)		inferencing
high = intermediate < low	region 2a (FP, RB, RP, TT)		integration S2
high < intermediate = low	region 2b (FP, RB, RP, TT)		inferencing
high (= intermediate) < low	region 3 (RB, RP, (TT))	region 4	representation

Research questions:

1. a. Does *context* affect the assumption of causal coherence? If so, at what point in processing?
   b. What is the effect of *context* on inferring and processing causal coherence?
   c. What is the effect of *context* on the representation of an implicitly marked coherence relation?

2. a. Does *causal strength* affect the assumption of causal coherence? If so, at what point in processing?
   b. What is the effect of *causal strength* on inferring and processing causal coherence?
   c. What is the effect of *causal strength* on the representation of an implicitly marked causal coherence relation?

Research question 1 can be answered by looking at regions 1 and 3. In both regions the presence of a context speeds up the reading process. The effect is region 1 can be explained when we take into account that readers speed up a little after they have began reading. In case on the no-context condition, region 1 was at the beginning of the item, what affected the reading times. The effect observed at region 3, where the representation of the S1 and S2 was tested can be interpreted as that the presence of a context helps the reader to integrate the S1 and S2 sentence with the S3 into a coherent discourse, since the context sentence gives an additional reason for the event in the S3. This can be interpreted as a positive answer to question 1b. The absence of a main effect for context at region 2ab and interaction effect between context and causal strength at regions 1 and 2ab are important observations with respect to answering questions 1a and 1aa. These results suggest that the context does not affect the assumption of causal coherence, inferencing causal links or the construction of the causal coherence relation between the S1 and S2. Below I will discuss these interpretations, and what they mean for the CbD hypothesis more extensively.

The observations in table (2) with respect to research question 2, which looks into the effect of causal strength, need some additional discussion about where the different processes (inferencing, integration, etc.) take place. This will be addressed below, in the section about causal strength. A more general interpretation of the observations is that it is apparent that causal strength influences inferring and processing causal coherence, both in the S1 and the S2. The observed effect in *total reading time* at region 3 can be interpreted as that both in high and intermediate causal strength conditions, the reader has constructed a causal coherence relation between S1 and S2, whereas in the low causal strength condition this does not seem to be the case. This interpretation answers question 2a. Based on this, the effects in region 1, where the low causal strength condition yields longer reading times than the intermediate causal strength condition can be interpreted with respect to the assumption of causal coherence: in the low causal strength condition, the assumption of causal coherence is dropped in the S1, answering question 2aa. Question 2b can be answered when we look at region 3 where the representation of the relation between S1 and S3 is tested, and the lack of difference between the high and intermediate causal strength conditions. No difference is found between the conditions with respect to the

representation of causal coherence. The observed effects between these two conditions at regions 1 and 2ab will be discussed in the section about causal strength.

Causality-by-Default and context

There are two interesting observations with respect to the factor *context*: the absence of an effect in the S2, and the presence of an effect in the S3 and the closing sentence.

First off, the complete absence of an effect for context in the S2 (regions 2a and 2b) indicates processing coherence itself is not directly affected by the presence of absence of context. This interpretation of the observations goes against the idea that the addition of a context sentence should lead to easier inferencing of causal links in order to construct a causal coherence relation , and therefore shorter reading times in the S2 for at least the intermediate condition, where inferences were likely to be made, and were necessary to construct causal coherence. That no difference is found at the point where causal coherence is inferred can be seen as support for the CbD hypothesis: the assumption of causality is not affected by a context that generates causal expectation, and must therefore be in place with or without the presence of a context.

Alternatively, it could be argued that making the inferences was so easy in the high and intermediate causal strength conditions, and so hard in the low causal strength condition, that the context was not relevant or strong enough to affect the inferencing process. This seems unlikely though, since the difference in causal strength between the high and intermediate conditions, as well as the effect of context is the low condition, have been shown to be significant in the pre-test.

Whether context can be seen to influence the integration of the S2 with the S1 clause depends on at where in the discourse the reader starts integrating the two clauses. Li (2010) observed that although integration seems to happen incrementally in the presence of a connective, when coherence marking is left implicit the moment of integration was delayed. This difference between marked and unmarked coherence can be explained by assuming that in the latter case, readers need to process more information due to the absence of an overt coherence signal. It could be argued that in case of low causal strength and no context information, readers may postpone the integration of the two clauses S1 and S2 until more information was available in the following clause, the S3, resulting in a context effect in this

sentence. Speculatively, this interpretation of postponing is backed up by the interaction effect observed in *First pass reading time* at region 3 (S3), where in the no-context condition, the reading time difference between high and intermediate causal strength that arises at the end of the S2 sentence, is continued into the S3, while in the context condition, no difference between high and intermediate causal strength is observed in the S3.

The observation with respect to context found in S3, where the presence of context resulted in shorter reading times overall, could also be interpreted with respect to the integration of the representation of the S1 and S2 into the discourse. The function of the S3 sentence was to test the representation constructed on the basis of the preceding discourse. The observed results indicate that overall, it is easier to construct a representation of a discourse that includes the context sentence and integrate this with the S3 test sentence, than when the representation relies only on the S1 and S2 sentence. Additionally, for the reading time measures that show later processing, the effect of context proceeds into region 4, which is the region where the reader has the opportunity for text wrap up, and rounds of the discourse as a short story. The continuation of the context effect up to the end of the discourse shows that, even though context might not affect initial processing or the integration of two causally related discourse segments, it does affect the overall integration and representation of the complete discourse.

The effect of context in this experiment can be summarized as truly a global effect that does not influence local causal coherence processing, but does affect the overall representation of a discourse. Based on the interpretations above, I can conclude that the assumption of causality as a default coherence interpretation is supported by the lack of effect for context during processing of the coherence relation, which is in line with prediction 4a.

Causality-by-Default and causal strength

First I will discuss the effect of causal strength with respect to the representation of the causal coherence relation. The function of the S3 sentence was to test the coherence relation and the representation of this relation constructed from the two preceding sentences, S1 and S2. The results (see table (2) indicate that both in the high and in the intermediate causal strength condition a causal representation is constructed, since no reading time difference were found between these two conditions in all reading time measures. When causal strength is low, the observed longer reading times can be interpreted as representing a mismatch between the constructed representation and the content of the S3 sentence. Since the S3 sentence presupposes at least an event-sequence between the S1 and S2, a likely interpretation is that this link is not constructed in the low causal strength condition. The lack of a difference in reading times between high and intermediate causal strength for the reading time measures that reflect later processes shows that for both degrees of causal strength causal coherence is inferred at this point in processing[5]. On the other hand, the observed results indicate that in case of low causal strength, the degree of causality is too low and no causal coherence is interpreted.

The spill-over of the effect of the difference between high and low causal strength into region 4 indicates that also when readers have obtained all the information available, they have difficulty constructing a causal link in case of low causal strength. A likely interpretation of this persisting effect is that there is a limit on the minimal level of causal strength necessary when constructing causal coherence. This interpretation can be seen as a support for the CbD hypothesis, in that when a minimal degree of causality is present (both in high and intermediate causal strength), a similar causally coherent representation is constructed. The lack of causal coherence in the low causal strength condition could be seen as problematic for the CbD hypothesis, since in this condition, like for example in the low causal relatedness condition in the Myers et al experiment, a complex and elaborate causal link between S1 and S2 is possible.

The observed effects of *causal strength* with respect to inferring and processing causal coherence indicate that there is a direct link between causal strength and causal coherence.

[5] This interpretation is possible since the high causal strength condition is taken as a baseline for which I presuppose that causal coherence is inferred.

The difference in reading times between the high and low causal strength conditions at the beginning of the S2 (region 2a, see table (2)) shows that the cause sentence affects processing right away. Since this part of the sentence only presents the subject of the S2 but not the event that the second sentence will be about (the consequence of the S1), an explanation for the reading time difference observed at this region must necessarily be found in the preceding discourse. The degree of causal strength in the S1 sentence must be interpreted as the source for this difference at the start of S2, where the reader is predicted to start integrating the two sentences. On the one hand, these observed results support the Causality-by-Default hypothesis: two levels of causal strength, high and intermediate, affect integration in similar way, and no three-way division for causal strength, where high causal strength would result in faster processing than intermediate causal strength, and intermediate would be faster than low causal strength was not found. On the other hand, the CbD hypothesis predicts an initial assumption of causal coherence, which could be argued to lead to similar integration over the three conditions, since the full coherence relation is not visible at this point. How then could the reading times in the low causal strength condition be longer than in the other two conditions? Since CbD explicitly includes sentence characteristics as a source of possible influence, the most viable explanation is that causal strength of the verb in S1 (opposed to the causal relatedness between the two verbs in S1 and S2) is indeed a characteristic that influences the assumption of causal coherence.
An interpretation of these results could be that in case of low causal strength, there is no support to maintain the assumption of causal coherence (i.e. no causal strength). This leads to the tentative conclusion for now that it is the absence or presence of causal strength that affects whether the assumption of CbD is dropped or not, but not the degree of causal strength (high vs. intermediate).

At the end of the S2 consequence sentence, after the verb phrase (region 2b), the pattern for the reading times in the intermediate causal strength condition is reversed with respect to the previous region: suddenly the reading times in the intermediate condition slow down when compared to the high causal strength condition and are comparable to those in the low condition (see table (2) and figure (2)). This could be seen as problematic for the CbD hypothesis, since I just concluded that only the absence of causal strength affects the

assumption of causality. And since We have seen that, based on the results in the S3, an intermediate level of causal strength leads to the construction of causal coherence, why would a difference arise at this point in processing?

An explanation for this effect at region 2b can be found in the results observed for the preceding sentence S1 (region 1, see table (2)). The S1 sentence in the intermediate causal strength condition was read faster than in the high and low causal strength conditions. Even though the observed effect in this sentence is somewhat modulated due to the different verb phrases that create the levels of causal strength, no difference in word count was found. From the Causality-by-Default point of view this difference seems unexplainable: since the causal assumption for high and intermediate should at least be equal, and therefore processing times are not expected to differentiate. At the same time, based on the results at the beginning of the second sentence, the lack of difference between high and low is puzzling. It is here that Asher and Lascarides' SDRT distinction between causal laws and causal scripts can provide some clarification (Asher & Lascarides 2003). As described in section 3.2, Asher and Lascarides argue that there are (at least) two ways a causal relation can be constructed: through causal law, or based on scriptal knowledge of a specific situation. In case of the former, the causal relation at hand is comparable to the relation between *pushing* and *falling* in *Max pushed John. He fell.* There is a strong causal link that is based on indefeasible world knowledge. Roughly speaking, when a person is pushed, he either falls, or he does not, but not falling would be seen as denial of expectation since the purpose of pushing is to make someone/something move. In case of the latter, the causal relation is seen as less strong and is based on situational knowledge. This knowledge of what happens in specific situations, for example in: *Max fell. John helped him up*. helps to create coherence. The relation between falling and helping is exists not because of indefeasible world knowledge, but because we know from experience that helping someone is what one might do in the situation of a fall. Moving to processing theories, an explanation that for example Myers et al. (1987.) used for the difference in recall between strong and intermediate causal coherence, was that in the former case, the causal link was entailed based on world-knowledge, while in the latter, inferences related to the specific situation were necessary. This can be directly related to Asher and Lascarides' theoretical difference.

And now back to the reading time difference that was observed between high and intermediate causal strength conditions. Compare *Johan hits a doorman.* with *Johan threatens a doorman.* as examples of cause-sentences that differ in degree of causal strength, and where either a causal law, or situational knowledge applies. Ignoring the difference in sentence length between these examples, the longer reading times for *Johan hits a doorman.* can be explained quite logically. Since hitting somehow entails knowledge about the possibility of someone getting hurt, whereas threatening does not directly entail a logical consequence, the reading times difference is possibly due to this entailment. The high causal strength sentences lead to immediate entailments about possible consequences, causing longer processing times. In the intermediate condition, causality is still expected (threatening is expected to have a consequence) therefore the reader does not drop the assumption of causal coherence and slow down processing. But the causal link can only be inferred at the point where the complete situational knowledge can be accessed, at the end of the S2. This explanation then leads immediately to an explanation for the observed difference between high and intermediate at the end of the S2 (region 2b, see table (2)), where the verb phrase of the S2 provides the information about the complete causal relation. Since we know from the results in S3 and the final sentence in the discourse (regions 3 and 4) that in the intermediate causal strength condition a causal coherence relation is indeed constructed, the longer reading times at the end of the S2 when compared to the reading times in the high causal strength condition can be interpreted as inferencing effects. It is at this point in processing that the reader is able to infer the causal link necessary to construct coherence between the S1 and S2.

Although this explanation for the observed processing differences between high and intermediate causal strength is somewhat elaborate, it does not seem improbable and fits the observed affects well.

Now that the difference in observed effects between high and intermediate causal strength in region 1 and region 2b can be explained as a difference between the positions where causal inferences can be made, these effects, together with the lack of difference between these two conditions at region 2a strongly support the CbD hypothesis. Since in both the high and intermediate causal strength condition the reading times in region 2a are not

slowed down when compared to those in the low causal strength (of which we know from the results in the S3 (region 3) that no causal coherence relation was constructed).

Based on the observations and interpretations above, some cautious overall conclusions can be drawn about the CbD, causal strength and their interaction.

A low degree of causal strength can cause the assumption of causal coherence to be dropped, even before the S2 is encountered. When the assumption persists, a difference in degree of causal strength can lead to differences in processing: the position where causal inferences are made differ between high and intermediate causal strength levels. A high degree of causal strength results in a higher expectation of causal coherence, leading to immediate inferencing effects, whereas in case of an intermediate degree of causal coherence, the assumption of causality is maintained, but no direct inferences are made. Here it is the causal relatedness between the two segments that causes inferencing and results in the integration of the second clause. Also, I hope to have convincingly shown that the assumption of causal coherence as in Causality-by-default, originating in cognitive mechanisms, and the expectation of causality, originating in causal strength in the discourse, are two separate factors in discourse processing, since despite the absence of causal expectation (leading to immediate causal inferencing) for intermediate causal strength, the assumption of causality is still proven to be present at this causal strength level.

10. Remarks and suggestions

Besides the experimental results and conclusions with respect to the CbD hypothesis and interacting factors, this thesis has also tested a novel way of testing representation: with a S3 sentence that targets a specific discourse coherence relation. This method seemed to have worked, in that from the reading time differences, representational differences could be distilled. Yet a point that I made before about the interpretation of recall results remains: how do we interpret the results in the low causal strength condition? Can it really be said, based on the processing differences, that no causal links were constructed between S1 and S2 in the low causal strength condition? Ideally, these results should be compared to the reading times of the S3 sentence in the case where only a temporal relation between S1 and S2 is possible. Yet it seems a very complex task to construct a narrative in which this could be tested. One possibility is to compare reading times for items where a reader can choose between a (weak) causal coherence interpretation and an overtly marked temporal coherence. Since temporal marking is underspecified for causal coherence, the S3 sentence could be used to find out what coherence relation the reader has constructed. .

Another point that should be looked at in more detail is the testing of causal entailment in high causal strength conditions, and in what way this differs from inferencing in case of intermediate causal strength, since this difference featured prominently in the above discussion. Causal strength affects processing in that high causal strength seems to lead to more immediate entailments due to causal expectation, where in the case of intermediate causal strength inferences necessary for causal link are made later in the consequence clause. In order to test whether this difference truly exists in such a way as suggested above, it should be experimentally tested. A possible method to test whether the positions where inferences are made differ between the degrees of causal strength could be through *denial of expectation*. When in high causal strength condition, causal expectations lead to inferences while processing the S1, a reading time difference should be observed when the S1 is followed by an S2 where these causal expectations are denied, when compared to the reading times in intermediate causal strength, since in this condition, inferences are based partly on the S2.

Additionally, in the experiment presented here, I have only looked at objective forward cause-consequence relations. It would be very interesting to see in what way the CbD hypothesis and the factors presented in this thesis affect other types of causal coherence relations, and whether the conclusions presented here also hold for other types of relations.

11. References

Albrecht, J.E. and Myers, J.E. (1995). The role of context in accessing distant information during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19: 1459-1468

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

Asher, N. and Lascarides, A. (1998) Bridging, *Journal of Semantics*, 15: 83-113, Oxford University Press.

Asher, N. and Lascarides, A. (1995) Lexical disambiguation in a discourse context. *Journal of Semantics,* 12: 69-108. Oxford University Press.

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

Caenepeel, M. (1991). Event structure versus discourse coherence. In Caenepeel, M., Delin, J., Oversteegen, L., and Sanders, J., editors, Proceedings of the DANDI Workshop on Discourse Coherence.

Cozijn, R. (2000). Integration and inference in understanding causal sentences. Unpublished doctoral dissertation, University of Tilburg, The Netherlands.

Haberlandt, K., and Bingham, G. (1982). The role of scripts in the comprehension and retention of texts. *Text*, 2, 29-46.

Graesser, A. C., Singer, M., and Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101: 371-395.

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

Keenan, J.M., Baillet, S.D., and Brown, P. (1984). The effects of causal cohesion on comprehension and memory. *Journal of Verbal Learning and Verbal Behavior*, 23: 115–126.

Kehler, A. (2002). *Coherence, Reference, and the Theory of Grammar*. Stanford, CA: CSLI Publications.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95: 163–182.

Kintsch, W. and Van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review,* 85: 363-394.

Kuperberg, G.R., Paczynksi, M. and Ditman, T. (2010b). Establishing causal coherence across sentences: an ERP study. *Journal of Cognitive Neuroscience*.

Landauer, T. K. and Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104: 211–40.

Lascarides, A. and Asher, N. (1993). Temporal interpretation, discourse relations and common-sense entailment. *Linguistics and Philosophy*, 16: 437-493.

Li, F. (2010). Causality in on-line discourse processing: What eye-tracking reveals about the role of causal relations and connectives, Unpublished MA-thesis, Utrecht University

Michotte, A. (1963). The perception of causality. (T.R. Miles & E. Miles, Trans.). London: Methuen. (English translation of Michotte, 1954).

Millis, K.K. and Just, M.A. (1994). The influence of connectives on sentence comprehension. *Journal of Memory and Language,* 33: 128-147.

Moeschler, J. (2001), How to infer temporal relations in discourse?, paper XIX AESLA Conference, May 2001.

Mulder, G. (2008). Understanding causal coherence relations. Doctoral dissertation, Utrecht University, The Netherlands.

Myers, J. L., O Brien, E. J., Albrecht, J. E., and Mason, R. A. (1994). Maintaining global coherence during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20: 876–886.

Myers, J.L., and Duffy, S.A. (1990). Causal inferences and text memory. *ThePsychology of Learning and Motivation*, 25: 159–173.

Myers, J.L., Shinjo, M., and Duffy, S.A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language*, 26: 453–465.

Noordman, L.G.M. and Vonk, W. (1998). Memory based processing in understanding causal information. *Discourse Processes* 26: 191-212.

Noordman, L.G.M. & Vonk, W. (1997). The different functions of a conjunction in constructing a representation of the discourse. In: J. Costermans & M. Fayol (eds.), Processing interclausal relationships. Studies in the production and comprehension of text, 75-93. Hillsdale, New Jersey: Lawrence Erlbaum.

Rayner, K., Raney, G. E., and Pollatsek, A. (1995). Eye movements and discourse processing. In: R. F. Lorch & E. J. O'Brien (eds.), Sources of coherence in reading, 9-36. Hillsdale, NJ: Erlbaum.

Sanders, T.J.M. (2005). Coherence, causality and cognitive complexity in discourse. In: M. Aurnague, M. Bras, A. Le Draoulec & L. Vieu (eds.), Proceedings/Acts SEM 05, First International Symposium on the exploration and modeling of meaning, 105 114.

Sanders, T.J.M. and Noordman, L.G.M. (2000). The role of coherence relations and their linguistic markers in text processing. Discourse Processes 29: 3760.

Sanders, T., Spooren, W. and Noordman, L. (1993). Coherence relations in a cognitive theory of discourse representation. *Cognitive Linguistics* 4-2: 93-133.

Sanders, T., Spooren, W. and Noordman, L. (1992). Toward a taxonomy of coherence relations. *Discourse Processes* 15: 135.

Trabasso, T., and van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language* 24: 912 630.

Traxler, M.J., Bybee, M.D. and Pickering, M.J. (1997). Processing causal and diagnostic statements in discourse. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2: 88-101.

Van Dijk, T. A., and Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Van den Broek, P.W. (1994). Comprehension and memory of narrative texts inferences and coherence. In M.A. Gernsbacher (ed.), Handbook of psycholinguistic. 539–588. San Diego, CA: Academic Press.

Van den Broek, P. W. (1990). Causal inferences and the comprehension of narrative text. In: A.C. Graesser and G.H.Bower (eds.), The psychology of learning and motivation, 275-196. New York: Academic.

Wolfe, M.B.W., Magliano, J.P., and Larsen, B. (2005). Causal and Semantic Relatedness in Discourse Understanding and Representation. *Discourse Processes* 39: 165 187.