Discourse Processes

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hdsp20

The Role of Connectives and Stance Markers in the Processing of Subjective Causal Relations

Yipu Wei, Jacqueline Evers-Vermeul, Ted M. Sanders & Willem M. Mak

To cite this article: Yipu Wei, Jacqueline Evers-Vermeul, Ted M. Sanders & Willem M. Mak (2021) The Role of Connectives and Stance Markers in the Processing of Subjective Causal Relations, Discourse Processes, 58:8, 766-786, DOI: 10.1080/0163853X.2021.1893551

To link to this article: https://doi.org/10.1080/0163853X.2021.1893551

Published online: 22 Mar 2021.

DISCOURSE PROCESSES
2021, VOL. 58, NO. 8, 766–786
https://doi.org/10.1080/0163853X.2021.1893551

Routledge
Taylor & Francis Group

The Role of Connectives and Stance Markers in the Processing of Subjective Causal Relations

Yipu Wei, Jacqueline Evers-Vermeul, Ted M. Sanders, and Willem M. Mak

School of Chinese as a Second Language, Peking University, Beijing, P.R. China; Utrecht Institute of Linguistics OTS, Utrecht University, Utrecht, The Netherlands

ABSTRACT

Interpreting subjectivity in causal relations takes effort: Subjective, claim-argument relations are read slower than objective, cause-consequence relations. In an eye-tracking-while-reading experiment, we investigated whether connectives and stance markers can play a facilitative role. Sixty-five Chinese participants read sentences expressing a subjective causal relation, systematically varied in the use of stance markers (no, attitudinal, epistemic) in the first clause and connectives (neutral *suoyi* "so", subjective *kejian* "so") in the second clause. Results showed that processing subjectivity proceeds highly incrementally: The interplay of the subjectivity markers is visible as the sentence unfolds. Subjective connectives increased reading times, irrespective of the type of stance marker being used. Stance markers did, however, facilitate the processing of modal verbs in subjective relations. We conclude that processing subjectivity involves evaluating how the argument supports the claim and that connectives, modal verbs, and stance markers function as processing instructions that help readers achieve this evaluation.

Introduction

Understanding language involves making a mental representation of the people and objects referred to and of the propositions that are made about the world. To fully understand discourse, however, language users also need to understand when speakers are not just referring to the outside world but rather are expressing their opinions and attitudes (Van Dijk, 1982; Van Dijk & Kintsch, 1983). Various linguistic cues are used to differentiate such subjective expressions from mere descriptions of objective real-world facts. Within a clause, adverbials can express attitudes toward certain propositional elements, as in example (1) on the size of a ferry. Some linguistic cues have a larger scope over the entire clause or sentence. For instance, *perhaps* in (2) and *may* in (3) mark the content of the clause as the hesitant judgment of the author. Other cues attribute the content to another person instead of the author, such as *John said* in (4). How people process these cues and interpret another person's opinions and attitudes is an important question for theories of discourse representation and processing.

(1) The passenger capacity of this ferry is **ridiculously** large.
(2) **Perhaps**, the passenger capacity of this ferry is large.
(3) The passenger capacity of this ferry **may** be large
(4) **John said** the passenger capacity of this ferry is large.

In linguistic theories, the involvement of a speaker whose opinion is conveyed is called subjectivity (Finegan, 1995; Langacker, 1990; Lyons, 1977). In terms of coherence relations in discourse, a basic distinction is drawn between subjective and objective relations (Sanders & Sweetser, 2009). For

CONTACT Yipu Wei yipu.wei@outlook.com Yiheyuan Road No.5, Haidian District, 100871, Beijing, P.R. China.

instance, (5) is more subjective compared with example (6), because it expresses a claim-argument relation instead of a consequence-cause relation. The claim-argument relation in (5) involves a speaker who expresses opinions, arguments, attitudes, and so on, whereas the consequence-cause relation in (6) does not (both examples taken from Traxler et al., 1997b, p. 485). The involvement of such a speaker, or locutionary agent in Finegan's terminology, contributes to a higher degree of subjectivity.

(5) Heidi could imagine and create things, because/she won/first prize/at the art show.
(6) Heidi felt very proud and happy, because/she won/first prize/at the art show.

Readers are sensitive to such differences in the degree of subjectivity. Traxler et al. (1997b; cf. also Traxler et al., 1997a) found that subjective relations such as (5) led to longer processing times at the second clause compared with objective relations such as (6). The processing delays were observed at the prefinal region, *first prize* in (5), when it becomes clear that the second clause is an argument for the claim in the first clause. The eye-tracking study provides evidence for the incremental interpretation of sentences: As soon as readers are informed on the subjectivity of the relation, their reading speed decreases.

The processing of subjectivity can be influenced by linguistic markers such as connectives. English uses *because* to express both subjective and objective relations, as is illustrated in examples (5) to (8), which implies that *because* is underspecified for subjectivity. Other languages, however, differentiate their use of connectives (see Degand & Pander Maat, 2003; Li et al., 2013; Pander Maat & Sanders, 2000; Pit, 2003, 2006; Stukker & Sanders, 2012; Zufferey, 2012). For instance, the Dutch counterpart of (7) is preferably marked with the connective *want* "because", whereas the more objective connective *omdat* "because" is used in (8). Similarly, the Dutch connective *dus* "so" is also used more often to express the reasoning of a speaker, whereas the objective connective *daarom* "that's why" more frequently marks the relation between two clauses as real-world fact, as is illustrated with examples (9) and (10). English uses *so* in both of these cases.

(7) English: The passenger capacity of this long-distance ferry must be large,
because it takes one hour before everyone is boarded.
Dutch: De passagierscapaciteit van deze veerboot moet wel groot zijn,
want het duurt een uur voordat iedereen aan boord is.

(8) English: Passengers of the ferry have arrived early at the gate,
because it takes one hour before everyone is boarded.
Dutch: Passagiers van de veerboot zijn vroeg bij de poort aangekomen,
omdat het een uur duurt voordat iedereen aan boord is.

(9) English: It takes one hour before everyone is boarded on the ferry,
so the passenger capacity of this ferry must be large.
Dutch: Het duurt een uur voordat iedereen aan boord is,
dus de passagierscapaciteit van deze veerboot moet groot zijn.

(10) English: It takes an hour before everyone is boarded on the ferry,
so the passengers have arrived early at the gate.
Dutch: Het duurt een uur voordat iedereen aan boord van de veerboot is,
daarom zijn de passagiers vroeg bij de poort aangekomen.

Thus, while the connective *because* only has the function of marking causality, *want* "because" as a specific subjective connective also indicates that the relation between the two segments can be attributed to a person making the claim. These functions of connectives affect processing: The subjective connective *want* "because" leads to an immediate processing delay compared to the

objective connective *omdat* "because", well before the content of the second segment makes clear that the relation is subjective (Canestrelli et al., 2013). The effect in Dutch arises earlier than the effect that was found in English in sentences with the underspecified connective *because* (Traxler et al., 1997a). Is this difference in processing between *because* in English and *want* "because" in Dutch due to the difference between underspecified connectives and specified connectives in terms of subjectivity, or is it simply due to language differences between Dutch and English or differences between experiments? Mandarin Chinese provides a chance to examine this comparison with one language and in one experiment.

In Chinese, three types of connectives are available: an underspecified causal connective *suoyi* "so", a specific subjective connective *kejian* "so", and a specific objective connective *yin'er* "as a result". In sentences with the underspecified *suoyi* "so" such as (11), readers can only find out the relation is subjective on the basis of the propositional content of the second clause. If a specific subjective connective is present, as in (12), this informs readers at an earlier processing stage that the relation is subjective. An experimental processing study has shown that in the latter case, the reading time on the final region was faster than in sentences such as (11) (Li et al., 2017).[1] This finding is comparable to the results of the experiments with the English connective *because*.

(11) Changtu youlun xuyao tiqian yi ge xiaoshi zhunbei dengchuan, **suoyi** youlun yiding zaikeliang feichang keguan.

"The long-distance ferry takes one hour in advance for boarding, **so** the passenger capacity of the ferry must be considerable."

(12) Changtu youlun xuyao tiqian yi ge xiaoshi zhunbei dengchuan, **kejian** youlun yiding zaikeliang feichang keguan.

"The long-distance ferry takes one hour in advance for boarding, **so** the passenger capacity of the ferry must be considerable."

As we have illustrated in (1) to (4), connectives are not the only linguistic indicators of subjectivity. In this article, we therefore investigate when and how readers of Chinese make use of other linguistic cues of subjectivity in the clause preceding and/or after the connective and how these cues interact with the signal provided by connectives. In the next section, we discuss the types of linguistic cues marking subjectivity that have been shown to co-occur with subjective and objective causal connectives: stance markers and modal verbs. Then, we present the methodological considerations underlying our reading experiment on Chinese in which we used eye-tracking-while-reading to examine the interplay between stance markers, modal verbs, and connectives during the processing of subjective causal relations.

Collocation patterns and processing effects of subjectivity markers

In this section, we look at the collocational patterns found in authentic data on Chinese language use. If language users display certain preferences in the way they combine various elements expressing subjectivity, this may also be predictive of the ease with which they process certain combinations.

Subjectivity markers in discourse

In a distinctive collocational analysis, Wei et al. (2020) examined the collocates of the underspecified connective *suoyi* and the specific subjective *kejian*, especially collocates expressing subjectivity. Some linguistic cues appeared more often in the context of *suoyi* compared with that of *kejian*, such as the modal verbs *keneng* "may" and *yinggai* "should" and the cognition verbs *zhidao* "know", *xiang* "think", and *renwei* "consider". By contrast, other linguistic cues were identified as the collocates of *kejian*, such

as the expressions of surprisal *juran* "unexpectedly" and *jingran* "surprisingly" and indicators of importance such as *zhongyao* "important".

These findings can be related to different types of stance markers (Conrad & Biber, 2000), evaluations (Bednarek, 2009), or perspective markers (Dancygier & Sweetser, 2005). Conrad and Biber (2000, p. 57) suggest three subtypes of stance markers (see for similar classifications Bednarek, 2006, 2009; Thompson & Hunston, 2000):

- Epistemic stance, which indicates how certain the speaker or writer is or from where the information comes (e.g., *probably, according to the President*)
- Attitudinal stance, which indicates feelings or judgments about what is said or written (e.g., *surprisingly, unfortunately*)
- Style stance, which indicates how something is said or written (e.g., *honestly, briefly*)

Reformulating the findings in Wei et al.'s (2020) collocational study in these terms, we can say that *suoyi* patterns with modal verbs and epistemic stance markers, whereas *kejian* patterns with attitudinal stance markers. In the following sections we review the results of earlier processing experiments investigating the effects of these cues on processing, thereby also generating research questions and hypotheses for the current study.

Modal verbs as subjectivity markers

In the collocational study, Wei et al. (2020) observed that modal verbs appear more in the context of *suoyi* than in the context of *kejian*. This may be because modal verbs provide information about subjectivity, in the absence of such information at the connective *suoyi*.

The role of modal verbs as expressions of subjectivity is supported by processing evidence. In a visual world paradigm experiment, Wei et al. (2019) have shown the influence of modal verbs on people's visual attention and the interaction between the effect of modal verbs and that of connectives. Participants heard sentences describing either a subjective causal relation or an objective causal relation while they looked at an image depicting the situation described in the sentences and an image of the speaker, who was presented explicitly as the source of information. The subjective relations in the experiment contained a modal verb (*yiding* "must" or *keneng* "may") in the second clause, either preceded by the underspecified connective *suoyi* "so" as in (11) or by the subjective connective *kejian* "so" as in (12). With *suoyi*, an increased attention to the speaker was observed at the modal verb. By contrast, after *kejian* the modal verb did not have such an effect in directing attention to the speaker. These findings show that modal verbs encode information that has a similar effect as the information in subjective connectives: They direct the attention of the listener to the speaker in the picture. When this information had already been provided by the subjective connective in the prior context, the modal verb does not have this effect anymore since the subjectivity information encoded by the modal verb is redundant.

In the experiment reported in the current article, we tested how the information of modal verbs interacts with the information provided by other cues of subjectivity. We first tested the interplay of modal verbs and connectives marking different degrees of subjectivity. In accordance with previous findings on the processing of connectives marking different degrees of subjectivity (Canestrelli et al., 2013; Li et al., 2017), we made the following predictions. First, with a specific subjective connective marking the subjective relation, we expected a delay at or immediately after *kejian*. Second, if the relation is connected by the underspecified connective *suoyi*, the processing delay is expected when the participants read the modal verb, since that is where the information on the subjectivity of the relation is first presented.

However, processing times may also be affected by collocation patterns, since collocation processing benefits—processing advantages in reading found for words frequently co-occurring in language

use—have been reported repeatedly (Durrant & Doherty, 2010; Vilkaitė & Schmitt, 2017). According to the collocation processing benefits hypothesis, the second prediction above may be contradicted by the collocation pattern reported in corpus research: Modal verbs collocate more often with the underspecified connective *suoyi* than with the specific subjective connective *kejian*. Thus, the collocation of *suoyi* and modal verbs may lead to a reading facilitation at the modal verb in the context of *suoyi*.

Epistemic stance markers as subjectivity markers

Epistemic stance markers have also been shown to influence the processing of subjective relations. Traxler et al. (1997b) found that epistemic stance markers such as *John said/thought* and *perhaps* facilitate the processing of subjective relations with *because*. The extra processing time of subjective relations compared with objective ones was canceled out by the presence of stance markers. In Canestrelli et al.'s (2013) reading studies, epistemic stance markers such as the Dutch counterpart of *according to Peter* in (13c) canceled out the processing asymmetry between clauses containing the objective connective *omdat* in (13a) and clauses containing the subjective connective *want* in (13b) (see Introduction), which resulted in comparable reading times for (13c) and (13a).

(13)

a.

Hanneke was buiten adem, **omdat** ze vier trappen was afgerend om de post te halen.

'Hanneke was out of breath, **because** she ran down four stairs to get the mail.'

b.

Hanneke had haast, **want** ze was vier trappen afgerend om de post te halen.

'Hanneke was in a hurry, **because** she ran down four stairs to get the mail.'

c.

Volgens Peter had Hanneke haast, **want** ze was vier trappen afgerend om de post te halen.

'According to Peter, Hanneke was in a hurry, **because** she ran down four stairs to get the mail.'

(Adapted from Canestrelli et al., 2013, p. 1403)

In both the English and the Dutch study, the epistemic stance markers facilitate the comprehension of readers by explicitly attributing the content of the first clause to a character (*John* or *Peter*) or indicating the presence of a speaker who is not sure (*perhaps*) about the following content. Therefore, at a later region, when subjectivity is marked by the specific connective *want* or by the propositional content in the second clause, readers have already processed the information in the first segment as subjective, and hence the reading times in the second clause in (13c) are shorter than those in (13b). This facilitative effect of epistemic stance markers demonstrates an overlap in epistemic meaning with subjective connectives, an overlap that is consistent with the collocation patterns discussed before.

Stance markers such as *John said/thought* and *according to Peter* influenced the processing of subjective relations marked by *want* "because" in (13c) and *because* in (5), which display a claim-argument structure. However, subjective relations can also be formulated in argument-claim structures such as (9), (11), and (12) in which subjective connectives such as Dutch *dus* "so" and Chinese *kejian* "so" mark the second segment as a claim and the relation as a whole as subjective. In these cases, the role of epistemic stance markers may be different. In the argument-claim type of relation illustrated in (14), the stance marker *according to Peter* marks the first segment as an argument made by Peter. But the claim that *she must have been in a hurry* is not necessarily the opinion of Peter.

(14) According to Peter, Hanneke ran down four stairs to get the mail, so she must have been in a hurry.

Our first research question therefore is as follows:

How does the presence of epistemic stance markers affect the processing patterns at the connective region and at later regions in subjective relations with an argument-claim structure?

Attitudinal stance markers as subjectivity markers

We now turn from epistemic stance markers to attitudinal stance markers. Canestrelli et al. (2016) examined the processing difference between subjective and objective connectives in subjective claim-argument relations containing attitudinal markers such as *ridiculously* in (1): *The passenger capacity of this ferry is ridiculously large*. Although such attitudinal stance markers also express subjectivity, they did not affect the processing of subjectivity as much as epistemic stance markers like *according to Peter* did. In other words, the increased reading times after *want* in comparison with *omdat* were not canceled out by attitudinal stance markers such as *ridiculously*. This lack of effects may be because the evaluative marker *ridiculously* only modified a specific element within the clause instead of the entire clause. In the present experiment we therefore used materials with attitudinal stance markers that modified the first clause as a whole. Our second research question therefore is the following:

Do attitudinal stance markers affect processing in the same way as epistemic stance markers do?

In terms of the comprehension of subjectivity, this question concerns whether the influence of stance markers on the processing asymmetry is due to the general degree of subjectivity expressed by all stance markers or due to the particular dimension of subjectivity expressed by the stance marker at hand: Epistemic stance markers indicate the speaker's degree of certainty about the expressed opinion, whereas attitudinal stance markers indicate the speaker's attitude/feelings. Earlier, we have made predictions on the processing asymmetry between *kejian* and *suoyi* at the connective itself (or the region immediately after it) and the modal verb region. If the extra time to process subjective relations is related to subjectivity in general, regardless of whatever dimensions of subjectivity are involved, all types of stance markers should have a similar effect of alleviating the processing load, thereby canceling out the processing delay associated with subjective connectives at the connective region and the predicted processing delay at the modal verb region in the *suoyi* context. If the processing load of subjectivity, however, is due to establishing a particular dimension of subjectivity, epistemic stance for instance, different stance markers should have different effects. Accordingly, the processing asymmetry between subjective connectives and underspecified connectives may be canceled out by epistemic stance markers but not by other types of stance markers.

Collocation patterns may also change the way readers process connectives. Wei et al. (2020) have shown that epistemic stance markers pattern with *suoyi* instead of *kejian*, while attitudinal stance markers pattern with *kejian*. In line with the collocation processing benefits hypothesis, two predictions can be made on the processing of connectives. On the one hand, in the contexts with epistemic stance markers, the reading times of *suoyi* may be facilitated, but the reading time of *kejian* may not. In other words, the predicted processing asymmetry between suoyi and *kejian* (*kejian* is associated with longer reading times than *suoyi*) may be enlarged by the presence of epistemic stance markers, instead of canceled as predicted by the subjectivity hypothesis. On the other hand, the presence of attitudinal stance markers in the context may shorten the reading times of *kejian* but not of *suoyi*. In this case, this effect is in the same direction as the predicted subjectivity effect, leading to a cancellation of the processing difference between *suoyi* and *kejian*. However, where exactly this collocation processing benefit takes place in reading is not straightforward. It is also possible that the collocation effect surfaces at a later region, when readers are integrating information from the context.

Methods

Participants

Sixty-five participants took part in this reading experiment (44 women; mean age = 26.4, *SD* = 3.3; age range, 18–36). All participants were native speakers of Mandarin Chinese and were paid for their participation. Informed consent was obtained from all participants.

Materials

The materials consisted of 48 sets of test items and 48 fillers.[2] All test items were two-sentence narratives: a sentence expressing a subjective causal relation and a spill-over sentence. The first sentence consisted of an introductory clause describing factual events and was followed by a second clause with a connective and a judgment or conclusion.

In a 3 × 2 (Stance × Connective) design, we created six versions of each test item, manipulating the use of stance markers in the first clause (no stance marker, epistemic stance marker, and attitudinal stance marker) and connectives (*suoyi, kejian*) in the second clause (see Table 1 for an example item). The two conditions without stance marker were created to set a baseline of how Chinese speakers process subjective relations marked by different connectives without the influence of stance markers. The conditions with stance markers allowed us to compare the effects of epistemic versus attitudinal stance markers in affecting the processing of subjective relations. In all items we inserted a modal verb immediately after the subject in the second clause. The number of characters in the sentences varied from 38 to 57 (*M* = 45.83, *SD* = 3.65). The length of the second clause (the target clause) varied from 13 to 20 characters (*M* = 15.19, *SD* = 1.71).

There were 48 items sets (288 two-sentence narratives in total), divided over six lists according to a Latin square design. Each list contained one version of a sentence and eight items in each condition.

Participants were assigned to read one of the six lists in the experiment. Sentences were presented in three lines on the screen, with critical regions not positioned at the beginning of a line or at the end of a line. We checked the naturalness of the test items using acceptability and subjectivity ratings from 60 native speakers of Mandarin Chinese (see the data and report we have made available online https://doi.org/10.17605/OSF.IO/W9E2V). There were no differences subjectivity between the conditions. However, sentences with *kejian* were rated as more acceptable than sentences with *suoyi*, and sentences without stance markers were rated as more acceptable than sentences with stance markers.

One-third of the sentences in the experiment were followed by a verification statement. The verification statements were about the content of single clauses and not about the relation between clauses. Participants were informed to judge the statements by pressing a "True" or "False" button.

Apparatus

The experiment was conducted on an EyeLink-1000 eye tracker (SR Research), sampling at 500 Hz (every 2 ms). The experiment was controlled by the software *ZEP* (version 1.6.3, Veenker, 2013). A high-speed camera was affixed to a desktop mount to measure the eye movements during reading. Items were presented on a 36.4 × 27.2 cm computer screen (screen refresh rate, 85 Hz) with font size 30 (font, FZSongHei-B07). A drift check was performed before the appearance of every sentence. This drift check point was located at the position where the sentence would begin.

Procedure

Participants were tested individually in a sound-treated lab booth. They first received an instruction on the procedure of the experiment and tasks, which included reading sentences on the screen and judging the verification statements randomly following the sentences. Before starting the test, the

Table 1. Example Test Item

Clause	Language	Test Item
Clause with/without stance marker	Chinese	*Chinese: Zhejia xiaoqu cheku* ***Ø*** (no stance marker)***/tingshuo*** (epistemic stance marker)***/jingran*** (attitudinal stance marker) *zai baitian shijian bei dao,*
	Gloss translation	This living:district garage **Ø/hearsay/surprisingly** during day time PASS rob,
	Free translation	"**Ø/As reported,/Surprisingly**, T/the garage of this district was robbed during daytime,"
Clause with connective	Chinese	***suoyi/kejian*** *xiaoqu anbao* ***keneng*** *zuo de bu daowei.*
	Gloss translation	**CONJ** district security **may** do PRT NEG enough.
	Free translation	"**so** the security of this district **may** not have done enough (work)."
Spill-over sentence	Chinese	*Yixie jumin yaoqiu zengjia xiaoqu shexiangtou.*
	Gloss translation	Some residents request increase living:district camera.
	Free translation	"Some residents requested to increase the number of cameras in this district."

experimenter adjusted the height of the seat so participants could sit comfortably, and the distance between the participants and the screen was adjusted to 550 to 600 mm to make sure their eyes could be measured properly. Instead of using a chin rest, we put a target sticker on the participant's forehead so that gaze positions could be corrected for small head movements. The practice trial and the real experiment started with a 13-point calibration test followed by a validation of this calibration.[3] The practice trial (with three items) acquainted participants with the experimental procedure, after which the experiment began. The experiment took about 30 minutes.

Analysis procedure

Example (15) demonstrates the regions we analyzed within the second clause, our target clause. The first region was the connective region. The subject region contained the words between the connective and the modal verb, usually the subject. The modal verb region consisted of a modal verb (*keneng* "may" or *yiding* "must") and three or four characters (one or two words) after the modal verb, and the final region contained the final words of the target clause.

(15)

[***Suoyi*/*kejian***]$_{\text{Connective region}}$ [*xiaoqu anbao*]$_{\text{Subject region}}$ [**keneng** *zuo de*]$_{\text{Modal verb region}}$ [*bu daowei.*]$_{\text{Final region}}$

[**CONJ**]$_{\text{Connective region}}$ [district security]$_{\text{Subject region}}$ [**may** do PRT]$_{\text{Modal verb region}}$ [NEG enough.]$_{\text{Final region}}$

"So the security of this district may not have done enough (work)."

Blinks and missing observations due to skipped regions were excluded from the data, just like observations that were 2 SDs above or below both the item mean and the subject mean. Then, we computed four reading time measures: *first-pass reading time, first-pass gaze duration, regression path duration*, and *total reading time. First-pass reading time* (the time between the onset of the first-pass first fixation in a region and the end time of the last fixation before leaving the region in any direction) provides useful information of the initial processing time of a region. *First-pass gaze duration* (the sum of fixations on a region before moving progressively to another region) is used as another way of measuring early-stage processing. *Regression path duration* (the time between the onset of the first-pass first fixation in a certain region and the end time of the last fixation before leaving the region in a forward direction) includes rereading of previous regions and reflects integration of a word with the previous context. This measure has been shown to be sensitive to the effect of subjectivity (e.g., Canestrelli et al., 2013). *Total reading time* measures the overall reading time of a region. For all four measures, a linear mixed effects regression analysis (Baayen et al., 2008) was performed on the reading time of each region using lme4 package (Bates et al., 2015) in R (R Core Team, 2015; version 3.1.3).

The reading times were log-transformed to adjust for normality. We started with a full model including the two factors connective and stance and the interaction between connective and stance. Treatment coding was used in the analysis. The no stance marker condition was taken as the baseline of the factor stance and the *suoyi* condition as the baseline of the factor connective. The effects of connective and stance were examined by comparing a specific type of connective/stance marker type with the baseline levels. The intercepts of items and subjects were included as random factors. We then tested whether excluding the interaction effects from a model would decrease the model fit significantly. The interaction effects whose exclusion did not cause a significant decrease of the model fit were dropped from the final model. Since the main research questions of this study concern the influence of different connectives and stance markers on reading, we did not drop any fixed factors from the model.

Results

We included data from 60 participants, after excluding data from 5 participants because of poor quality (e.g., severe drifts, too many blinks). The outliers were less than 0.5% of the total number of observations (*first-pass reading time*, 0.2%; *first-pass gaze duration*, 0.2%; *regression path duration*, 0.5%; *total reading time*, 0.3%). The participants whose data were included in the analysis all had an accuracy rate of above 84% (27/32) in judging the verification statements.

In the following report of the findings, we focus on the significant results and important comparisons of reading times. For each region (connective, subject, modal verb, and final region), the results on *first-pass reading time, first-pass gaze duration*, and *regression path duration* are summarized in response to our predictions: the processing effects of connectives and modal verbs without the presence of stance markers, the effects of stance markers, and the interaction effects of stance markers with connectives. *Total reading time*, as a measure of later-stage processing, is reported and interpreted separately from the other measures. Table 2 shows the means and SDs for all four measures.

Connective region

Dropping the interaction between connective and stance did not reduce the model fit significantly in any of the three first-pass measures (*first-pass reading time*: χ^2 (2, n = 1,304) = 1.04, p = .59; *first-pass gaze duration*: χ^2 (2, n = 1,303) = 1.41, p = .50; *regression path duration*: χ^2 (2, n = 1,289) = 0.51, p = .77). Therefore, the final analytical models for the connective region only included the factors of connective and stance and the random factors.

For all three measures, an effect of connective at this region showed that the reading times at the subjective connective *kejian* were longer than the reading times of the underspecified connective *suoyi* (*first-pass reading time*: β = 0.05, *SE* = 0.02, t(1,230) = 2.86, p = .004; *first-pass gaze duration*: β = 0.05, *SE* = 0.02, t(1,226) = 2.50, p = .01; *regression path duration*: β = 0.06, *SE* = 0.02, t(1,213) = 2.51, p = .01). This finding is consistent with the prediction that subjective connectives in general take longer processing times.

Table 2. Mean First-Pass Reading Times, First-Pass Gaze Durations, and Regression Path Durations (and SDs) per Region Aggregated by Participant

	Connective Region	Subject Region	Modal Verb Region	Final Region
First-pass reading time				
No stance marker + *suoyi*	243 (50)	344 (111)	444 (161)	396 (169)
No stance marker + *kejian*	257 (70)	330 (106)	426 (136)	376 (148)
Epistemic stance marker + *suoyi*	244 (61)	361 (137)	406 (154)	387 (127)
Epistemic stance marker + *kejian*	270 (73)	345 (126)	415 (174)	393 (148)
Attitudinal stance marker + *suoyi*	256 (64)	363 (125)	407 (157)	393 (148)
Attitudinal stance marker + *kejian*	273 (80)	336 (106)	427 (166)	371 (116)
First-pass gaze duration				
No stance marker + *suoyi*	254 (62)	393 (133)	471 (180)	433 (168)
No stance marker + *kejian*	260 (71)	373 (130)	470 (174)	435 (171)
Epistemic stance marker + *suoyi*	252 (62)	414 (175)	434 (169)	445 (194)
Epistemic stance marker + *kejian*	280 (80)	383 (150)	450 (190)	446 (174)
Attitudinal stance marker + *suoyi*	265 (72)	409 (156)	450 (202)	454 (168)
Attitudinal stance marker + *kejian*	279 (96)	389 (134)	457 (177)	410 (129)
Regression path duration				
No stance marker + *suoyi*	263 (71)	435 (173)	504 (204)	553 (225)
No stance marker + *kejian*	287 (130)	422 (184)	510 (216)	560 (256)
Epistemic stance marker + *suoyi*	274 (81)	477 (248)	466 (192)	571 (255)
Epistemic stance marker + *kejian*	303 (106)	432 (182)	494 (212)	575 (233)
Attitudinal stance marker + *suoyi*	338 (187)	439 (193)	488 (228)	631 (250)
Attitudinal stance marker + *kejian*	322 (173)	440 (186)	500 (224)	513 (205)

Since we did not find an interaction effect of stance and connective at this region, the effect of connective held for all stance conditions. In other words, the prediction that the presence of stance markers may cancel the processing asymmetry between the subjective connective and the underspecified connective was not confirmed.

There was an effect of stance, but only for the *regression path duration*: This reading time of connectives under the attitudinal stance marker condition was longer compared with the no stance marker condition ($\beta = 0.07$, $SE = 0.03$, $t(1,213) = 2.41$, $p = .02$). The *regression path duration* did not significantly differ between the epistemic stance marker condition and the no stance marker condition ($\beta = 0.05$, $SE = 0.03$,$t(1,216) = 1.69$,$p = .09$) or between the epistemic stance marker condition and the attitudinal stance marker condition ($\beta = -0.02$,$SE = 0.03$,$t(1,218) = -0.75$,$p = .45$).

Subject region

Concerning the subject region, excluding the interaction between connective and stance did not reduce the model fit for the three measures (*first-pass reading time*: χ^2 (2, n = 2,055) = 0.35, $p = .84$; *first-pass gaze duration*: χ^2 (2, n = 2,057) = 0.61, $p = .74$; *regression path duration*: χ^2 (2, n = 2,028) = 2.05, $p = .36$). Therefore, the final models for this region only included the factors of connective and stance and the random factors.

For all three measures, the connective *kejian* led to a general facilitation compared with the connective *suoyi* across all stance conditions in reading the subject region (*first-pass reading time*: $\beta = -0.04$, $SE = 0.02$, $t(1,950) = -2.06$, $p = .04$; *first-pass gaze duration*: $\beta = -0.04$, $SE = 0.02$, $t(1,951) = -2.19$, $p = .03$). We did not expect this facilitation effect of *kejian* at the subject region and discuss this finding in the Discussion.

We did not find a significant influence of stance markers in this region: The epistemic stance marker condition did not differ from the no stance marker condition (*first-pass reading time*: $\beta = 0.01$, $SE = 0.02$, $t(1,954) = 0.19$, $p = .85$; *first-pass gaze duration*: $\beta = -0.001$, $SE = 0.02$, $t(1,954) = -0.05$, $p = .96$; *regression path duration*: $\beta = 0.01$, $SE = 0.03$, $t(1,926) = 0.46$, $p = .64$). Also, the attitudinal stance marker condition was not different from the no stance marker condition (*first-pass reading time*: $\beta = 0.03$, $SE = 0.02$, $t(1,954) = 1.27$, $p = .21$; *first-pass gaze duration*: $\beta = 0.03$, $SE = 0.02$, $t(1,954) = 1.29$, $p = .20$; *regression path duration*: $\beta = 0.02$, $SE = 0.03$, $t(1,926) = 0.78$, $p = .44$).

Modal verb region

At the modal verb region, excluding the interaction between connective and stance did not affect the model fit significantly in any of the three measures (*first-pass reading time*: χ^2 (2, n = 2,161) = 4.32, $p = .12$; *first-pass gaze duration*: χ^2 (2, n = 2,160) = 0.80, $p = .67$; *regression path duration*: χ^2 (2, n = 2,133) = 1.11, $p = .57$). The final models for this region only contained the factors of connective and stance and the random factors.

For all three measures, no significant effects of connective were observed in this region (*first-pass reading time*: $\beta = -0.001$, $SE = 0.02$, $t(2,055) = -0.06$, $p = .96$; *first-pass gaze duration*: $\beta = 0.01$, $SE = 0.02$, $t(2,053) = 0.45$, $p = .65$; *regression path duration*: $\beta = 0.01$, $SE = 0.02$, $t(2,026) = 0.49$, $p = .62$). Against our predictions, the different degrees of subjectivity encoded in connectives did not make readers process modal verbs differently when the stance markers were absent.

However, we found significant effects of stance in the *first-pass gaze duration* and the *regression path duration*: The reading time of the modal verb region was shorter under the epistemic stance condition compared with the no stance marker condition (*first pass gaze duration*: $\beta = -0.06$, $SE = 0.03$, $t(2,054) = -2.25$, $p = .03$; *regression path duration*: $\beta = -0.06$, $SE = 0.03$, $t(2,028) = -2.13$, $p = .03$). Such a facilitation effect was also found for the attitudinal stance marker condition compared with the no stance marker condition, but only in the *first-pass gaze duration* ($\beta = -0.05$, $SE = 0.03$, $t(2,055) = -2.02$, $p = .04$). When taking the epistemic stance marker condition as the baseline, we found no significant differences between the attitudinal stance marker condition and the epistemic stance marker condition

(*first-pass reading time*: β = 0.01, *SE* = 0.03, *t*(2,058) = −0.47, *p* = .64; *first-pass gaze duration*: β = −0.01, *SE* = 0.03, *t*(1,310) = −0.35, *p* = .73; *regression path duration*: β = 0.01, *SE* = 0.03, *t*(1,307) = 0.23, *p* = .82).

The facilitation effect of both epistemic stance markers and attitudinal stance markers in the processing of modal verbs partially answers the second research question. We did not find evidence to differentiate the processing effect of the two types of stance markers.

Final region

For the final region, excluding the interaction between connective and stance had a significant effect on the model fit for the *first-pass gaze duration* (χ^2 (2, *n* = 2,101) = 6.29, *p* = .04) and the *regression path duration* (χ^2 (2, *n* = 2,039) = 9.28, *p* = .01) but not for the *first-pass reading time* (χ^2 (2, *n* = 2,100) = 2.37, *p* = .31). Hence, the final analytical model for the *first-pass gaze duration* and the *regression path duration* included the factors of connective and stance and their interaction. The model for the *first-pass reading time* only included connective and stance. The results for the *first-pass gaze duration* and *regression path duration* are given in Figure 1. The analytical models of these measures included the interaction terms. The effect of connective in the baseline condition (no stance marker) was not significant (*first pass gaze duration*: β = −0.02, *SE* = 0.03, *t*(1,992) = −0.63, *p* = .53; *regression path duration*: β = −0.03, *SE* = 0.04, *t*(1,931) = −0.70, *p* = .49). The interaction terms, which show whether the effect of connective in the other stance conditions is different from that in the no stance marker condition, showed that this was neither the case for the attitudinal stance marker condition (*first-pass gaze duration*: β = −0.07, *SE* = 0.05, *t*(1,992) = −1.50, *p* = .13; *regression path duration*: β = −0.12, *SE* = 0.06, *t*(1,931) = −1.92, *p* = .06) nor for the epistemic stance marker condition (*first-pass gaze duration*: β = 0.05, *SE* = 0.05, *t*(1,991) = 1.01, *p* = .31; *regression path duration*: β = 0.07, *SE* = 0.06, *t*(1,931) = 1.11, *p* = .27).

The model with "no stance marker" as the baseline condition did not provide a comparison between the effect of connective in the epistemic stance marker condition and the effect of connective in the attitudinal stance marker condition. We therefore releveled the model by using the epistemic stance marker condition as the baseline so we could directly compare the effect of connective in these conditions. In this analysis, the effect of connective in the epistemic condition (the baseline for stance in this model) was not significant (*first-pass gaze duration*: β = 0.03, *SE* = 0.03, *t*(1,991) = 0.79, *p* = .43; *regression path duration*: β = 0.04, *SE* = 0.04, *t*(1,931) = 0.87, *p* = .38). However, the interaction term comparing the effect of connective for the epistemic stance marker condition and the effect of connective in the attitudinal stance marker condition was significant (*first-pass gaze duration*: β = −0.12, *SE* = 0.05, *t*(1,990) = −2.49, *p* = .01; *regression path duration*: β = −0.19, *SE* = 0.06, *t*(1,930) = −3.01, *p* = .003) (see Table 2 for mean reading times under each condition).

Reading times on the connective *suoyi*, the baseline of connective, were longer in the attitudinal stance marker condition than in the epistemic stance marker condition (*regression path duration*: β = 0.09, *SE* = 0.04, *t*(1,931) = 2.11, *p* = .04). The findings in the final region provide evidence for collocation benefits in processing, not at the word in collocation with the connective itself, but rather at the end of the sentence containing the collocation word.

Total reading time

The analytical models of the *total reading time* of the four regions included only the two factors (stance and connective) but not their interactions, because excluding the interaction effect of these two factors did not decrease the model fit significantly in any of these regions (connective region: χ^2 (2, *n* = 1,810) = 1.37, *p* =.50; subject region: χ^2 (2, *n* = 2,121) = 1.04, *p* = .60; modal verb region: χ^2 (2, *n* = 2,215) = 1.25, *p* = .54; final region: χ^2 (2, *n* = 2,106) = 0.35, *p* = .84).

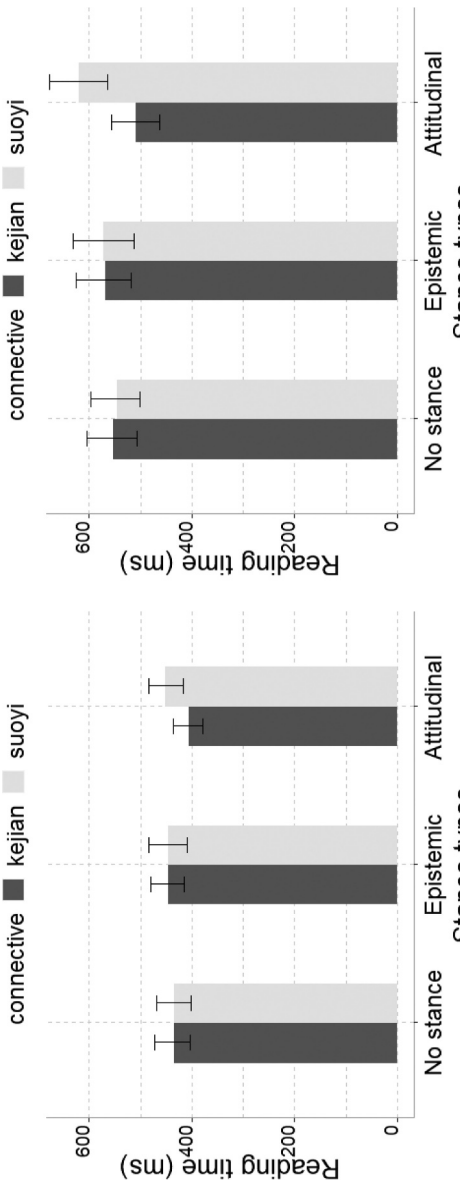

Figure 1. Mean reading times of the final region: first-pass gaze duration (left) and regression path duration (right). Error bars in the figure represent SEMs.

Table 3. Mean Total Reading Times (and SDs) per Region Aggregated By Participant

	Connective Region	Subject Region	Modal Verb Region	Final Region
No stance marker + *suoyi*	317 (103)	551 (197)	636 (241)	499 (198)
No stance marker + *kejian*	329 (97)	530 (191)	598 (225)	505 (225)
Epistemic stance marker + *suoyi*	314 (88)	572 (224)	605 (254)	538 (262)
Epistemic stance marker + *kejian*	336 (88)	520 (191)	601 (261)	511 (216)
Attitudinal stance marker + *suoyi*	321 (125)	543 (219)	614 (265)	521 (232)
Attitudinal stance marker + *kejian*	325 (122)	510 (190)	605 (251)	482 (149)

The *total reading times* of the four regions under different conditions are summarized in Table 3. For this measure, we only found a significant effect of the factor connective at the subject region. Across all stance conditions, the *total reading time* of the subject region was shorter when the connective was *kejian* in comparison to the *suoyi* conditions ($\beta = -0.07$, $SE = 0.02$, $t(2{,}017) = -2.96$, $p = .003$). That is to say, although the subjective connective *kejian* took longer to read itself in first-pass reading, it subsequently facilitated reading.

Discussion

Previous work has shown that subjective connectives lead to longer processing times than objective connectives. In this study we investigated the interplay of connectives with modal verbs and stance markers in the processing of subjective causal relations.

Effects of connectives

As a starting point, we tested the effects of connectives at three regions: the connective, the subject, and the modal verb when no stance markers were available in the previous context. Our data confirmed the effect of subjective connectives in argument-claim causal relations in Chinese: The subjective connective *kejian* "so" led to an immediate processing delay compared with the underspecified connective *suoyi* at the connective region. This result is comparable with the findings by Canestrelli et al. (2013) with Dutch *want* "because" expressing claim-argument causal relations. The processing delay supports previous claims that subjective relations are more difficult to process. Subjective connectives triggered the establishment of subjective relations immediately: This subjectivity effect was found in both Dutch and Chinese and in both the claim-argument relation and the argument-claim relation.

We found a facilitation effect of the subjective connective *kejian* compared with the underspecified connective *suoyi* at the subject region following the connective. The reason why *kejian* required less reading time than *suoyi* might be that *kejian* is a better fit for a subjective context, as substantiated by the acceptability ratings reported in our online repository. However, some alternative explanations could be obtained given the evidence from previous reading experiments. When the connective was underspecified, an effect of subjectivity was found at a later region—the prefinal region in Traxler et al. (1997b) and the final region in Li et al. (2017). In the present experiment, we found an effect of the specific subjective connective in the same direction but at a position where readers could not yet know from the content of the sentence that the relation was subjective. One explanation is that the effect is due to parafoveal processing of forthcoming words (i.e., the modal verb in this study; Rayner, 1998; Vasilev & Angele, 2017), which have been found across languages, especially for Chinese, in which bigger preview benefits have been found (Yang et al., 2009). From this perspective, the shorter reading times of the subject region under the *kejian* condition can be due to the facilitation effect of subjective connectives in providing preceding processing cues of subjectivity.

There is, however, an alternative explanation: The processing asymmetry may be due to a delaying effect of the underspecified connective—*suoyi* allows both subjective and objective interpretations,

whereas *kejian* only allows a subjective interpretation. Under the *suoyi* condition, readers need to maintain multiple interpretations when processing the propositional content, which may increase the processing load, as the maintenance of ambiguous representations in comprehension entails larger memory loads (Just & Carpenter, 1992; King & Just, 1991). In other words, it may not be the subjective connective *kejian* that facilitates processing but rather the underspecified connective *suoyi* that slows down processing.

The modal verb functions as a cue for readers that the second clause is a claim based on the argument made in the first clause, that is, the relation is subjective. Hypothetically, the processing time of the modal verb region should be shorter under the *kejian* condition compared with the *suoyi* condition because *kejian* already provides the information that S2 is subjective at the very start of the second clause. However, we did not find any effects of connective type at the modal verb region under the no stance marker conditions: The subjective connective *kejian* did not make the processing of the later modal verb region easier in comparison with the *suoyi* condition. One explanation from the view of parafoveal processing might be that readers processed the modal verb already at the subject position. If this was the case, the processing asymmetry at the subject region demonstrated the effect of modal verbs in processing subjective relations.

The longer processing times that surfaced at the connective region when the specific connective *kejian* served as an early cue of subjectivity, and at the later subject region when the connective was underspecified, indicate that the processing of subjective relations require more cognitive effort. The visual world paradigm eye-tracking study described in the Introduction (Wei et al., 2019) provides a possible interpretation as to what these cognitive efforts could be attributed to. In that study, subjective connectives in both Dutch and Chinese, in comparison with objective connectives, led to more attention to a speaker in the visual context, who was presented as a person responsible for making an argument or judgment. The extra processing time of subjective connectives in the current reading study correspond to the growing attention to the speaker introduced by subjective connectives in the visual world paradigm study. Hence, the longer reading times associated with the processing of subjectivity might be attributed to the process of tracking the speaker, who is the source of information. If this explanation is on the right track, this seems to indicate that linguistic cues such as connectives function as processing instructions that help the reader determine who the source of the information is.

Effects of stance markers

On the basis of a subjectivity account and a collocation processing benefits account, we arrived at competing hypotheses on the processing effects of stance markers in relation to the use of different connectives. The subjectivity account predicted a facilitation effect of epistemic stance markers on the processing of the specific connective *kejian* but not for *suoyi*. The collocation processing benefits account, by contrast, suggested shorter processing times for *suoyi* than *kejian* in the epistemic stance marker condition, since epistemic stance markers occur more frequently with *suoyi* than with *kejian* in corpus data.

General effects of stance markers in connective, subject, and modal verb regions

The first research question on the effect of epistemic stance markers was as follows:

> *How does the presence of epistemic stance markers affect the processing patterns at the connective region and at later regions in subjective relations with an argument-claim structure?*

On the basis of both accounts, we predicted an interplay between connective and stance in influencing the reading times of subjective relations. However, at the first two regions of the target clause (the connective region and the subject region), the type of stance marking did not change the reading time differences between the subjective connective *kejian* and the underspecified connective *suoyi*. *Kejian* required longer reading time than *suoyi* at the connective region regardless of whether a stance marker was available or not and regardless of which type of stance marker was used. The facilitation effect of

kejian compared with *suoyi* at the subject region after the connective also held for all three stance-marking conditions.

The lack of interaction effects between connective and stance in these two early regions may be due to the particular structure of the sentences in the current experiment. The processing delay associated with subjective connectives compared with objective/underspecified connectives is about establishing a subjective claim in the second segment. Epistemic stance markers influenced the processing asymmetry between subjective connectives and objective connectives in the claim-argument type of subjective relations (Canestrelli et al., 2013; Traxler et al., 1997b), because epistemic stance markers explicitly mark the first clause as a claim made by an intentional mind. However, in the current study, the stance markers only had scope over the first clause, the argument, and not over the second clause. Therefore, the stance markers did not directly mark the second segment as a claim. This might explain why the influence of stance markers on the processing differences between the two connective conditions was limited.

Despite the lack of an interaction between connective and stance, the presence of attitudinal stance markers in the preceding context led to longer reading times of both connectives compared with the condition without stance markers. Epistemic stance markers also introduced a processing delay in the same direction. This could be due to increased difficulty to integrate the second clause when there was more information in the context with attitudinal/epistemic stance markers. Since no differences were found between the two types of stance markers or between the subjective connective and the underspecified connective with regard to this effect, the longer reading times cannot be rashly attributed to the processing of subjectivity.

Most importantly, the reading times of the modal verb region in the second clause were significantly reduced by the presence of epistemic and attitudinal stance markers in the preceding clause, even though these markers did not have scope over the second clause. One possible interpretation is related to a process of tracking the source of information in the representation of subjectivity (Wei et al., 2019). At the modal verb, readers obtained a manifest cue that the second clause was a subjective claim. With either an epistemic or an attitudinal stance marker in the preceding context, the subjectivity information encoded by the modal verb can be linked to this preconstructed source of information. Accordingly, no extra effort to establish a new source of information is needed at the modal verb for either of these stance marker conditions.

The interplay between the processing times of modal verb and stance markers can also be explained in terms of validation during reading. According to Singer (2013), language processing involves an immediate validation of current input with information in prior context and world knowledge (see also Richter, 2015). Comprehenders routinely validate incoming messages (e.g., a statement with a truth condition) with what has been established in discourse (e.g., the truth condition of a prior sentence), suggesting a close relation between validation and comprehension (Richter et al., 2009). Inconsistencies in this validation process would lead to longer reading times, as shown by Cook and O'Brien (2014). In addition, the size of such a validation effect is dependent on the relation between textual input and world knowledge (Cook & O'Brien, 2014) and varies across individuals with different working memory spans and knowledge access (Singer & Doering, 2014). In this current study, a stance marker, regardless of its type, establishes a subjective context, and a modal verb appearing in the second clause can trigger a validation process of evaluating subjective information. With a stance marker in the prior context, the consistency between the subjective information encoded by the modal verb and that expressed by the stance marker would lead to faster validation compared with sentences without a stance marker, which could very likely be interpreted as objective.

In the offline acceptability test (see the data made available online https://doi.org/10.17605/OSF.IO/W9E2V), epistemic and attitudinal stance marker conditions obtained lower acceptability scores compared with the no stance marker conditions, which would normally predict slower reading times in these marked stance conditions. However, the results of the online reading experiment show the opposite effect of stance markers. Thus, these effects cannot be interpreted as the result of a difference in acceptability.

Distinctive effects of epistemic and attitudinal stance markers in the final region

Our second research question on the effect of stance markers was the following:

Do attitudinal stance markers affect processing in the same way as epistemic stance markers do?

As discussed in the previous section, we did not find a difference between epistemic and attitudinal stance markers in the influence they exerted on the processing of the first three regions of the second clause. However, at the final region we did find an interaction effect of connective and stance, as the processing of the attitudinal stance condition differed from both other stance marker conditions.

When there was no stance marker in the context, the reading times of the final region were not significantly different between the *suoyi* and the *kejian* condition. The similarity in processing times between these conditions make sense because at this final region readers are supposed to know that the relation is subjective—even in the condition with the smallest amount of linguistic instruction on subjectivity (no stance marker + underspecified *suoyi* condition)—since the modal verb provides the subjectivity information. By the final region, readers have been well informed that the second clause is a claim and that the relation is formulated in an argument-claim structure. Therefore, no extra processing time is needed under the *suoyi* condition to process the final region. This also explains why no difference was found between clauses connected by *suoyi* and those connected by *kejian* in the conditions with an epistemic stance marker in the preceding context.

By contrast, when the context contained attitudinal stance markers, which express the speaker's attitude toward the content, there was a processing advantage of the subjective connective *kejian* compared with *suoyi* at this final region. Note that this facilitative effect of attitudinal stance markers in combination with *kejian* is in line with the collocation patterns in the corpus-based study presented in Wei et al. (2020), which demonstrated that *kejian* co-occurs more often with attitudinal stance markers than *suoyi*. This effect can be explained with regard to the way argument-claim relations are structured: With an attitudinal stance marker such as *surprisingly* in (16b) in the first clause, the content of the first clause is assumed to be factual. In an argument-claim relation, the factuality of the argument makes the claim more reliable. This might explain why the argument-claim relation was probably most accessible to readers when *kejian* was used in combination with attitudinal stance markers. This is in contrast to the epistemic stance markers: The factuality of the first clause in (16a) is put in doubt by *jushuo* ("it is said"), because it is presented as an event described by someone else. If the factuality of the first clause is doubted, the conclusions drawn on the basis of that information are also doubtful, which is probably why readers did not benefit from the presence of epistemic stance markers at this stage of processing as much as they did in the attitudinal stance condition.

(16)

a. Liu Yishan **jushuo** meinian you yi zhengge yue de daixin nianjia, **suoyi/kejian** ta de gongsi yiding gei yuangong de fuli bucuo.

'**It is said (that)** Liu Yishan has one whole month vacation with salary every year, **suoyi/kejian 'so'** her company must provide good welfare for its employees.'

b. Liu Yishan **jingran** meinian you yi zhengge yue de daixin nianjia, **suoyi/kejian** ta de gongsi yiding gei yuangong de fuli bucuo.

'**Surprisingly** Liu Yishan has one whole month vacation with salary every year, **suoyi/kejian 'so'** her company must provide good welfare for its employees.'

This study focused on the processing of coherence relations consisting of two clauses. By adding stance markers to the first clause of the relation, we were able to clearly present how different stance markers influence the processing of subjectivity in coherence relations. However, the effects of such markers in a larger context might be different if there are contextual cues outside the scope of the coherence relation. Canestrelli et al. (2016) tested the influence of stance markers such as *amazingly* and *horribly* as

contextual cues preceding a coherence relation in argumentative letters to the editor, in comparison with informative news texts without such evaluative cues. These contextual cues led to longer reading times of the forthcoming sentence in general compared with the sentence embedded in the unmarked informative context. According to the Mental Space Theory, these evaluative cues activate a mental space of the author, and readers are more careful in interpreting the following sentences when this mental space is constructed (Canestrelli et al., 2016; see also Sanders et al., 2009). It is noteworthy that the evaluative cues in discourse did not cancel out the complexity associated with subjective coherence relations. In both argumentative letters with stance markers and informative newspaper texts, subjective relations required longer processing time compared with objective ones (Canestrelli et al., 2016).

Incremental sentence processing

The online reading of subjective relations marked by connectives, modal verbs, and stance markers exhibits an incremental process of sentence comprehension. Note that the subjectivity ratings of the sentences as a whole did not differ across conditions. However, to establish a claim-argument or argument-claim relation needed for the processing of subjectivity, readers use different linguistic cues throughout the sentence in online reading. Connectives set up a coherence relation between clauses. Some connectives, such as Dutch *want* and *dus* and Chinese *kejian*, also specify the degree of subjectivity of this relation by attributing the relation to a locutionary agent and emphasizing the claim-argument or argument-claim structure of the relation. Modal verbs such as *yiding* "must" and *keneng* "may" explicitly mark the second clause as a claim that is attributed to a source of information and trigger a validation process of subjective information with prior discourse. Stance markers not only express subjectivity but also specify the dimension of subjectivity that is involved: opinions as in (16a) or attitudes or feelings as in (16b).

These linguistic cues function as processing instructions by incrementally providing cues. Subjective connectives immediately trigger the construction of a subjective relation, which requires more cognitive effort. Modal verbs explicitly inform readers on the subjectivity of the relation, independent of the choice of connective. This is the point where preceding stance markers exhibit an effect: The processing of modal verbs is facilitated by both the epistemic and attitudinal stance markers irrespective of connective type. This finding shows that as long as the subjectivity is established in the preceding context, no matter how this information is expressed, epistemically or attitudinally, readers can benefit from the previously established discourse in later processing, which is consistent with the validation account substantiated by previous studies. If such information is not established in the preceding context, however, a modal verb functions as a novel cue of subjectivity and leads to a processing delay. The online reading finding also converges with the effect of modal verbs in the visual world paradigm experiment conducted by Wei et al. (2019): If the information of subjectivity provided by the context is not sufficient, modal verbs direct comprehenders' attention to a picture of the speaker.

At the end of the sentence, readers wrap up all relevant information incrementally provided in the sentence. The information on the dimensions of subjectivity conveyed by stance markers in the preceding context affects sentence processing as it instructs readers how well the forthcoming argument supports the claim in a subjective relation. If the speaker/author uses an epistemic marker to express uncertainty about the proposition in the first clause, this also weakens the strength of the argument-claim relation as a whole.

Conclusion

Parts of the collocation patterns in language use reported in Wei et al. (2020) were reflected in online processing. For instance, at the end of the sentence, readers were sensitive to the co-occurrence tendency of *kejian* and attitudinal stance markers. However, at other points the collocation patterns were not reflected in the reading times. People did not process modal verbs differently in response to

different connectives in the context, which was predicted on the basis of the distribution patterns of modal verbs and connectives. Explorations on collocation patterns are useful to obtain general distributional information on the use of linguistic cues expressing subjectivity, but future experimental research is needed to investigate whether and how distributional information is used by comprehenders in processing subjectivity and causality and to what extent they are sensitive to such information.

The current study shows that readers make use of the information provided by different markers of subjectivity incrementally to construct a discourse representation that encodes subjectivity. Meanwhile, by evaluating the factuality of the event described in the first clause, readers also evaluate how well the argument functions as a support for the claim formulated in the subjective relation. These processes, which contribute to the mental representation of subjective relations, are reflected in online reading times. Future research could explore how contextual features specifying different dimensions (e.g., factuality of an event, reliability of a speaker, style, etc.) influence the processing of causal relations.

Notes

1. English does not have a specific subjective counterpart of *kejian*, which is why we used the same connective *so* to translate the connectives in (11) and (12). In English, the subjectivity of the relation is derived from the content of the clauses.
2. A complete list of test items, a discussion of their acceptability and subjectivity ratings, and the datasets and scripts used in the analyses and details on model selection can be found online (see Data availability statement).
3. Calibration results were considered successful when all 13 calibration points were well aligned horizontally and vertically; validation results were acceptable only when Eyelink reported "Good" (average error < 1.0 degree).

Disclosure statement

No potential competing interest was reported by the authors.

Data availability statement

Data that support the findings of this study are openly available in Open Science Framework at https://doi.org/10.17605/OSF.IO/W9E2V.

Funding

Part of this article is based on Chapter 5 of the first author's Ph.D. thesis (Wei, 2018) which was supported by the Chinese Scholarship Council under grant number 2013077220042. This research also received support from the Starting Grant for Young Scholars of Peking University under grant number 7101502221.

ORCID

Yipu Wei http://orcid.org/0000-0002-0128-4098

References

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*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bednarek, M. (2006). *Evaluation in media discourse: Analysis of a newspaper corpus*. Continuum.

Bednarek, M. (2009). Dimensions of evaluation: Cognitive and linguistic perspectives. *Pragmatics & Cognition*, *17*(1), 146–175. https://doi.org/10.1075/pc.17.1.05bed

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. https://doi.org/10.1080/01690965.2012.685885

Canestrelli, A. R., Mak, W. M., & Sanders, T. J. M. (2016). The influence of genre on the processing of objective and subjective causal relations: Evidence from eye-tracking. In N. Stukker, W. Spooren, & G. Steen (Eds.), *Genre in language, discourse and cognition* (pp. 51–73). De Gruyter Mouton.

Conrad, S., & Biber, D. (2000). Adverbial marking of stance in speech and writing. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 56–73). Oxford University Press.

Cook, A. E., & O'Brien, E. J. (2014). Knowledge activation, integration, and validation during narrative text comprehension. *Discourse Processes*, *51*(1–2), 26–49. https://doi.org/10.1080/0163853X.2013.855107

Dancygier, B., & Sweetser, E. E. (2005). *Mental spaces in grammar: Conditional constructions*. Cambridge University Press.

Degand, L., & Pander Maat, H. L. W. (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). LOT.

Durrant, P., & Doherty, A. (2010). Are high-frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory*, *6*(2), 125–155. https://doi.org/10.1515/cllt.2010.006

Finegan, E. (1995). Subjectivity and subjectivisation: An introduction. In D. Stein & S. Wright (Eds.), *Subjectivity and subjectivisation: Linguistic perspectives* (pp. 1–15). Cambridge University Press.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–149. https://doi.org/10.1037/0033-295X.99.1.122

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, *30*(5), 580–602. https://doi.org/10.1016/0749-596X(91)90027-H

Langacker, R. W. (1990). Subjectification. *Cognitive Linguistics*, *1*(1), 5–38. https://doi.org/10.1515/cogl.1990.1.1.5

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. https://doi.org/10.1075/cld.4.1.03li

Li, F., Mak, W. M., Evers-Vermeul, J., & Sanders, T. J. M. (2017). On the online effects of subjectivity encoded in causal connectives. *Review of Cognitive Linguistics*, *15*(1), 34–57. https://doi.org/10.1075/rcl.15.1.02li

Lyons, J. (1977). *Semantics*. Cambridge University Press.

Pander Maat, H. L. W., & Sanders, T. J. M. (2000). Domains of use or subjectivity: The distribution of three Dutch causal connectives explained. In E. Couper-Kuhlen & B. Kortmann (Eds.), *Cause - condition - concession – Contrast: Cognitive and discourse perspectives* (pp. 57–82). Mouton de Gruyter.

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

Pit, M. (2006). Determining subjectivity in text: The case of backward causal connectives in Dutch. *Discourse Processes*, *41*(2), 151–174. https://doi.org/10.1207/s15326950dp4102_3

R Core Team. (2015). *R: A language and environment for statistical computing (Version 3.1.3)* [computer software]. R Foundation for Statistical Computing. http://www.R-project.org/

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Richter, T. (2015). Validation and comprehension of text information: Two sides of the same coin. *Discourse Processes*, *52*(5), 337–355. https://doi.org/10.1080/0163853X.2015.1025665

Richter, T., Schroeder, S., & Wöhrmann, B. (2009). You don't have to believe everything you read: Background knowledge permits fast and efficient validation of information. *Journal of Personality and Social Psychology*, *96*(3), 538–558. https://doi.org/10.1037/a0014038

Sanders, T. J. M., Sanders, J., & Sweetser, E. E. (2009). Causality, cognition and communication: A mental space analysis of subjectivity in causal connectives. In T. J. M. Sanders & E. E. Sweetser (Eds.), *Causal categories in discourse and cognition* (pp. 19–60). Mouton de Gruyter.

Sanders, T. J. M., & Sweetser, E. E. (2009). Introduction: Causality in language and cognition – What causal connectives and causal verbs reveal about the way we think. In T. J. M. Sanders & E. E. Sweetser (Eds.), *Causal categories in discourse and cognition* (pp. 1–18). Mouton de Gruyter.

Singer, M. (2013). Validation in reading comprehension. *Current Directions in Psychological Science*, *22*(5), 361–366. https://doi.org/10.1177/0963721413495236

Singer, M., & Doering, J. C. (2014). Exploring individual differences in language validation. *Discourse Processes*, *51*(1–2), 167–188. https://doi.org/10.1080/0163853X.2013.855534

Stukker, N., & Sanders, T. (2012). Subjectivity and prototype structure in causal connectives: A cross-linguistic perspective. *Journal of Pragmatics*, *44*(2), 169–190. https://doi.org/10.1016/j.pragma.2011.06.011

Thompson, G., & Hunston, S. (2000). Evaluation: An introduction. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 1–27). Oxford University Press.

Traxler, M., Sanford, A. J., Aked, L. M., & Moxey, L. M. (1997a). Processing causals and diagnostics in discourse. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*(1), 88–101. https://doi.org/10.1037/0278-7393.23.1.88

Traxler, M. J., Bybee, M. D., & Pickering, M. J. (1997b). Influence of connectives on language comprehension: Eye tracking evidence for incremental interpretation. *The Quarterly Journal of Experimental Psychology Section*, *50A*(3), 481–497. https://doi.org/10.1080/027249897391982

Van Dijk, T. A. (1982). Opinions and attitudes in discourse comprehension. In J. F. Le Ny & W. Kintsch (Eds.), *Language and comprehension* (pp. 35–51). North Holland.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press.

Vasilev, M. R., & Angele, B. (2017). Parafoveal preview effects from word N + 1 and word N + 2 during reading: A critical review and Bayesian meta-analysis. *Psychonomic Bulletin and Review*, *24*(3), 666–689. https://doi.org/10.3758/s13423-016-1147-x

Veenker, T. J. G. (2013). *The Zep experiment control application (Version 1.6.3)* [Computer software]. Utrecht Institute of Linguistics OTS, Utrecht University.

Vilkaitė, L., & Schmitt, N. (2017). Reading collocations in an L2: Do collocation processing benefits extend to non-adjacent collocations? *Applied Linguistics*, *40*(2), 329–354. https://doi.org/10.1093/applin/amx030

Wei, Y. (2018). *Causal connectives and perspective markers in Chinese: The encoding and processing of subjectivity in discourse* [Doctoral dissertation], Utrecht University. LOT. https://www.lotpublications.nl/Documents/482_fulltext.pdf

Wei, Y., Evers-Vermeul, J., & Sanders, T. J. M. (2020). The use of perspective markers and connectives in expressing subjectivity: Evidence from collocational analyses. *Dialogue & Discourse*, *11*(1), 62–88. https://doi.org/10.5087/dad.2020.103

Wei, Y., Mak, W. M., Evers-Vermeul, J., & Sanders, T. J. M. (2019). Causal connectives as indicators of source information: Evidence from the visual world paradigm. *Acta Psychologica*, *198*(Article), 102866. https://doi.org/10.1016/j.actpsy.2019.102866

Yang, J., Wang, S., Xu, Y., & Rayner, K. (2009). Do Chinese readers obtain preview benefit from word n + 2? Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(4), 1192–1204. https://doi.org/10.1037/a0013554

Zufferey, S. (2012). "Car, parce que, puisque" revisited: Three empirical studies on French causal connectives. *Journal of Pragmatics*, *44*(2), 138–153. https://doi.org/10.1016/j.pragma.2011.09.018