On the online effects of subjectivity encoded in causal connectives

Fang Li,* Willem M. Mak,** Jacqueline Evers-Vermeul** and Ted J. M. Sanders**
* Beijing International Studies University / ** Utrecht University

Causal relations between sentences differ in terms of subjectivity: they can be objective (based on facts) or subjective (based on reasoning). Subjective relations lead to longer reading times than objective relations. Causal connectives differ in the degree to which they encode this subjectivity. The Chinese connectives *kejian* 'so' and *yin'er* 'so' specify a high and low degree of subjectivity, respectively, whereas *suoyi* 'so' is underspecified for subjectivity. In an eye-tracking experiment we compare the effect of the specificity of these connectives in subjective and objective relations. In objective relations, the specificity of the connective has no effect on reading times. In subjective relations, reading times are shorter in sentences with the specified connective *kejian* than in sentences with the underspecified connective *suoyi*. These results suggest that readers prefer to interpret a relation as objective. Computing subjective relations requires extra processing time, which is diminished when the connective encodes the subjectivity.

Keywords: causal connectives, subjectivity, discourse processing, eye-tracking

1. Introduction[1]

1.1 Causal coherence relations and subjectivity

It is widely accepted that discourse is more than merely a collection of utterances: it shows coherence. This coherence can be characterized in terms of 'coherence relations' (also called 'rhetorical relations'; see Knott & Dale, 1994; Mann & Thompson,

1. This chapter is an adapted version of Chapter 5 of the first author's Ph.D. dissertation (Li, 2014). The research was enabled by a Huygens Scholarship from Nuffic, by Education Reform Project of Beijing International Studies University (project number 200232), and by Special Item Fund of Beijing International Studies University, 391002/002.

Review of Cognitive Linguistics 15:1 (2017), **34–57**. DOI 10.1075/rcl.15.1.02li
ISSN 1877-9751 / E-ISSN 1877-976X

1986, 1988; Sanders, Spooren, & Noordman, 1992, 1993). Coherence relations (e.g., cause-consequence, additive, and concessive relations) are conceptual relations that hold between the propositional content of two discourse segments (Sanders et al., 1993), which can (but need not) be explicitly marked by connectives such as *because*, *and*, and *but*. Causal coherence relations, such as the ones in (1)–(3), are typical examples of coherence relations.

(1) Temperatures were below minus ten degrees for more than a month. Many kingfishers died last year.

(2) The neighbors' car is not in the driveway, so they probably left.

(3) What are you doing tonight, because there is a good movie on?

In order to understand these text fragments, readers need to interpret the coherence relations contained in them. The coherence relations expressed in (1), (2), and (3) have something in common: a causal relation can be established between an antecedent segment P and a consequent segment Q. Still, several aspects can be distinguished. One apparent distinction concerns the use of connectives. In (1), no connective is used to explicitly mark the causal relation, whereas (2) and (3) respectively contain the causal connectives *so* and *because*.

Other distinctions can be made on the basis of conceptual differences between causal relations; for example, in terms of causal domains. According to Sweetser (1990), there are three causal domains: the content domain, the epistemic domain, and the speech-act domain. Example (1) illustrates causal relations in the content domain; that is, the "real-world causality" holding between observable facts/events (Sweetser, 1990, p. 77): the observable fact P that *the temperatures were below minus ten degrees for more than a month led* to the other observable fact Q that *many kingfishers died*. Sentence (2) is an instance of a causal relation in the epistemic domain, in which the speaker's knowledge P (*the neighbors' car is not in the driveway*) is involved as the basis for a logical conclusion Q (*they probably left*). Sentence (3) typifies the speech-act domain in which one clause provides justification for the speech act in the other clause. The speaker asks what the addressee is doing that night and justifies this question (a type of speech act) with the reason clause *there is a good movie on*, i.e., "I am asking you the question Q because P". These conceptual differences between (1) on the one hand, and (2) and (3) on the other are also referred to as external versus internal (Halliday & Hasan, 1976; Martin, 1992), subject matter versus presentational (Mann & Thompson, 1988), and semantic versus pragmatic (Moeschler, 1989; Sanders, 1997; Sanders et al., 1992; van Dijk, 1979).

Recently, a cognitively-oriented approach has started to view these differences from the perspective of 'subjectivity', taken as the degree of 'speaker involvement' (Degand & Pander Maat, 2003; Pander Maat & Sanders, 2001; Stukker & Sanders,

2012; Stukker, Sanders, & Verhagen, 2008). Across languages, links have been established between the notion of subjectivity and causal domains: it is assumed that different domains of causality reflect different degrees of subjectivity (Degand & Pander Maat, 2003; Stukker & Sanders, 2012). Causal relations in the content domain, such as (1), are assumed to be objective, because the speaker's role is limited to reporting or describing causality that holds between observable facts in the outside world. By contrast, causal relations in the epistemic domain, as the one in (2), are assumed to be subjective, because the speaker is actively involved in grounding a belief or claim on an appropriate argument (be it a piece of real-world evidence, some shared knowledge, or a certain conclusion established in the preceding context). Causal relations in the speech-act domain are also subjective because the speaker is highly involved in providing justifications for the speech act (e.g., raising a question, making a promise, or issuing a command) that she performs.

In the current study, we distinguish between subjective and objective causal relations. In particular, epistemic relations and content relations are selected as the targets of study because they typically represent the subjective causal relation and the objective causal relation, respectively. Although the speech-act domain is also associated with a high degree of subjectivity, it rarely occurs in written discourse (Spooren, Sanders, Huiskes, & Degand, 2010; Li, Evers-Vermeul, & Sanders, 2013). The present experiment uses written materials, so we do not include the speech-act domain as a target of study.

1.2 Causal connectives and subjectivity in discourse processing

Causal connectives are explicit linguistic markers of causal relations between adjacent segments (Murray, 1997; Stukker & Sanders, 2012). In psycholinguistics, connectives are considered as operating instructions for interpretation: connectives instruct the reader to relate the content of the connected segments in a specific type of relationship during online processing (Britton, 1994; Koornneef & Sanders, 2013; Mak & Sanders, 2013). In the terms of Noordman and Vonk (1997), a connective functions as an integration device, facilitating the construction of two connected text segments into a single representation during online processing (see also Kintsch, 1988; Millis & Just, 1994). On the basis of these ideas, it can be predicted that the presence of an appropriately-used causal connective speeds up the processing of the immediately following discourse segment. Indeed, several online reading studies have shown that participants spend less time reading a clause when it is introduced by a causal connective than when the connective is absent (see Cozijn, 2000; Haberlandt, 1982; Millis & Just, 1994).

Causal connectives not only instruct the reader to establish a causal link between two clauses, but they also often provide information with respect to the degree of subjectivity of the relation that they mark. Studies on languages such as French, German, Dutch, Polish, and Mandarin Chinese have consistently observed such subjectivity encoding in causal connectives (see Dancygier, 2009; Evers-Vermeul, Degand, Fagard, & Mortier, 2011; Keller, 1995; Li et al., 2013; Pander Maat & Sanders, 2001; Pit, 2003; Zufferey, 2012). For example, in Dutch *want* 'because' is encoded with a high degree of subjectivity whereas *omdat* 'because' is encoded with a low degree of subjectivity (i.e., it involves real-world causality) (Degand & Pander Maat, 2003; Pit, 2003; Sanders, Sanders, & Sweetser, 2012; Sanders & Spooren, 2009; Verhagen, 2005). Both types can be referred to as 'specified causal connectives', as they specify the degree of subjectivity the relation is involved with.

In other languages, there are certain causal connectives that have a more general usage, such as the English *because* (see Couper-Kuhlen, 1996; Ford, 1993; Knott & Sanders, 1998; Sweetser, 1990). According to Knott and Sanders (1998), *because* is undefined for the parameter of 'source of coherence'; i.e., with respect to the distinction between semantic/objective and pragmatic/subjective causality. We will refer to these causal connectives of general usage as 'underspecified causal connectives'.

Previous reading-time experiments have shown that the subjectivity information encoded in connectives affects online processing: in general, causal relations or connectives that are associated with different degrees of subjectivity show different patterns during online processing (Traxler, Bybee, & Pickering, 1997a; Traxler, Sanford, Aked, & Moxey, 1997b). These studies focus on the processing of content versus epistemic causal relations, both marked with the underspecified causal connective *because*. Traxler et al. (1997a) used English sentence-pairs such as (4) and (5) in their study.

(4) The goalkeeper won the game **because** the other team didn't score any goals.

(5) The goalkeeper knew how to play the game **because** the other team didn't score any goals.

Within each pair, one sentence expresses a causal relation in the content domain, as in (4), and the other sentence expresses a causal relation in the epistemic domain, as in (5). In both cases, *because* is used to connect the consequent clause and the antecedent clause. As an underspecified causal connective, *because* provides no information as to whether the unfolding causal relation is likely to be subjective or objective. In this case, the predicate (*didn't score*) is the first possible position where the necessary information becomes available for readers to figure out that the sentence in (5) expresses a causal relation with a high degree of subjectivity. The results from Traxler et al. (1997a) show that, at this position, a processing

delay occurred for epistemic compared to content relations. These longer reading times reflect the effect of subjectivity on discourse processing. A higher degree of subjectivity results in a higher processing effort.

This processing finding conforms to the so-called 'subjective complexity hypothesis' proposed by Sanders (2005): a subjective relation is cognitively more complex than an objective relation. This hypothesis is corroborated by different types of evidence. First, epistemic reasoning is often based on real-world content relations (Noordman & de Blijzer, 2000; Pander Maat & Degand, 2001; Sanders, 2005; Sanders & Spooren, 2015). Second, we know that diachronically, subjective causal relations are derived from objective causal relations (Sweetser, 1990; Traugott, 1995). Third, a consistent finding in developmental studies is that content relations are acquired before epistemic relations (Evers-Vermeul, 2005; Evers-Vermeul & Sanders, 2011; Spooren & Sanders, 2008; van Veen, 2011). This finding has been explained by the cumulative cognitive complexity approach (Bloom et al., 1980; Evers-Vermeul & Sanders, 2009), which claims that complex relations are acquired later than simple ones. Subjective epistemic relations are cognitively more complex than objective content relations, which is why the former take longer to acquire than the latter.

Traxler et al. (1997a) used underspecified causal connectives. What if the causal connectives already specify the degree of subjectivity of the causal relation they mark? Canestrelli, Mak, and Sanders (2013) compared reading times of Dutch content and epistemic causal relations, respectively marked by *omdat* 'because' and *want* 'because', basing their experimental materials on those in Traxler et al. (1997a). Again, the eye-tracking results indicated a processing delay for epistemic relations: the *want* condition generated longer processing times than the *omdat* condition. Importantly, this processing delay associated with epistemic relations was observed earlier in the Dutch materials than in the English ones: in the *want* condition, there was a delay before the predicate region (Canestrelli et al., 2013), that is before the position where the reader can derive the subjectivity of the relation on the basis of the content of the sentence. Hence, this effect is due to the subjectivity encoded in the connective itself. Moreover, there is no effect on the predicate region, showing that the extra processing effort at that region is not necessary when the connective already shows that the relation is subjective.

Taken together, previous experimental research on Dutch and English has shown that objective causal relations are processed faster than subjective causal relations during reading. Furthermore, the processing asymmetry between objective and subjective conditions occurs at different positions in the sentence, depending on the specificity of the connective with respect to subjectivity. With Dutch materials, the processing delay for subjective relations was observed at or immediately after the specified causal connective *want*, but it occurred at the predicate position

(at which the bulk of the two propositions P and Q were available) when English materials and the underspecified causal connective *because* were used.

However, the above claims about the processing consequences of the specificity of connectives (with respect to subjectivity) are abstracted between experiments and between languages. So far, no direct comparison has been made between the role of underspecified versus specified causal connectives, because studies looked at languages with either only an underspecified causal connective (i.e., *because*) or only specified causal connectives (i.e., *omdat* and *want*). In order to establish the effect of specificity, a direct comparison between a specific connective and an underspecified connective within the same language is necessary. Therefore, we set up an eye-tracking experiment with causally related clauses in Chinese. In this experiment, we compared specified and underspecified causal connectives in objective and subjective causal relations. In the next section, we discuss why we selected the Chinese language and we present the hypotheses for this study.

2. Current study: Connective selection and hypotheses

As shown in Table 1, Chinese possesses both specified causal connectives (including objective and subjective ones) and underspecified causal connectives. For English, no objective causal connective has been identified. Existing analyses suggest that in English the demarcation between subjective and objective categories is realized by cue phrases rather than connectives (Knott & Dale, 1994; Knott & Sanders, 1998):

Table 1. Overview of causal connectives in English, Dutch, and Mandarin Chinese[2]

Language	Marker	Objective	Subjective	Underspecified
English	Reason		*since*	*because*
	Result		*so*	
Dutch	Reason	*doordat*, *omdat*	*want*	
	Result	*daardoor, daarom*	*dus*	
Mandarin Chinese	Reason	*youyu*, sentence-initial *yinwei*	*jiran*, intersentential *yinwei*	
	Result	*yin'er, yushi*	*keijan*	*suoyi, yinci*

2. The Chinese connective *yinwei* 'because' is equally felicitous in an antecedent clause that precedes the main clause (sentence-initial *yinwei*), and heading a post-posed antecedent clause (inter-sentential *yinwei*). The former usage is exemplified by the sentence "*Yinwei* my legs felt numb, I got out of the car to have a walk". The latter usage is exemplified by the sentence "Zhang San still loves Xiaoli, *yinwei* he has come back".

for that reason (objective) and *it follows that* (subjective). For Dutch, there seems to be no underspecified causal connective. For Chinese result connectives, there are instances of each of the three categories. Therefore, Chinese materials are used in the current experimental study.

2.1 *Kejian*, *suoyi*, and *yin'er*

We choose to use the Mandarin Chinese connectives *kejian*, *suoyi*, and *yin'er* (all translated as 'so/therefore') in the experiment. On the basis of a quantitative corpus-based study on the use of connectives in informative, narrative and argumentative texts, Li et al. (2013) have shown that *kejian* is highly subjective and *yin'er* is highly objective across genres, whereas *suoyi* is underspecified with respect to subjectivity. Similar observations about the usage of *kejian* and *yin'er* have been made by Xing (2001). The corpus fragments (6) – (9) exemplify typical uses of the three connectives. The usage of *kejian* is restricted to subjective causal relations such as (8): Li et al. (2013) found that 221 out of 225 cases were subjective epistemic relations, that the remaining four fragments were subjective speech-act relations, and that *kejian* is never used in content causal relations such as (6) and (7). *Yin'er* is typically used to mark content causal relations (62%) between facts in observable reality, such as (6). Hence, the specified connectives *yin'er* and *kejian* provide readers with instructions as to whether an objective respectively a subjective causal relation has to be established. The underspecified causal connective *suoyi* can be used in both objective and subjective relations, as (7) and (9) illustrate, and does not provide readers with instructions regarding the degree of subjectivity of the unfolding causal relation.

(6) *Zhe zhong lan meigui bei zhiru yi zhong neng ciji lan sesu chansheng de jiyin,* ***yin'er*** *huaban chengxian lanse.*
'This type of blue roses has been implanted with a hormone that can stimulate the production of blue pigments, **as a result**, the petals are blue.'

(7) *Nashi women dou zhu zai zhe yidai,* ***suoyi*** *women jingchang pengjian.*
'At that time we both lived in this area, **so** we often encountered each other.'

(8) *Dianshiji de huamian yanse te dan,* ***kejian*** *xianxiangguan yijing laohua.*
'The color of the picture on the TV is particularly light, **so** the kinescope has already been aging.'

(9) *Ta shou shang daile hao ji ge jiezhi,* ***suoyi*** *wo yiwei ta jiehun le.*
'He wore quite a few rings on the hand, **so** I thought he was married.'

Kejian consists of two morphemes: *ke* 'can' and *jian* 'see'. Both morphemes contribute to the subjectivity of *kejian*: *ke* 'can' concerns epistemic modality, and hence

involves someone's perspective on and assessment of the plausibility or truthfulness of a certain state of affairs; *jian* 'see' has been demonstrated to be a subjectivity indicator that encodes perspective (see Tao (2007) for evidence that *jian* indicates subjectivity in the context of existential/presentative constructions). For both *yin'er* and *suoyi*, however, the words themselves do not refer to someone's perspective.

Apart from the differences in terms of the degree of subjectivity encoded in them, *kejian*, *suoyi*, and *yin'er* are very similar in their usage. They are all used to introduce the consequent clause, so they are called 'result connectives'. They are all conjunctions. In terms of frequency, though, they are not the same. *Suoyi* is much more frequent than *kejian* and *yin'er.* According to the Lancaster Corpus of Mandarin Chinese (McEnery & Xiao, 2004), the frequency of *suoyi* is 3.55 per 10,000 words, and the frequency of *kejian* and *yin'er* is 0.63 and 1.50 per 10,000 words, respectively.

2.2 Hypotheses

As we have put forward in the introduction, the present study explores the way in which the degree of subjectivity encoded in causal connectives affects online discourse processing. The English version of one set of our test items, (10) through (13), will suffice here for the purpose of illustrating our hypotheses. A full explanation of the materials and the design will be given in the next section.

(10) My bike ran over a nail when I rode on it, ***yin'er*** the tire leaked.

(11) My bike ran over a nail when I rode on it, ***suoyi*** the tire leaked.

(12) The tire became flat immediately after I pumped it up, ***kejian*** the tire leaked.

(13) The tire became flat immediately after I pumped it up, ***suoyi*** the tire leaked.

Sentences (10) and (11) express a causal relation in the content domain (i.e., the fact that the bike ran over a nail led to the fact that the tire leaked). The wording is identical, except for the connective (i.e, *yin'er* vs. *suoyi*). Likewise, the wording in sentences (12) and (13) is identical except for the connective (i.e., *kejian* vs. *suoyi*). They express a causal relation in the epistemic domain (i.e., the speaker drew a conclusion "the tire leaked" on the basis of the fact that the tire deflated quickly).

First, we expect to replicate the results in Traxler et al. (1997a). We predict that, when both the content relation and the epistemic relation are marked with the underspecified connective *suoyi*, as in (11) vs. (13), a processing asymmetry will occur between these two conditions at the predicate of the second clause. More precisely, because subjective relations are more complex than objective ones (see Section 1.2), we expect a processing delay at the predicate of the epistemic cases.

Second, we expect differences in the processing of sentences such as (12) and (13). If *kejian* indeed instructs readers to construct a subjective causal relation, we expect a slowdown effect in cases like (12) immediately after the connective, parallel to the slowdown effect found in the Dutch experiments with the subjective connective *want* (Canestrelli et al., 2013). The underspecified causal connective *suoyi* provides no instruction in terms of subjectivity. Therefore, we assume that in cases like (13) the epistemic nature of the relation has to be constructed on the basis of the content of the clauses, probably at the predicate of the second clause (i.e., *leaked*), where the bulk of the propositions become available. Accordingly, we predict that at the predicate of the second clause, epistemic relations with *suoyi* are read slower than epistemic relations with *kejian*.

Third, for the comparison between the content conditions (10) and (11), we have no clear prediction. One possibility is that *yin'er* speeds up the processing of the second clause. Y*in'er* immediately instructs the reader to construct an objective causal relation, and in this way facilitate the processing of the actual content causal relation. Alternatively, no processing differences may arise between pairs such as (10) and (11). Traxler et al (1997b) claim that readers prefer to build the simplest discourse representation. If this claim holds true for sentences (10) and (11), readers will construct a content causal relation in every sentence with an objective or an underspecified causal connective, because these connectives do not instruct them to construct a more complex subjective relation, contrary to what would be the case in Example (12). In other words, if a content interpretation is the default interpretation because it is the simplest one, *yin'er* will not facilitate the processing of the content relation compared to *suoyi*.

3. Method

3.1 Participants

Forty-four undergraduate/postgraduate students from Utrecht University participated in the experiment (19 male, 25 female, mean age 27.8, age range 20–36), and were paid for their participation. All of them were native speakers of Chinese.

3.2 Materials and design

Preparing test items
Sixty sets of mini stories were constructed. Table 2 provides a sample set.

Table 2. Sample items for the experiment

Condition	Sample item
A. content relation & *yin'er*	*Meng Na yi nian lai baoshou weibing de zhemo,* ***yin'er*** *ta bi yiqian shou le bu shao. Ta shi liang ge haizi de muqin.* 'For a year Meng Na has been suffering from stomach trouble, ***yin'er*** she has become much thinner now than before. She has two children.'
B. content relation & *suoyi*	*Meng Na yi nian lai baoshou weibing de zhemo,* ***suoyi*** *ta bi yiqian shou le bu shao. Ta shi liang ge haizi de muqin.* 'For a year Meng Na has been suffering from stomach trouble, ***suoyi*** she has become much thinner now than before. She has two children.'
C. epistemic relation & *kejian*	*Meng Na na tiao kuzi xianzai xiande hen fei,* ***kejian*** *ta bi yiqian shou le bu shao. Ta shi liang ge haizi de muqin.* 'That (old) pair of trousers now look very baggy on Meng Na, ***kejian*** she has become much thinner now than before. She has two children.'
D. epistemic relation & *suoyi*	*Meng Na na tiao kuzi xianzai xiande hen fei,* ***suoyi*** *ta bi yiqian shou le bu shao. Ta shi liang ge haizi de muqin.* 'That (old) pair of trousers now look very baggy on Meng Na, ***suoyi*** she has become much thinner now than before. She has two children.'

There were four experimental conditions per set: (a) content causal relations marked with *yin'er*, (b) content causal relations marked with *suoyi*, (c) epistemic causal relations marked with *kejian*, and (d) epistemic causal relations marked with *suoyi*.

This design involves two critical manipulations. One manipulation concerned the relation type (content vs. epistemic). The other concerned the choice of connective. For content causal relations, the objective causal connective *yin'er* was used. For epistemic causal relations, we used the subjective causal connective *kejian*. As to the underspecified causal connective, we used *suoyi* in both types of relations. This selection of connectives was based on results from the corpus-based study (Li et al., 2013) introduced earlier.

The selection of test items

To make sure that the constructed causal relations were without ambiguity (i.e., they do not have a subjective interpretation and an objective interpretation that are on a par with each other), an online assessment questionnaire was conducted. The questionnaire consisted of objective and subjective paraphrases of the constructed causal relations. Objective paraphrases were in the form of "Q *shi yinwei* 'be because' P", and subjective paraphrases were in the form of "P, *zhe biaomin* 'this suggests' Q". The objective paraphrase "Q *shi yinwei* P" is appropriate only for sentences expressing objective causal relations. If a sentence is still acceptable after

adding *shi yinwei* between Q and P, then the sentence expresses a content causal relation; if not, the sentence should be interpreted differently (Shen, 2003). "P, *zhe biaoming* Q" is called a subjective paraphrase because all epistemic causal relations can be paraphrased in this way, whereas content causal relations cannot. The cue phrase *zhe biaoming* 'this suggests' requires that incoming information should contain one's ideas or conclusions, so it is fit only for epistemic causal relations.

Applying the two paraphrase patterns to the 60*2 constructed causal relations, 240 paraphrases (60*2 relations*2 paraphrases) were generated, and were then divided into four lists using a Latin square design. Sixty native speakers of Chinese, who were undergraduate students from Zhejiang University, took the questionnaire. They were asked to rate the appropriateness of the paraphrases on a five-point scale, with 5 representing "very acceptable", 1 "not acceptable at all", and 2–4 representing the intermediate degrees. Each participant rated only one list of paraphrases. Table 3 summarizes the item's means and the standard deviations for the four conditions.

Table 3. Mean, minimum and maximum ratings (and standard deviations) for the paraphrases of the constructed causal relations

Causal relation	Paraphrase	Mean 60 sets	Mean 40 sets	Min. 40 sets	Max. 40 sets
Objective	1. objective	4.07 (0.49)	4.13 (0.40)	3.24	4.80
	2. subjective	2.33 (0.57)	2.23 (0.46)	1.39	3.00
Subjective	3. objective	2.42 (0.62)	2.30 (0.55)	1.33	3.00
	4. subjective	4.09 (0.61)	4.30 (0.37)	3.20	4.89

We scrutinized the mean scores of each item, and excluded those items in which the correct conditions (i.e., 1 & 4) scored below 3.00, and those items in which the incorrect conditions (i.e., 2 & 3) scored above 3.00. We also excluded some items for which the mean scores of the correct conditions were approximately the same as the mean scores of the incorrect conditions. Following these criteria, we selected 40 sets of stories as test items for the experiment. A repeated measures ANOVA was then conducted on the item's means. We found a significant interaction between relation type and the type of paraphrase ($F(1, 39) = 980.66$, $p < .001$). Pairwise comparisons with Bonferroni adjustments indicated that neither the correct conditions 1&4 differed ($p = .36$), nor the incorrect conditions 2&3 ($p = 1.00$); meanwhile, the correct conditions scored significantly higher than the incorrect conditions (1 vs. 2: $p < .001$; 4 vs. 3: $p < .001$). It follows that the selected content relations and epistemic relations are equally appropriate. None of them simultaneously have a subjective interpretation and an objective interpretation that are on a par with each other.

Creating the stimuli

The selected test items, 40 sets of stories, were then divided into four lists using a Latin square design. Each participant in the experiment read only one list. To avoid having participants see a story more than once, each list included only one condition of each story. However, when all stories were taken into account, all four experimental conditions appeared in each list. Thereby, the two factors involved in this design, relation (objective vs. subjective) and connective (specified vs. underspecified), are within-subjects variables.

To avoid strategic processing, 60 texts containing a wide variety of non-causal coherence relations (such as additive, conditional, concessive, temporal, etc.) were used as fillers. The filler items were matched with the experimental items in terms of text length. One pseudo-randomization was used for all lists. Thirty verification questions were included to encourage reading for comprehension. Half of the questions followed the experimental items and half followed the filler items. Furthermore, half of the questions required a "yes" answer and the other half required a "no" answer. The questions never probed inter-clausal relations.

3.3 Procedure

The experiment was conducted in the Eye-tracking Lab of the Utrecht institute of Linguistics OTS, Utrecht University. The materials were presented on a computer screen, and a desktop-mounted EyeLink 1000 eye-tracker was used to record eye movements. Participants were tested individually in the lab. They were seated comfortably and then were asked to read an instruction on the screen. They were informed that they had to look at a fixation point to make a text available, and that they should press the "yes" or "no" button in response to verification questions. They were also instructed not to move their head or blink excessively during the experiment. After the instruction, the eye-tracker was adjusted, and then a thirteen-point calibration and validation procedure was carried out. Upon successful calibration and validation, the experiment started with five practice items, two of which were followed by verification questions. If participants performed the procedures appropriately with the practice items, the real test started. Each participant read 105 texts in the experiment, which took about 25 minutes on average.

3.4 Critical regions and measures

The second clause of each text was the target for analysis. Table 4 shows the way in which a target clause was divided.

Table 4. Three regions per target sentence

	Connective region	Subject region	Predicate region
Chinese item	*suoyi/yin'er/kejian*	*ta bi yiqian*	*shou le bushao.*
Translation	so	she (now) than before	(has got) thinner much.

As the names imply, the connective region contained the connective, the subject region contained a subject, usually a noun phrase and sometimes also a time or place adverbial, and the predicate region contained the predicate. Note that the results at the connective region cannot be interpreted straightforwardly, because the connectives differ in frequency and form. Therefore we focus on the regions after the connective. The target sentence was displayed near the center of the screen, in order to prevent 'edge effects' (Koornneef & van Berkum, 2006; Rayner, 1998).[3]

Following Traxler et al. (1997a), we used four reading-time measures. Three of them concerned first-pass reading. *First-pass reading time* (FP) is the total reading time (including fixations and saccade durations) spent on a region before the eyes leave the region either in a progressive manner or a regressive manner. *First-pass total gaze duration* (TG, referred to as *right-bounded time* in Traxler et al., 1997a) is the sum of durations of fixations that fall within a region before the region is left progressively. Different from *first-pass reading time*, *first-pass total gaze duration* can include fixation durations after regressions. *Regression path duration* (RP) is the time between the start time of the first fixation in a region and the end time of the last fixation before the region is left progressively. It includes not only *first-pass total gaze* durations of the region, but also the rereading (as a result of regression) time of the previous regions. The fourth measure, *total fixation duration* (TF, referred to as *total time* in Traxler et al., 1997a), includes not only fixations during first-pass reading, but also the second-pass reading, the third-pass reading, and so on. It is the sum of all fixation durations in a region, which captures the time required for reanalysis when a piece of information is not completely processed during the first reading (Rayner & Sereno, 1994).

Prior to analysis, we deleted the data of four participants due to excessive eye blinks, or poor drift correction. This left us with the data of forty participants. For further clean-up, any observation more than two standard deviations from the participant's mean and the item's mean were discarded from further analyses. On the basis of these criteria, we removed 0.64% of the reading-time data (244 cells of measurements) from analysis.

---

3. The reading times associated with line beginnings and endings are likely to be contaminated by return sweeps. According to Rayner (1998), the first and last fixations on a line are generally 5–7 letter spaces from the ends of a line, the first fixation on a line tends to be longer than other fixations, and the last is shorter. This type of effect can be referred to as 'edge effects'.

Next, we evaluated our participants' performance on verification questions. All participants scored above 80% (mean score 92%). Participants' high average score on the verification task guaranteed that they were reading for comprehension. On the grounds that participants were reading for comprehension, we concluded that their eye movement data reflected the natural reading processes we aimed to study.

4. Analysis and results

All data were analyzed using the R packages *lme4* (Bates, Maechler, & Bolker, 2012) and *languageR* (Baayen, 2011). *Linear mixed effects regression analyses* (LMER) (Baayen, 2008) were performed on the log of the means of the four reading time measures, respectively.[4] We used both *subject* and *item* as random effects (see

Table 5. Mean reading times (and standard deviations) per region (in ms)

Measure & condition	Connective region	Subject region	Predicate region
First-pass reading time:			
A: content relation with *yin'er*	284 (159)	345 (171)	430 (266)
B: content relation with *suoyi*	233 (71)	363 (225)	422 (249)
C: epistemic relation with *kejian*	251 (101)	366 (212)	367 (226)
D: epistemic relation with *suoyi*	242 (79)	367 (225)	416 (259)
First-pass total gaze duration:			
A: content relation with *yin'er*	287 (161)	398 (218)	452 (280)
B: content relation with *suoyi*	241 (78)	375 (225)	436 (256)
C: epistemic relation with *kejian*	259 (110)	383 (215)	402 (251)
D: epistemic relation with *suoyi*	243 (79)	389 (236)	457 (280)
***Regression path duration*:**			
A: content relation with *yin'er*	294 (176)	436 (289)	484 (314)
B: content relation with *suoyi*	248 (99)	392 (230)	463 (282)
C: epistemic relation with *kejian*	270 (148)	405 (233)	434 (295)
D: epistemic relation with *suoyi*	258 (118)	403 (246)	530 (417)
***Total fixation duration*:**			
A: content relation with *yin'er*	372 (242)	523 (342)	510 (321)
B: content relation with *suoyi*	284 (138)	486 (307)	508 (354)
C: epistemic relation with *kejian*	314 (175)	528 (353)	461 (313)
D: epistemic relation with *suoyi*	344 (202)	594 (400)	569 (392)

4. Log-transformations were performed on the means in order to meet the requirement of normality. Parametric statistical techniques, such as Linear mixed effects models, require the difference between the conditions to be normally distributed, e.g., the differences between condition A and condition B need to approximate a bell-shaped curve.

Baayen, Davidson, & Bates, 2008). As fixed effects, we included *relation* (objective vs. subjective), *connective* (underspecified vs. specified), and the interaction of *relation* and *connective* in the models. Whenever an interaction effect was observed, we performed extra pairwise comparisons between conditions.

In Table 5, we present the mean reading times and standard deviations for each reading-time measure and for each region under investigation.

4.1 First-pass reading

At the subject region, no significant effects were observed, but at the predicate region, we observed an interaction effect between relation and connective (FP: $\beta = -0.13$, $SE = 0.05$, $t = -2.39$, $p = .02$; TG: $\beta = -0.13$, $SE = 0.05$, $t = -2.63$, $p = .01$; RP: $\beta = -0.17$, $SE = 0.06$, $t = -2.92$, $p = .004$). In the subjective sentences, pairwise comparisons showed that participants spent shorter times at this region when *kejian* was used than they did when the connective was underspecified (i.e., *suoyi)* (FP: $\beta = -0.10$, $SE = 0.04$, $t = -2.65$, $p = .008$; TG: $\beta = -0.10$, $SE = 0.04$, $t = -2.57$, $p = .01$; RP: $\beta = -0.14$, $SE = 0.04$, $t = -3.12$, $p = .002$). Meanwhile, no significant difference was observed between the reading times of content causal relations with *yin'er* and those with *suoyi* (FP: $\beta = 0.03$, $SE = 0.04$, $t = 0.78$, $p = .45$; TG: $\beta = 0.04$, $SE = 0.04$, $t = 1.10$, $p = .28$; RP: $\beta = 0.03$, $SE = 0.04$, $t = 0.86$, $p = .38$). Furthermore, no significant differences were observed between content causal relations with *suoyi* and epistemic causal relations with *suoyi* (FP: $\beta = -0.02$, $SE = 0.04$, $t = -0.46$, $p = .64$; TG: $\beta = 0.05$, $SE = 0.04$, $t = 1.23$, $p = .25$; RP: $\beta = 0.08$, $SE = 0.04$, $t = 1.86$, $p = .07$).

4.2 Total fixation duration

For *total fixation duration*, there was an interaction effect at all three regions (the connective region: $\beta = -0.28$, $SE = 0.06$, $t = -4.92$, $p < .001$; the subject region: $\beta = -0.18$, $SE = 0.06$, $t = -3.09$, $p = .001$; the predicate region: $\beta = -0.19$, $SE = 0.05$, $t = -3.51$, $p < .001$). Pairwise comparison between the epistemic conditions showed that there were faster reading times with the specified causal connective *kejian* than with the underspecified causal connective *suoyi* at both the subject region ($\beta = -0.12$, $SE = 0.04$, $t = -2.86$, $p = .005$) and the predicate region ($\beta = -0.16$, $SE = 0.04$, $t = -4.04$, $p < .001$). For the pairwise comparison between the content causal relations, no effect was found.

In addition, pairwise comparisons between subjective and objective relations with an underspecified connective showed that the epistemic relations with *suoyi* led to longer reading times than the content relations with *suoyi* (the connective

region: $\beta = 0.16$, $SE = 0.04$, $t = 4.24$, $p < .001$; the subject region: $\beta = 0.18$, $SE = 0.04$, $t = 4.25$, $p < .001$; the predicate region: $\beta = 010$, $SE = 0.04$, $t = 2.65$, $p = .01$).

5. Discussion

Unlike previous processing studies, the current study simultaneously included specified causal connectives and underspecified causal connectives within the same language in the experimental design. This enabled us to directly compare the roles of these two types of causal connectives during online processing. Our design ensured that these conditions were highly comparable, as the materials were constructed in such a way that critical sentences in the conditions to be compared differed only in one word, i.e., the connective itself.

In line with our first hypothesis, the current eye-tracking experiment replicated the results reported in Traxler et al. (1997a). In both experiments, it was found that epistemic relations with an underspecified causal connective (i.e. *suoyi* in the current experiment and *because* in Traxler et al., 1997a) cost longer total reading times than content relations with the same causal connective. This replication offers new evidence in support of the so-called subjective complexity hypothesis: a subjective relation is cognitively more complex than an objective relation (Sanders, 2005).

As predicted, the results also showed that the epistemic causal relations marked with *kejian* resulted in shorter reading times at the predicate region than the epistemic causal relations marked with *suoyi*. The difference between the two epistemic conditions solely concerns the connective, in particular, the specificity of the connective with respect to subjectivity. Accordingly, we can conclude that the high degree of subjectivity encoded in *kejian* has facilitated the reading and interpretation of the subjective epistemic relation. Contrary to our prediction (which was based on the results of Canestrelli et al. in Dutch), however, we found this effect at the end of the sentence, not at the region immediately following the connective.

Crucially, these results provide further insight into how a high degree of subjectivity encoded in the connective affects online processing. As discussed above, at the predicate region, when the connective was underspecified, epistemic causal relations resulted in longer total reading times than content causal relations. This finding has been related to the inherent cognitive complexity of epistemic causal relations in general (Evers-Vermeul & Sanders, 2011). However, when epistemic causal relations were marked with the subjective connective *kejian*, the difficulty associated with the processing of the relation per se decreased: the predicate region of the *kejian* clauses showed shorter processing times than the predicate region of the *suoyi* clauses. The processing difficulty induced by the high degree of subjectivity of the relation is reduced when the connective already marks the relation as subjective.

In the content relations, the objective causal connective *yin'er* did not facilitate the processing of the unfolding relation more than the underspecified connective *suoyi*: at the subject and the predicate region, we observed no processing differences between the *yin'er* condition and the *suoyi* condition. Several explanations for this finding come to mind. First, this result conforms to Traxler et al.'s (1997b) proposal that readers might prefer to build the simplest possible discourse representation during interpretation, in this case the objective relation. This would suggest that, on reading the underspecified causal connective *suoyi*, readers construct an objective content causal relation (i.e., the simplest causal relation) in the same way as they do after encountering the objective connective *yin'er*.

An alternative explanation for the lack of a speed-up effect after *yin'er* concerns the formality of this connective. According to the Taiwan Mandarin Spoken Wordlist, which was derived from the transcripts of a total of 42 hours of speech recording by Academia Sinica, *yin'er* never appeared in conversations. By contrast, *suoyi* was very frequently used in speech (2276 occurrences), and *kejian* was also found in informal conversations (4 occurrences). Our materials were written in an informal style. We refrained from using big and formal words, for the sake of the materials' easy readability. It might be the case that the formal style of *yin'er* did not entirely match the informal nature of the materials. Accordingly, it can be argued that *yin'er* might have a processing advantage (over *suoyi*) due to its specificity, which, however, is offset by the processing difficulty induced by its formal style. More research is needed to disentangle the role of style from that of the specificity of the connective. Our study could be replicated with another language that has both objective, subjective and underspecified causal connectives, or alternatively, the same experimental design could be used with Chinese materials in a more formal style.[5]

A third factor, the difference in frequency of use between *yin'er* and *suoyi*, does not seem to be a likely candidate for explaining the lack of differences in the processing of content relations marked with these connectives. For the epistemic cases, we found that any facilitative effect of a high word frequency did not outweigh the facilitative effect of the specificity of the connective: a speed-up effect was observed for the epistemic relations marked with *kejian* as compared to the ones marked with *suoyi*, in spite of the fact that *suoyi* is more frequent than *kejian*. Moreover,

5. Note that selecting another, less formal, objective Chinese connective is not an option. Although Chinese does have an objective as well as an underspecified result marker (see Table 1), respectively *yushi* and *yinci* (both translated as 'so'), a direct comparison of the processing effects of these connectives runs into different interpretation problems, because *yushi* can also be used to mark temporal relations. Replacing *yin'er* with *yinci* does not solve the formality problem of the objective connective *yin'er*, because *yinci* is an underspecified instead of an objective connective.

the effects were found at the final region of the sentence and in the rereading times. It is highly unlikely that these effects would be attributable to the difference in frequency of the connectives.

Finally, we would like to discuss what possibly makes texts encoded with a high degree of subjectivity difficult to process during online reading. There are different theoretical explanations. Canestrelli et al. (2013) have argued that subjective information requires more cognitive effort and longer processing time than objective information, because the former contains an additional proposition compared to the latter, a meta-representation of the speaker's or other's beliefs or conclusions, which could be in the form of 'I/someone think(s)' (see also De Smet & Verstraete, 2006; Ross, 1970; Rutherford, 1970). *Kejian* directly encodes the speaker's perspective (see Li et al., 2013), so it is likely to trigger the representation of the proposition 'I think' immediately, which should increase the cognitive cost, and hence should lead to longer processing time immediately after the connective. In the case of the underspecified causal connective *suoyi*, this representation 'I think' has to be constructed on the basis of the propositional content of P and Q, the bulk of which is often available at the predicate of the second clause. This should lead to a slowdown in processing at the predicate region. Clearly, this 'extra information account' can explain the major results of the current experiment in a very sensible way. Moreover, this explanation is also plausible because it is in line with native speakers' intuitions about the conceptual difference between *kejian* and *suoyi.* Compare (14a), (14b), and (14c).

(14) *Meng Na na tiao kuzi xianzai xiande hen fei,*
'That old pair of trousers now look very baggy on Meng Na,
a. ***suoyi wo renwei*** *ta bi yiqian shou le bu shao.*
'**so I think** she has become much thinner now than before.'
b. ****kejian wo renwei*** *ta bi yiqian shou le bu shao.*
'**so I think** she has become much thinner now than before.'
c. ***kejian*** *ta bi yiqian shou le bu shao.*
'**so** she has become much thinner now than before.'

As shown in (14a), the underspecified causal connective *suoyi* allows the addition of the words *wo renwei* 'I think' right after it, whereas (14b) illustrates that adding *wo renwei* 'I think' after the subjective connective *kejian* makes the originally appropriate sentence in (14c) unacceptable. It can be argued that the connective *kejian* is encoded with subjectivity information that contains the additional proposition 'I think', and thus adding the same proposition after *kejian* creates redundancy. The underspecified causal connective *suoyi* is not encoded with any subjectivity information, not to mention the proposition 'I think'. So adding 'I think' to the *suoyi* clause will not create any problems.

A straightforward way to verify this account is to set up a design that includes two conditions: an epistemic causal relation marked with *kejian*, such as (14c), and an epistemic causal relation marked with "*suoyi* I think", such as (14a). If it is the construction of the extra proposition 'I think' that increases the processing difficulty for subjective information, then there should be no processing differences between the two conditions towards the end of the second clause. In both cases, the representation of the proposition 'I think' should be constructed at the beginning of the second clause. Canestrelli et al. (2013) have provided some evidence in support of this account (that subjective information contains an extra proposition) with an eye-tracking experiment using Dutch materials. We expect to provide further evidence for this account with the above design using Chinese materials.

In the literature, there are some suggestions (Fauconnier, 1998; Sanders, Sanders, & Sweetser, 2009) that language users build, and constantly update, a network of mental spaces when they communicate. In this view, to represent the proposition 'I/someone think(s)' would involve setting up and instantiating the mental space to present the speaker's or another person's thoughts, which would result in a higher processing cost. In addition to these strictly cognitive explanations, Sperber et al. (2010) speculate about other reasons why readers may spend more effort on other people's thoughts: 'epistemic vigilance' (Canestrelli, 2013). The speaker's or others' thoughts are not necessarily true, so readers will more carefully evaluate those thoughts/claims during reading, which is associated with a processing cost (Sperber et al., 2010). Further research will be needed to ascertain whether the subjective complexity observed in the current and previous processing studies can be related to these cognitive processes.

References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, R. H. (2011). *LanguageR: Data sets and functions with "Analyzing linguistic data: A practical introduction to statistics"*. R package version 1.4.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412.

Bates, D., Maechler, M., & Bolker, B. (2012). *Lme4: Linear mixed-effects models using S4 classes*. R package version 0.999999-0.

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

Canestrelli, A. R. (2013). *Small words, big effects?: Subjective versus objective causal connectives in discourse processing*. Ph.D. dissertation, Utrecht: LOT. Available online: http://www.lotpublications.nl/Documents/325_fulltext.pdf.

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

Couper-Kuhlen, E. (1996). Intonation and clause-combining in discourse: The case of *because*. *Pragmatics*, 6(3), 389–426.

Cozijn, R. (2000). *Integration and inference in understanding causal sentences*. Ph.D. dissertation, Tilburg University.

Dancygier, B. (2009). Causes and consequences: Evidence from Polish, English, and Dutch. In T. J. M. Sanders & E. Sweetser (Eds.), *Causal categories in discourse and cognition* (pp. 91–118). Berlin: Mouton de Gruyter.

Degand, L., & Pander Maat, H. (2003). A contrastive study of Dutch and French causal connectives on the Speaker Involvement Scale. In A. Verhagen & J. van de Weijer (Eds.), *Usage-based approaches to Dutch* (pp. 175–199). Utrecht: LOT.

De Smet, H., & Verstraete, J. C. (2006). Coming to terms with subjectivity. *Cognitive Linguistics*, 17(3), 365–392.

Evers-Vermeul, J. (2005). *The development of Dutch connectives: Change and acquisition as windows on form-function relations*. Utrecht: LOT. Available online: http://www.lotpublications.nl/Documents/110_fulltext.pdf.

Evers-Vermeul, J., Degand, L., Fagard, B., & Mortier, L. (2011). Historical and comparative perspectives on subjectification: A corpus-based analysis of Dutch and French causal connectives. *Linguistics*, 49(2), 445–478.

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

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

Fauconnier, G. (1998). Mental spaces, language modalities, and conceptual integration. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (pp. 251–279). Mahwah, NJ: Lawrence Erlbaum.

Ford, C. E. (1993). *Grammar in interaction: Adverbial clauses in American English conversations*. Cambridge: Cambridge University Press.

Haberlandt, K. F. (1982). Reader expectations in text comprehension. In J. Le Ny & W. Kintsch (Eds.), *Language and comprehension* (pp. 239–249). Amsterdam: North-Holland.

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

Keller, R. (1995). The epistemic 'weil'. In D. Stein & S. Wright (Eds.), *Subjectivity and subjectivisation: Linguistic perspectives* (pp. 16–30). Cambridge: Cambridge University Press.

Kintsch, W. (1988). Role of knowledge in discourse comprehension. *Psychological Review*, 95, 163–182.

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

Knott, A., & Sanders, T. J. M. (1998). The classification of coherence relations and their linguistic markers: An exploration of two languages. *Journal of Pragmatics*, 30, 135–175.

Koornneef, A. W., & Sanders, T. J. M. (2013). Establishing coherence relations in discourse: The influence of implicit causality and connectives on pronoun resolution. *Language and Cognitive Processes*, 28(8), 1169–1206.

Koornneef, A. W., & van Berkum, J. J. A. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language*, 54, 445–465.

Li, F. (2014). *Subjectivity in Mandarin Chinese: The meaning and use of causal connectives in written discourse*. Ph.D. dissertation, Utrecht University. Utrecht: LOT. Available online: http://www.lotpublications.nl/Documents/365_fulltext.pdf.

Li, F., Evers-Vermeul, J., & Sanders, T. J. M. (2013). Subjectivity and result marking in Mandarin: A corpus-based investigation. *Chinese Language and Discourse*, 4(1), 74–119.

Mak, W. M., & Sanders, T. J. M. (2013). The role of causality in discourse processing: Effects of expectation and coherence relations. *Language and Cognitive Processes*, 28(9), 1414–1437.

Mann, W. C., & Thompson, S. A. (1986). Relational propositions in discourse. *Discourse Processes*, 9, 57–90.

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

Martin, J. R. (1992). *English texts: System and structure*. Amsterdam: John Benjamins.

McEnery, T., & Xiao, R. (2004). *The Lancaster corpus of Mandarin Chinese*. Lancaster: Lancaster University.

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

Moeschler, J. (1989). Pragmatic connectives, argumentative coherence and relevance. *Argumentation*, 3, 321–339.

Murray, J. D. (1997). Connectives and narrative text: The role of continuity. *Memory & Cognition*, 25(2), 227–236.

Noordman, L. G. M., & de Blijzer, F. (2000). On the processing of causal relations. In E. Couper-Kuhlen & B. Kortmann (Eds.), *Cause, condition, concession and contrast: Cognitive and discourse perspectives* (pp. 35–56). Berlin: Mouton de Gruyter.

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* (pp. 75–93). Mahwah, NJ: Lawrence Erlbaum.

Pander Maat, H., & Degand, L. (2001). Scaling causal relations and connectives in terms of Speaker Involvement. *Cognitive Linguistics*, 12, 211–245.

Pander Maat, H., & Sanders, T. (2001). Subjectivity in causal connectives: An empirical study of language in use. *Cognitive Linguistics*, 12, 247–273.

Pit, M. (2003). *How to express yourself with a causal connective: Subjectivity and causal connectives in Dutch, German and French*. Ph.D. Dissertation, Utrecht University. Amsterdam: Rodopi.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422.

Rayner, K., & Sereno, S. C. (1994). Eye movements in reading: Psycholinguistic studies. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 57–82). New York: Academic Press.

Ross, J. R. (1970). On declarative sentences. In R. A. Jacobs & P. S. Rosenbaum (Eds.), *Readings in English transformational grammar* (pp. 222–272). Washington: Georgetown University Press.

Rutherford, W. (1970). Some observations concerning subordinate clauses in English. *Language*, 46(1), 97–115.

Sanders, J., Sanders, T. J. M., & Sweetser, E. (2012). Responsible subjects and discourse causality: How mental spaces and connectives help identifying subjectivity in Dutch backward causal connectives. *Journal of Pragmatics*, 44, 191–213.

Sanders, T. J. M. (1997). Semantic and pragmatic sources of coherence: On the categorization of coherence relations in context. *Discourse Processes*, 24, 119–147.

Sanders, T. J. M. (2005). Coherence, causality and cognitive complexity in discourse. In M. Aurnague, M. Bras, A. L. Draoulec, & L. Vieu (Eds.), *Proceedings of the first international symposium on the exploration and modelling of meaning SEM-05* (pp. 31–46). Toulouse, France: Université de Toulouse-le-Mirail.

Sanders, T. J. M., Sanders, J., & Sweetser, E. (2009). Causality, cognition and communication: A mental space analysis of subjectivity in causal connectives. In T. J. M. Sanders & E. Sweetser (Eds.), *Causal categories in discourse and cognition* (pp. 19–59). Berlin: Mouton de Gruyter.

Sanders, T. J. M., & Spooren, W. (2009). Causal categories in discourse: Converging evidence from language use. In T. J. M. Sanders & E. Sweetser (Eds.), *Causal categories in discourse and cognition* (pp. 205–246). Berlin: Mouton de Gruyter.

Sanders, T. J. M., & Spooren, W. P. M. (2015). Causality and subjectivity in discourse: The meaning and use of causal connectives in spontaneous conversation, chat interactions and written text. *Linguistics*, 53(1), 53–92.

Sanders, T. J. M., Spooren, W. P. M., & Noordman, L. G. M. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, 15(1), 1–35.

Sanders, T. J. M., Spooren, W. P. M., & Noordman, L. G. M. (1993). Coherence relations in a cognitive theory of discourse representation. *Cognitive Linguistics*, 4(2), 93–133.

Shen, J. (2003). Compound sentences in three conceptual domains: Acting, knowing, and uttering. *Chinese Language*, 3, 195–204.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind and Language*, 25(4), 359–393.

Spooren, W. P. M., & Sanders, T. J. M. (2008). The acquisition order of coherence relations: On cognitive complexity in discourse. *Journal of Pragmatics*, 40, 2003–2026.

Spooren, W., Sanders, T., Huiskes, M., & Degand, L. (2010). Subjectivity and causality: A corpus study of spoken language. In J. Newman & S. Rice (Eds.), *Empirical and experimental methods in cognitive/functional research* (pp. 256–270). Stanford, CA: CSLI Publications.

Stukker, N., & Sanders, T. J. M. (2012). Causal connectives in discourse: A cross-linguistic perspective. *Journal of Pragmatics*, 44, 131–137.

Stukker, N., Sanders, T. J. M., & Verhagen, A. (2008). Causality in verbs and in discourse connectives: Converging evidence of cross-level parallels in Dutch linguistic categorization. *Journal of Pragmatics*, 40, 1296–1322.

Sweetser, E. E. (1990). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.

Tao, H. (2007). Subjectification and the development of special-verb existential/presentative constructions. *Language and Linguistics*, 8(2), 575–602.

Traugott, E. C. (1995). Subjectification in grammaticalization. In D. Stein & S. Wright (Eds.), *Subjectivity and subjectivisation: Lingustic perspectives* (pp. 31–54). Cambridge: Cambridge University Press.

Traxler, M. J., Bybee, M. D., & Pickering, M. J. (1997a). Influence of connectives on language comprehension: Eye-tracking evidence for incremental interpretation. *The Quarterly Journal of Experimental Psychology*, 50A(3), 481–497.

Traxler, M. J., Sanford, A. J., Aked, J. P., & Moxey, L. M. (1997b). Processing causal and diagnostic statements in discourse. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(1), 88–101.

van Dijk, T. A. (1979). Pragmatic connectives. *Journal of Pragmatics*, 3, 447–456.
van Veen, R. (2011). *The acquisition of causal connectives: The role of parental input and cognitive complexity*. Ph.D. dissertation. Utrecht: LOT. Available online: http://www.lotpublications.nl/Documents/286_fulltext.pdf.
Verhagen, A. (2005). *Constructions of intersubjectivity: Discourse, syntax, and cognition*. Oxford: Oxford University Press.
Xing, F. (2001). *Hanyu fuju yanjiu* [*A study of Chinese complex sentences*]. Beijing: Commercial Publishing House.
Zufferey, S. (2012). 'Car, parce que, puisque' revisited: Three empirical studies on French causal connectives. *Journal of Pragmatics*, 44(2), 138–153.

Authors' addresses

Fang Li
School of English Language, Literature and Culture / Center for Language and Cognition
Beijing International Studies University
100024
China
lifang@bisu.edu.cn

Willem M. Mak
Utrecht Institute of Linguistics OTS
Utrecht University
Trans 10
NL-3512 JK Utrecht
The Netherlands
w.m.mak@uu.nl

Jacqueline Evers-Vermeul
Utrecht Institute of Linguistics OTS
Utrecht University
Trans 10
NL-3512 JK Utrecht
The Netherlands
j.evers@uu.nl

Ted J. M. Sanders
Utrecht Institute of Linguistics OTS
Utrecht University
Trans 10
NL-3512 JK Utrecht
The Netherlands
t.j.m.sanders@uu.nl

About the authors

Fang Li, lecturer at Beijing International Studies University, mainly studies the use and function of causal connectives. In 2013, she published a corpus-based study on subjectivity and result marking in Mandarin in the journal *Chinese Language and Discourse.* In 2014, she received her doctor's degree in linguistics at Utrecht University (the Netherlands), and published *Subjectivity in Mandarin Chinese: The meaning and use of causal connectives in written discourse* in the LOT dissertation series.

Pim Mak is assistant professor at the Department of Languages, Literature, and Communication at Utrecht University. His main research interests are the processing of discourse by adults, and the development of the understanding of discourse phenomena in young children. He is an expert in the use of eye tracking, in reading studies with adults, and in visual-world experiments in young children.

Jacqueline Evers-Vermeul, assistant professor at the Department of Languages, Literature and Communication at Utrecht University, studies the representation, production and comprehension of text-structural aspects of discourse. She uses a variety of methods to study adult connective use (synchronic and diachronic corpus-based analyses), students' understanding of school texts and examination questions (on- and off-line reading experiments among children and teenagers), and children's connective acquisition (visual-world experiments among 2-and 3-year-olds, growth-curve analyses of dense longitudinal corpora).

Ted J. M. Sanders (1963, Ph.D. 1992) is a full professor of Discourse Studies and language use at Utrecht University. His research concentrates on the linguistics and psycholinguistics of coherence in discourse representation. He (co-)published in many edited books and international top journals. He co-edited special issues, like *Discourse Processes* (2004) and *Journal of Pragmatics* (2012), as well as the book volumes *Text representation* (John Benjamins, 2001) and *Causal categories in discourse and cognition* (Mouton, 2009).