

**Subjectivity in Mandarin Chinese:
The meaning and use of causal connectives in written
discourse**

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6111

e-mail: lot@uu.nl
<http://www.lotschool.nl>

Cover illustration: The author's office (the place where this dissertation was written) and its surroundings, photographed by Assaf Toledo.

ISBN: 978-94-6093-147-5
NUR 616

Copyright © 2014: Fang Li. All rights reserved.

**Subjectivity in Mandarin Chinese:
The meaning and use of causal connectives in written
discourse**

Subjectiviteit in Mandarijn Chinees: De betekenis en het gebruik van
causale connectieven in geschreven tekst

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan,
in het openbaar te verdedigen
op vrijdag 29 augustus 2014
des middags te 4.15 uur

door

Fang Li

geboren op 4 augustus 1977 te Hebei, China

Promotor: Prof.dr. T.J.M. Sanders

Co-promotor: Dr. J. Evers-Vermeul

This thesis was accomplished with financial support from Huygens Scholarship, granted by Nuffic.

For my family

獻給我的家人

TABLE OF CONTENTS

CHAPTER 1	1
Introduction	
CHAPTER 2	25
The relevance of subjectivity for the characterization of Mandarin causal connectives: A literature review	
CHAPTER 3	53
Subjectivity and result marking in Mandarin: A corpus-based investigation	
CHAPTER 4	95
On the subjectivity of Mandarin reason connectives: Robust profiles or genre-sensitivity	
CHAPTER 5	123
On the online effects of subjectivity information encoded in causal connectives	
CHAPTER 6	149
Conclusion and discussion	
REFERENCES	175
SAMENVATTING IN HET NEDERLANDS	189
ACKNOWLEDGEMENTS	203
CURRICULUM VITAE	209

CHAPTER 1

Introduction

1.1 The notion of subjectivity in language

Language is generally used as a vehicle to communicate messages. The message at times contains purely objective propositions, i.e., about events or facts that can be observed in the outside world. At other times, the message also involves the speaker's point of view. In the words of Kristeva (1989: 11), the speaker sometimes "imprints a specific seal" upon the message, even without being aware of it. Imagine a communicative context in a geography class. The teacher points at a map of the world, and talks about cities in different countries. She may state an established fact about a particular city by uttering a sentence such as (1), and perhaps she also expresses her personal evaluative judgment of that city by uttering a sentence such as (2).

- (1) Beijing is the capital of China.
- (2) Beijing is a wonderful city.

In the literature, the distinction between (1) and (2) is referred to in terms of the degree of subjectivity. It is generally recognized that (2) is associated with a higher degree of subjectivity than (1).

The notion of subjectivity was raised with the development of pragmatics, functional linguistics, and cognitive grammar by western linguists in the 1990s (see Stein & Wright 1995). Among these linguists, Lyons, Langacker, and Traugott deserve our special attention. They represent three major trends in the western theory of subjectivity. Lyons (1977) takes the notion of subjectivity as the speaker's expression of 'self' in an utterance. The term *self* refers to the speaker's cognition, perception, affect, attitude and intention with regard to the propositional content of the utterance. Sentence (2), for example, is considered to convey a high degree of subjectivity because the speaker expresses herself, i.e., her attitude towards the city of Beijing, which is clearly positive. By contrast, (1) does not concern the speaker's expression of *self*. Therefore, it is considered objective.

Very similar to Lyons, Traugott (1995, 2010) views subjectivity as a property of being speaker-related. Subjectivity is defined as the "relationship to the speaker and the speaker's beliefs and attitudes" (Traugott 2010: 30). Traugott approaches subjectivity from a diachronic perspective, paying attention to the process in which meaning becomes more and more dependent on the speaker's attitude towards the proposition, i.e., subjectification. For example, Traugott (1995: 41) mentioned that, in English,

an inference has gradually been semanticised in *while* on the basis of its original temporal meaning ‘during’, signaling the speaker’s perspective, i.e., “surprise concerning the overlap in time or the relations between event and ground”, which led to the adversative, concessive meaning ‘although’ (see also Keller 1995).

Langacker (1990) defines subjectivity in terms of on/off-stage conceptualization. In his view, the speaker conceptualizes objects or events from either an on-stage perspective or an off-stage perspective. The off-stage conceptualization refers to the situation in which the speaker’s perspective is implicit in the utterance, as in (2). Langacker considers the off-stage conceptualization maximally subjective on the grounds that the speaker is so absorbed in the perceptual experience that she loses all awareness of self (1990: 7). Compare (3) with (2).

- (3) I think Beijing is an amazing city.

Sentence (3) exemplifies the on-stage conceptualization, which refers to the situation in which the speaker is explicitly referred to as a conceptualizer in the utterance. In (3), the speaker is linguistically realized by the pronoun *I*. Langacker suggests that an utterance is objectified when the speaker (as a conceptualizer) is put on stage and becomes observable. Accordingly, (3) expresses a lower degree of subjectivity than (2). Nevertheless, (3) is associated with a higher degree of subjectivity than (1). After all, (3) involves the speaker’s attitude or perspective whereas (1) does not.

Langacker, Lyons, and Traugott mainly focus on the degree of subjectivity expressed by single, isolated clauses. However, subjectivity also resides in the coherence relation holding between connected clauses (Degand & Pander Maat 2003; Pander Maat & Sanders 2000, 2001; Sanders & Spooren 2009, 2013; Sweetser 1990; Zufferey 2012). Discourse analysts have drawn our attention to the distinction between relations that can be observed in the real world, such as (4), and relations that arise in one’s mind, as exemplified by (5).

- (4) Temperatures were below minus ten degrees for more than a month. As a result, many kingfishers died last year.
 (5) The lights in the house are off, so nobody is at home.

Sentences (4) and (5) are instances of causal coherence relations, which are the focus of the present study. The causal connection expressed in (4) exists between observable facts in the outside world. In (5), however, an observed fact *the lights in the house are off* gives rise to the conclusion *nobody is at home* in the speaker’s mind. Discourse analysts propose to analyze this kind of difference in terms of subjectivity. Essentially, a causal relation is

considered subjective when some thinking entity is involved in the construction of the relation. This thinking entity has been termed ‘Subject of Consciousness’ (henceforth SoC) – an animate subject, a person, whose intentionality is conceptualized as the ultimate source of the causal event, be it an act of reasoning or some real-world activity (Pander Maat & Sanders 2001: 251). Sentence (4) does not involve an SoC: the causal relation has an origin in a different source, located in the real world. Therefore, the causal relation expressed in (4) is objective. Meanwhile, it is not difficult to see that there is an SoC in (5): the causal relation holds between an argument and a conclusion in the speaker’s mind, though the speaker-SoC is not linguistically realized in the utterance. We can conclude that the causal relation expressed in (5) is associated with a higher degree of subjectivity than the one expressed in (4).

As a point to note, the SoC can be explicitly mentioned, or put on stage, as well. For example, (6) is a case of on-stage conceptualization. According to Langacker’s sense of subjectivity in terms of on/off-stage conceptualization, (6) is associated with a lower degree of subjectivity than (5).

(6) The lights in the house are off, so I think nobody is at home.

In addition, in free indirect speech such as (7), the narrator often introduces the perspective of other persons.

(7) The lights in the house were off, so nobody was at home, John thought.

In (7), John is the SoC responsible for the causal relation. A causal inference *nobody was at home* crossed John’s mind, when he saw *the lights in the house were off*. The narrator simply reports this causal connection arising in John’s mind. In the terminology of a mental space analysis, the causal relation expressed in (7) does not concern the speaker’s epistemic space (see Sanders & Redeker 1996; Sanders, Sanders & Sweetser 2009).¹ Only *John’s* epistemic space is at stake. The notion of subjectivity has been thus extended beyond speaker-relatedness: it could be character-related as well.

¹ Sanders, Sanders, and Sweetser (2012) suggest that many cases of free indirect speech are more complicated than (7). They concern blending of spaces, by which the speaker identifies with the (non-speaker) SoC and sees through the SoC’s eyes. For example, there is a blending of the epistemic space of the speaker/narrator (implicit SoC) with the space of the narrative subject *Jan* in the Dutch sentence *Jan zag dat de buren niet thuis waren* (‘Jan saw the neighbors were not at home’), *want hun licht was uit* (‘because their lights were out’).

1.2 Research questions

A growing number of studies on subjectivity have focused on the class of linguistic expressions generally referred to as causal connectives. In many languages, there is more than one causal connective at the language user's disposal. In English, for example, causal connectives include *because*, *since*, *so*, *as a result*, by means of which the speaker can explicitly mark that the connected clauses are causally (rather than merely additively) related to each other. Usage-based studies into causal connectives of Dutch, French, German, Polish, and Mandarin Chinese (abbr. Chinese) have pointed to a common phenomenon: causal connectives show some overlap in their usage in certain contexts, but are not always interchangeable without changing the meaning of the sentence (Deng 2007; Guo 2006; X. Li 2009; Zhao 2003; for European causal connectives, see the overview in Sanders & Sweetser 2009 and in Stukker & Sanders 2012). For instance, the overlap between the use of *so* and *as a result* is apparent in (8), whereas (9) reveals that these two causal connectives are not fully identical.

- (8) It rained heavily the whole day, so / as a result the picnic was canceled.
- (9) Peter wants to see you, so / *as a result he will probably ask you for an appointment.

Findings from a group of European studies have consistently shown that causal connectives are associated with varying degrees of subjectivity, and that is why the speaker/author prefers one causal connective over another under given circumstances such as (9) (Degand & Pander Maat 2003; Evers-Vermeul, Degand, Fagard & Mortier 2011; Pit 2003; Spooren, Sanders, Huiskes & Degand 2010; Zufferey 2012). In (9), the relation holds at the argument level, so speakers prefer to use a linguistic marker that is encoded with a higher degree of subjectivity (e.g., *so*). Similar to *so*, French *car* and *puisque*, and German *denn* (all translated by *because*) mark higher degrees of subjectivity, and thus they are typically used to express relations as in (9), which are constructed with high involvement of the speaker or some other SoC (see Stukker & Sanders 2010a, 2012). In contrast, French *parce que* and German *weil* (both translated by *because*) mark lower degrees of subjectivity, and are typically used to express objective causal relations as manifested in (8). Objective causal relations are not intended for argumentation, but focus on describing connections between events or circumstances observable in the world.

It is assumed that linguistic categorization reflects human cognition (Lakoff 1987; Lakoff & Johnson 1999). Accordingly, subjectivity has been proposed to be cognitively relevant, in that it organizes our knowledge of

causality and use of causal connectives, at least in several European languages (Canestrelli 2013; Sanders and Spooren 2009, submitted; Stukker & Sanders 2012; van Veen 2011).

Findings from discourse processing and language acquisition have provided support for the idea that subjectivity is cognitively relevant. It is found that causal relations or connectives that are associated with different degrees of subjectivity show different patterns in online processing and language acquisition. Psycholinguistic work has discovered that objective causal relations or connectives are processed faster than subjective causal relations or connectives during online reading (Canestrelli 2013; Canestrelli, Mak & Sanders 2013; Noordman & De Blijzer 2000; Traxler, Bybee & Pickering 1997a; Traxler, Sanford, Aked & Moxey 1997b). Research on child language has found that objective causal relations are acquired before subjective causal relations (Evers-Vermeul 2005; Evers-Vermeul & Sanders 2011; Spooren & Sanders 2008; van Veen 2011). These findings from language acquisition and discourse processing can be explained by the subjective complexity hypothesis (Sanders 2005): subjective relations are cognitively more complex than objective relations, which is why the former take longer to acquire and process than the latter.

Like other languages, Chinese displays a rich lexical repertoire of causal connectives. In the literature, the Chinese lexicon of causality has been studied with different approaches. Consequently, different accounts have been produced as to the way in which one causal connective differs from another. A fundamental question that the present dissertation takes interest in is whether it is feasible to study the full set of Chinese causal connectives with the subjectivity approach rooted in cognitive linguistics. In other words, we investigate the extent to which the alleged cognitively-plausible subjectivity account for connective use is generalizable to causal connectives in Mandarin discourse. As a first step in testing the generalizability of the subjectivity account, a review of Chinese literature on causal connectives and causal coherence relations will be presented. The research questions for this literature review are as follows.

Research question 1

How do Chinese linguists define subjectivity? Can their approaches be related to western approaches of subjectivity?

Research question 2

Which analytical categories have been used in previous studies on Chinese causal connectives? Can these categories be related to analyses in terms of subjectivity, and if so, how?

Research question 3

Do Chinese and European studies address the issue of subjectivity in causality with similar methods?

European languages show clear differences in the way they divide the domain of causality according to subjectivity. For example, in English there is a causal connective *because* that is very general in use, whereas several Dutch causal connectives are specific in meaning and use. Existing findings based on English data suggest that the demarcation of subjective and objective categories is realized by cue phrases rather than connectives (Knott & Dale 1994; Knott & Sanders 1998): *for that reason* (objective) and *it follows that* (subjective). In Polish, causality is most typically expressed via epistemic or intersubjective construals marked with connectives *bo* 'because' and *to* 'then', and it seldom focuses on objective relations in the real world (Dancygier 2009). This type of observation gives rise to the following question: How is causality categorized in Mandarin Chinese, a language that is typologically different from the European languages? We approach this question by conducting a corpus-based investigation into the meaning and use of Chinese causal connectives in terms of subjectivity. The research question for the corpus-based investigation is as follows.

Research question 4

Do Chinese causal connectives show systematic variation in terms of the degree of subjectivity they encode? If so, how? Are there language-specific properties in this respect?

Another important issue the present corpus-based research aims to address is related to genre, conventionally defined as a recognizable communicative event characterized by a set of communicative purposes identifiable and mutually understood by the members of the discourse community (Swales 1990; Trosborg 1997). This means that although the writer has a lot of freedom to use linguistic resources, she must conform to certain standard practices within the boundaries of a particular genre (see Bhatia 1993: 14). Swales (1990: 58) claims that genre, characterized by some set of recognizable communicative purposes, shapes the schematic structure of the discourse and constrains choices of content and style. These ideas give rise to the following question: Does genre have an impact on the meaning and use of causal connectives, in accordance with the degree of subjectivity that the text (or text category) is intended to express given its communicative purpose?

Genre-specific properties of connective use have not yet been investigated systematically, although studies on French, German, and Dutch causal connectives seem to indicate that the distribution of connectives over objective and subjective causal categories is likely to vary in relation to the genre (Degand & Pander Maat 2003; Pit 2003; Zufferey 2012). Subjective causal connectives such as French *car*, German *denn*, and Dutch *want*, which are all roughly translated as 'because', display consistent usage patterns

across text types such as newspapers, novels, and periodicals. By contrast, usage patterns of their objective counterparts, French *parce que*, German *weil*, and Dutch *omdat*, are less consistent across these text types (Stukker & Sanders 2012). Results from some Chinese studies also seem to imply that genre might play a role in shaping the meaning and use of causal connectives. Guo (2006) and Xing (2002) have reported that certain causal connectives have higher-frequency distributions in a specific genre than other causal connectives in Chinese discourse. The question is: if the connective is used in its less typical context, does it express different relations that are associated with different degrees of subjectivity? A systematic analysis is needed. Therefore, another important issue our corpus-based research aims to address is as follows.

Research question 5

Do subjectivity profiles of Chinese causal connectives vary across genres?

In psycholinguistic terms, connectives can be considered as operating instructions for interpretation: they 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). It is found that the presence of an appropriately used causal connective speeds up the processing of an immediately following discourse segment that is causally related to the previous segment: participants spent less time reading a clause when it was introduced by a causal connective than when the connective was absent (see Cozijn 2000; Haberlandt 1982; Millis & Just 1994). Furthermore, previous experimental studies have shown that subjective and objective causal connectives in Dutch show different patterns during online processing of causally related discourse (Canestrelli 2013; Canestrelli et al. 2013). Overall, subjective causal connectives generated longer processing times than objective causal connectives. This gives rise to the sixth research question of the present dissertation.

Research question 6

How are differences in degree of subjectivity between Chinese causal connectives reflected in online discourse processing?

In sum, this dissertation aims to address these six research questions, with an initial theoretical exploration in the form of a literature review, followed by corpus-based quantitative studies on the use and function of causal connectives in written Mandarin discourse and a processing experiment using eye-tracking. It is expected that the integration of corpus-linguistic work with psycholinguistic research can produce new insights into

the way in which subjectivity constrains the use and function of Chinese causal connectives, and into the universal and language-specific properties associated with this use.

1.3 The scope of causality to be covered

The notion of causality is in fact a cover term for a set of causal relations. As will become clear soon, opinions can diverge over the exact types of relations that causality denotes. Therefore, the first task is to define the scope of causality this dissertation is going to deal with. Given that the present dissertation focuses on Chinese expressions of causality, this section starts out with insights offered in the literature for categorizing causal coherence relations in Chinese discourse, and makes a link to relevant western classifications later on.

1.3.1 The general view of causality and Xing's (2001) taxonomical work

The causal relation is always referred to as a key component of coherence relations in the system of Chinese complex sentences, in spite of the variety of taxonomies of coherence relations proposed in the literature. The traditional *erfen fa* 'dichotomous approaches' draw a direct division between *yinguo* 'causal' and *fei yinguo* 'non-causal' relations (Huang 1990; Zhang 1959). Some other dichotomous approaches draw a division between *lianhe* 'coordination' and *pianzheng/zhucong* 'subordination' on the basis of structural variance, and then highlight the causal relation as a salient subcategory of the subordination relation (Ding et al. 1999; Gao 1948; Hu 1995; Li & Liu 1957; Wang 1957). The *duofen fa* 'multiple-division approaches' (the so-called *zhi cheng fa* 'direct naming approaches') classify complex sentences into ten or more types of relations, among which the *yinguo guanxi* 'causal relation' is always considered a crucial type, along with other types of relations, such as *tiaojian* 'conditional', *rangbu* 'concessive', *zhuanzhe* 'adversative', *chengjie* 'contingency', *mudi* 'purpose' relations and so on (Luo 1999; Lü & Zhu 1952; Shao 2001; Xu, Du & Peng 1987).

One of the most detailed studies on Chinese complex sentences is Xing's taxonomical work (2001), which distinguishes between three basic types – *yinguo* 'causal', *zhuanzhe* 'adversative', and *binglie* 'additive' – as demonstrated in Figure 1. Each of these three basic types is further classified into subcategories. The causal type is categorized into *yinguo ju* 'cause-and-effect sentence', *tuiduan ju* 'inferential sentence', *jiashi ju* 'hypothetical sentence', *tiaojian ju* 'conditional sentence', and *mudi ju* 'purpose sentence', according to the similarities and differences in the logical relation.

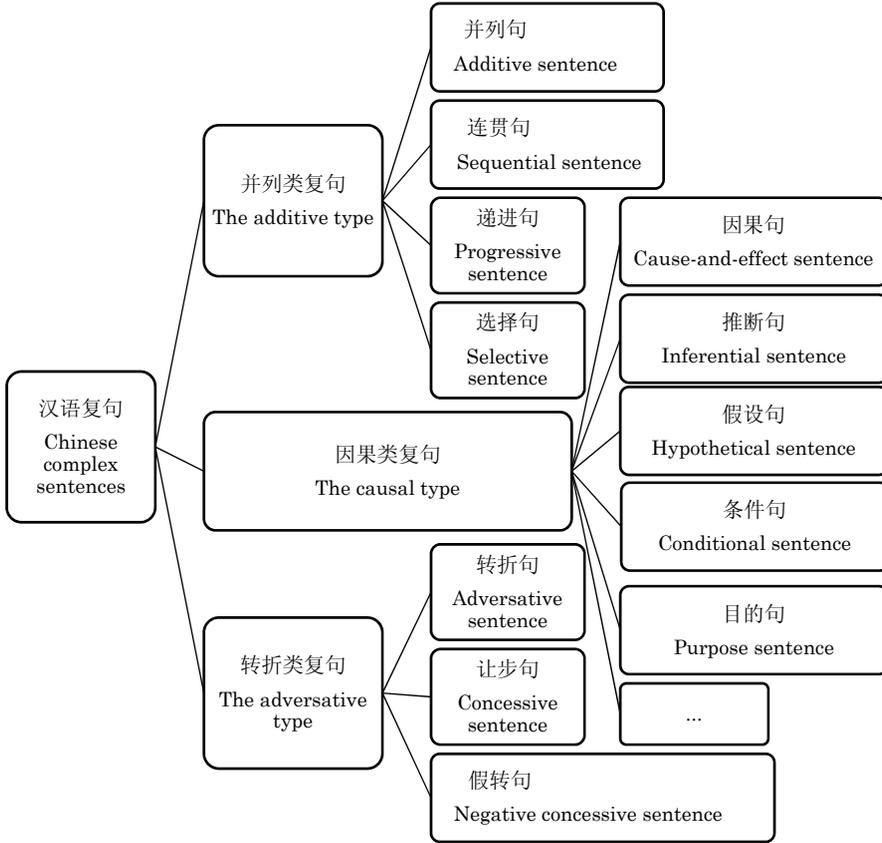


Figure 1. Xing's (2001) trichotomous taxonomy of Chinese complex sentences

In fact, Xing (2001) considers the taxonomy of causal complex sentences as an open system. In his view, new categories can be added to it at any time when valid linguistic evidence can be found to instantiate them.

According to Xing (2001), a complex sentence is considered causally related provided that it expresses a coherence relation of one thing leading to another. This is the broad sense of causality from Xing's point of view, on the basis of which five types of causal complex sentences are distinguished (see Figure 1), exemplified by sentences (10) through (14) (taken from Xing 2001: 39).

- (10) *Yinwei you ren zhao ta, ta hen gaoxing.*
'Because he was wanted by someone, he was very happy.'

- (11) **Jiran** *you ren zhao ta, ta yiding hen gaoxing.*
 ‘Since he was wanted by someone, he must have been very happy.’
- (12) **Ruguo** *you ren zhao ta, ta zhun hui hen gaoxing.*
 ‘If he is wanted by someone, he will certainly be very happy.’
- (13) **Zhiyao** *you ren zhao ta, ta ding hui hen gaoxing.*
 ‘Provided he is wanted by someone, he will surely be very happy.’
- (14) *Yiding yao you ren zhao ta, yibian rang ta gaoxing gaoxing.*
 ‘Be sure to ask someone to see him, **so as to** make him happy.’

A ‘cause-and-effect sentence’ such as (10), the full name of which is *shuomingxing yingguo ju* ‘descriptive causal sentence’, describes a relation between one event *he was wanted by someone* and a second event *his feeling happy*, where the second event is understood as the consequence of the first. This is actually what Xing considers as the narrow sense of causality. Xing (2001: 40) refers to such cause-and-effect sentences as the prototypical causal complex sentence, and claims that the type of relation expressed by cause-and-effect sentences represents the causal relation in the strictest sense.

An inferential sentence such as (11), the full name of which is *jushixing yingguo tuiduan ju* ‘inferential causal sentence according to facts’, infers the consequence on the basis of facts and rationality. The inferential sentence is different from the cause-and-effect sentence in that the former is not entirely factual whereas the latter is. In particular, the consequent of the former is always an inference rather than a fact. Crucially, rationality is involved in the process of inference-making. Nevertheless, as Xing observes, this type of inferred causal connection is rooted in evidence/facts. For example, in (11) the speaker grounds his inference *he must have been very happy* in the fact *he was wanted by someone*.

A hypothetical sentence such as (12) is also inferential, but it is different from the inferential sentence because the inference involved (e.g., *he will certainly be very happy*) is made on the basis of a supposition (e.g., *he is wanted by someone*) rather than on the basis of an established fact. That is why the full name of the hypothetical sentence is *jiashexing yingguo tuiduan ju* ‘inferential causal sentence based on hypothetical premise’.

A conditional sentence such as (13) is inferential, too. The inferred consequence *he will surely be very happy* only holds under a certain condition *he is wanted by someone*. It is called *tiaojianxing yingguo tuiduan ju* ‘inferential causal sentence based on conditional premise’ in full.

A purpose sentence such as (14), *mudixing yingguo yinhan ju* ‘purpose sentence with implicit causality’ in full, expresses a relation between a goal *making him happy*, and a certain action *asking someone to see him* as an instrument to achieve the goal. Xing (2001) suggests that a causal connection

underlies the goal and the instrumental action. From Xing's point of view, the underlying causal connection is not yet realized, but is looked forward to by the speaker.

1.3.2 Other major taxonomies of Chinese causal complex sentences

In the literature, causality is often defined in either a broad sense or a narrow sense, or both, by Chinese linguists (see Rong 2011 for an overview). As was introduced in the previous section, Xing (2001) provides a definition of causality in a sense broad enough to incorporate five types of causal relations. Many other studies (Dong 1999; Fan 1998; Huang 1990; Xing 2001; Zhang 1959) also define causality in a very broad sense in a similar way to Xing (2001). For example, Fan (1998) provides a taxonomy that contains *yiban yinguo ju* 'general cause-and-effect sentence' (equivalent to cause-and-effect sentence in Xing's terms), *tuilun yinguo ju* 'inferential causal sentence' (equivalent to inferential sentence in Xing's terms), and *mudi yinguo ju* 'purpose causal sentence' (equivalent to purpose sentence in the terms of Xing), but does not incorporate hypothetical sentences and conditional sentences into the causal system. In the narrow sense, Xing (2001) suggests that causality is simply the relation expressed in cause-and-effect sentences. The following sections serve to discuss taxonomies of causal complex sentences in Mandarin Chinese proposed by other representative works in the literature.

Taxonomies built upon traditional Chinese approaches

Before the comprehensive taxonomical work by Xing (2001) came out, most Chinese linguists classified causal complex sentences into only two types: *shuoming yinguo fuju* 'descriptive causal complex sentence' and *tuilun yinguo fuju* 'inferential causal complex sentence' (Hu 1995; Liu 1986; Wang 1985; B. Zhang 2008; Zhao 1985). Sentences (15) and (16), taken from Hu (1995: 366-367), serve to illustrate these two categories, but without the use of causal connectives.

(15) *Tian xia zheme da de yu, women dou chidao le.*

'It was raining so heavily, we were all late.'

(16) *Haizimen you la you tu, yiding shi shiwu zhongdu le.*

'The children suffered from diarrhea and vomiting, they must have eaten poisonous food.'

A descriptive causal complex sentence such as (15) reports a causal connection that exists in the observable world: a bad weather condition, as it always does in the natural world, hampered *us* from arriving on time. An inferential causal complex sentence such as (16) involves an inference and an

associated reasoning process: from the symptoms of diarrhea and vomiting, the speaker infers that food poisoning must have occurred to the children. These two types are identical to what Xing (2001) terms as cause-and-effect sentence and inferential sentence, respectively.

Lü (1982) proposes to distinguish between three types of reason: reason for fact, reason for action, and reason for inference. Sentences (17) – (19), taken from Lü (1982: 388), exemplify the three types of reason.

- (17) *Yinwei tian leng, gang li de shui dou jie le bing.*
 ‘**Because** it is cold, the water inside the jar is frozen.’
- (18) *Yinwei tian leng, wo you ba maoxianyi chuan shang le.*
 ‘**Because** it is cold, I put on the sweater again.’
- (19) *Tian yiding hen leng, yinwei gang li de shui dou jie le bing le.*
 ‘It must be very cold, **because** the water in the jar has been frozen.’

Lü’s taxonomy is consistent with the traditionally-defined sense of causality as being either descriptive or inferential, but subdivides descriptive causality into reason for fact and reason for action on the basis of the property of the consequent segment. Obviously, the consequent in (17) *the water inside the jar is frozen* is an observable fact, and the consequent in (18) *I put on the sweater again* is an action. These two cases taken together can be equated with the category of descriptive causality defined in the literature, given that both cases concern causal connections existing in the observable world. The consequent in (19) *it must be very cold* is an inference. In view of the evidence *the water in the jar has been frozen*, the speaker draws the inference about the temperature. In this sense, the category of reason for inference, as exemplified by (19), is identical to the category of inferential causality.

Taxonomies built upon western-originated approaches

Recently, some linguists have attempted to characterize Chinese causal complex sentences with approaches developed on the basis of western linguistic data. Representative works are Shen (2001, 2003, 2008) and Wong (2005). As will become clear, the taxonomies provided by these studies cohere with each other, and partially coincide with the taxonomies built upon traditional Chinese approaches. Moreover, the western-origin approaches to causality have brought another causal category to our attention, i.e., the speech-act category.

Shen (2001) has introduced Sweetser’s (1990) ‘domain theory’ to the Chinese scholarship. Although Shen’s observations are based on constructed examples rather than corpus data, he has managed to show that Chinese causal complex sentences display three domains of causality. Shen (2001, 2003, 2008) claims that the domain theory is useful in accounting for the

logical relations in Chinese causal complex sentences in a systematic way. Sentences (20) – (22), taken from Shen (2008: 405-406), are presented as instances of the content domain, the epistemic domain, and the speech-act domain, respectively.

- (20) *Zhangsan huilai le, yinwei ta hai ai Xiaoli.*
 ‘Zhang San has come back **because** he still loves Xiaoli.’²
- (21) *Zhangsan hai ai Xiaoli, yinwei ta huilai le.*
 ‘Zhang San still loves Xiaoli, **because** he has come back.’
- (22) *Zhangsan hai huilai ma? Yinwei Xiaoli zai deng ta.*
 ‘Will Zhang San come back still? **Because** Xiaoli is waiting for him?’

According to Shen, a causal relation in the content domain, such as in (20), concerns real-world causality holding between events (e.g., *Zhang San loves Xiaoli* and *Zhangsan has come back*) in the observable world. Relations in the epistemic domain, such as (21), involve the speaker’s knowledge (e.g., *Zhang San coming back*) as the basis for a logical conclusion (e.g., *Zhang San still loves Xiaoli*). For relations in the speech-act domain, the reason clause provides justification for the speech act embodied in the main clause. In the case of (22), the speaker asks whether Zhang San will come back and justifies this question (a type of speech act) with the reason clause *Xiaoli is waiting for him*, i.e., “I am asking you the question Q because P”.³

It should be apparent that the distinction between the content causal domain and the epistemic causal domain actually mirrors the one between descriptive causality and inferential causality proposed by the traditional Chinese studies represented by Hu (1995) and Xing (2001).

Following the approach developed on the basis of the English data by Quirk et al. (1985), Wong (2005) has shown that Chinese causal complex sentences can be classified into four direct reason relationships (see Figure 2). It bears noting that Wong does not report any instance of what Quirk et al. term an ‘indirect reason relationship’ (or, in Sweetser’s terms, the speech-act domain). This is not very surprising, because Wong uses a written corpus, and the speech-act domain has been shown to occur more frequently in spoken language (Spooren, Sanders, Huiskes & Degand 2010).

² In (20), there is a comma before the connective in the original Chinese sentence, but this comma is removed in the English translation. This is because according to Sweetser (1990) the commaless use of *because* secures a content reading. In her view, if a comma is inserted before the connective *because*, the comma intonation at the end of the main clause forces the alternative reading, wherein “Zhang San’s coming back” is asserted.

³ P and Q refer to the propositional content of the antecedent and that of the consequent, respectively.

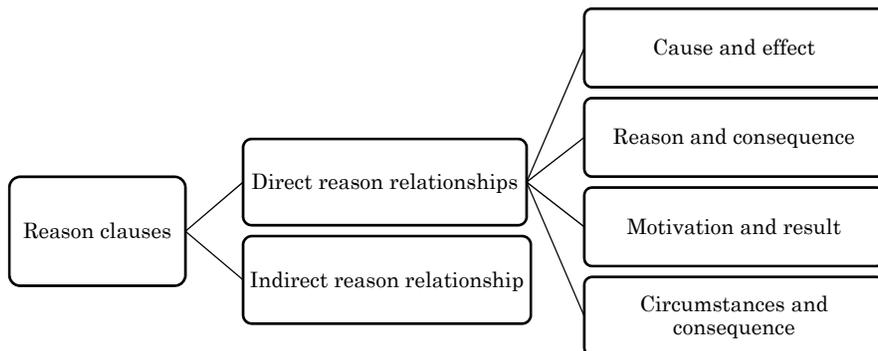


Figure 2. Taxonomy of Chinese reason clauses by Wong (2005) on the basis of Quirk et al.'s (1985) distinction between direct and indirect reason relationships

Sentences (23) – (26), taken from Wong's (2005: 135-140) collection of corpus data, serve to clarify the different types of direct reason clauses: cause and effect (23), reason and consequence (24), motivation and result (25), circumstances and consequence (26).

(23) *Yoyu tamen shi jinqin, bu neng lianyin, gongyuan wu yi zhi xiao hu chushi.*

'As the tigers are close relatives and they cannot mate with each other, no tiger cub has been born in the zoo yet.'

(24) *Renmen dui keguan shiwu guilü de renshi jiran shi ge guocheng, bu keneng yixiazi wancheng, name, zai gongzuo zhong yibosanzhe zongshi nanmian de.*

'Since one's perception of the world is accumulated as time goes by, in the course of learning how the world works, it is very likely to come across hurdles of different sorts.'

(25) *Yoyu zhang de chou, wo yixiang bu xihuan zhao jingzi.*

'Because I am not good-looking, I do not like looking at myself in the mirror.'

(26) *Malaixiya hangkong gongsi yinwei zhengfu suojian kaizhi he jiesheng waihui, zheng jihua jiang 20 jia boyin feiji de jiaohuo riqi tuichi 5 nian.*

'Because the Malaysian government reduced its expenditure to save its reserves of foreign currencies, Malaysia Airlines is currently planning to postpone the delivery date of twenty Boeing aircraft for five years.'

The cause-and-effect relation (23) describes a natural connection between facts in the real world: the situation that the tigers in the zoo cannot mate

with each other naturally leads to the fact that no tiger cub has been born in the zoo. The reason-and-consequence relation involves the speaker's inference of a connection. In (24), the speaker infers that the likelihood of coming across hurdles is high on the basis of his knowledge that one's perception of the world has to be accumulated through the time. The motivation-and-result relation expresses the result of the intention of an animate being. In (25), the fact that the speaker is not good-looking leads to the fact that she does not often look at herself in the mirror. Note that, in this causal relation, the speaker's intention not to face herself (as an ugly girl) plays an important role, which makes (25) different from (23).

The circumstances-and-consequence relation is a relationship between a premise and a conclusion. However, the examples given by Wong (2005), such as (26), do not count as good examples. In (26), the consequent "planning to postpone the delivery date" does not typically represent a conclusion. Instead, it can be taken as the concrete action of "planning". In that case, (26) would be better considered as an instance of the motivation-and-result relation. The government's expenditure-reduction policy has motivated those in charge of Malaysia Airlines to make detailed plans on how to postpone the delivery date of twenty Boeing aircraft for five years. Since no good Chinese examples can be found in Wong (2005) to illustrate the so-called circumstances-and-consequence relation, we can directly borrow an English example from Quirk et al. (1985: 1104), that is, (27).

- (27) Seeing that it is only three, we should be able to finish this before we leave today.

No connective is used in (27), but the relation is self-evident. The speaker draws the conclusion *we should be able to finish this before we leave today* from the premise *it is only three*.

A link can be established between this taxonomy by Wong (2005) and Shen's taxonomy. Wong's cause-and-effect relation and the motivation-and-result relation both capture the essence of what Shen characterizes as the content domain. In both cases, we see connections in which one fact leads to another in the outside world. As for the reason-and-consequence relation and the circumstances-and-consequence relation, either an inference is made on the basis of one's knowledge, or a conclusion is drawn from a certain premise. In both cases, the speaker is actively involved in inferring or arguing towards a causal connection in the situational context under consideration. In this sense, both can be related to what Shen characterizes as the epistemic domain.

In addition, Quirk et al.'s indirect reason relationship is equivalent to Sweetser's (1990) speech-act domain. According to Quirk et al., in an indirect reason relationship, "the reason is not related to the situation in the matrix

clause but is a motivation for the implicit speech act of the utterance” (1985: 1104). This is precisely how Sweetser defines the speech-act causal domain. In the case of (28), taken from Quirk et al. (1985:1104), the speaker asks about the position of the files, and motivates this speech act with the reason clause *as you’re in charge*, i.e., “I ask you this question because you are in charge of things related to this question”.

(28) As you’re in charge, where are the files on the new project?

1.3.3 The scope of causality to be dealt with in this dissertation

The relationship between major taxonomies of causal complex sentences presented in the previous sections is summarized in Figure 3.

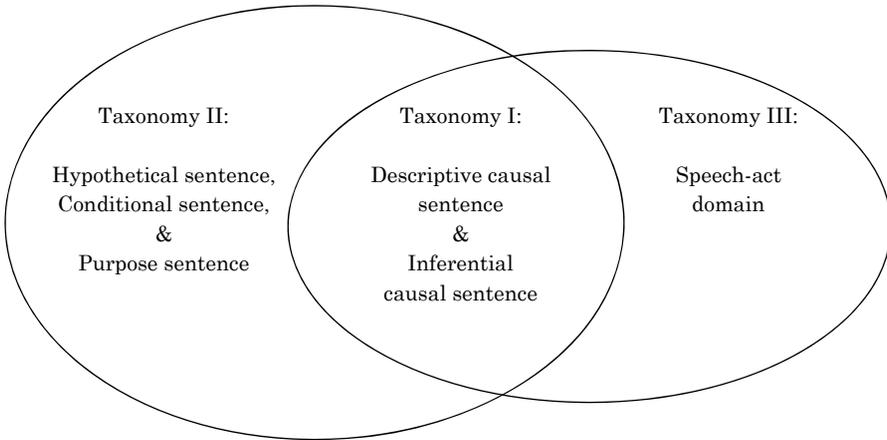


Figure 3. The relationship between previous taxonomies of causality

Taxonomy I, represented by Hu (1995), adheres to the sense of causality defined by many traditional Chinese studies (see also Liu 1986; Wang 1985; B. Zhang 2008; Zhao 1985). It contains purely descriptive causal sentences and inferential causal sentences. Taxonomy II, represented by Xing (2001), is created in accordance with the view of causality in a very broad sense (see also Dong 1999; Fan 1998; Huang 1990; Xing 2001; Zhang 1959). Xing (2001) proposes to incorporate all sentences expressing the relation of one thing leading to another into the causal system. Therefore, in his study, causal complex sentences are taken to include hypothetical, conditional, and purpose sentences, apart from those contained in Taxonomy I. Taxonomy III, represented by Shen (2001), is built upon western views of causality (see also Wong 2005). It has brought to our attention causal relations containing

speech acts, i.e., the speech-act domain or the so-called indirect reason relationship.

It is apparent from Figure 3 that Taxonomy II and III partially overlap with each other. The overlap between them is precisely Taxonomy I. That is, all three taxonomies built upon Chinese approaches as well as the western-originated approaches have drawn apart the causal connections already realized in the real world (i.e., the descriptive causal sentence/content domain) and the causal connections inferred by the speaker (i.e., the inferential causal sentence/epistemic domain).

At this point, the question arises as to how the scope of causality should be defined for the current dissertation. Primarily, the causal categories covered by Taxonomy I should be included in the scope of this study because the set of causal categories in question is indisputably considered to be an essential part of the causal system in all three taxonomies that have been observed from the literature. Furthermore, this dissertation will also cover the speech-act domain, which has not yet received much attention from Chinese linguists. In sum, the scope of this dissertation is set to encompass descriptive versus inferential causal sentences, complemented by the speech-act causal domain that was first described in the western literature. As was noted earlier, there are apparent correspondences between the descriptive causality and content causal domain on the one hand, and between the inferential causality and epistemic domain on the other. Hence, the scope of causality defined above covers the content, the epistemic, and the speech-act domains. For the sake of consistency and simplicity, I hereafter use Sweetser's terms to refer to these three types of causals. In the remainder of this section, I probe into conceptual factors that bind these three types (and not others) together, and, in the meantime, explain why hypothetical sentences, conditional sentences, and purpose sentences are excluded from the present study.

Causal complex sentences in the three domains have something in common with hypothetical, conditional, and purpose sentences. They all express a relation of implication: a certain antecedent P leads to a certain consequent Q. This relation of implication ($P \rightarrow Q$) can be formulated in two orders: either P precedes Q (e.g., "Because S_1 , S_2 " or " S_1 , so S_2 ", where S_1 and S_2 represent the first and the second clause respectively), or Q precedes P (e.g., S_1 , because S_2). However, in the western literature, hypothetical, conditional, and purpose relations are usually considered separate categories in existing classifications of coherence relations (Hobbs 1983, 1985; Longacre 1983; Mann & Thompson 1986, 1988; Sanders, Spooren & Noordman 1992). I am in agreement with this point of view. An important reason is that hypothetical, conditional, and purpose sentences bring more conceptual factors into the general relation of implication than the three domains of causality defined above. This is summarized in Table 1.

Table 1. A summary of the conceptual differences between the three domains of causality and hypothetical, conditional, and purpose relations

Relation	Interpretation
Content, epistemic, & speech-act	$P \rightarrow Q$ (P is presented as a fact.)
Hypothetical & conditional	Suppose $P \rightarrow Q$ (P is not realized yet.)
Purpose	P as a goal $\rightarrow Q$ (P is not realized yet.)

The antecedent Ps in all three domains of causality are generally facts. The interpretation always takes the pattern of “the fact P leads to a certain consequent Q”.⁴ Hypothetical, conditional, and purpose relations are associated with additional meanings, which initiate moderations to the basic operation, in particular to P (see Table 1). The hypothetical relation (29) and the conditional relation (30) are based primarily on suppositional or conditional antecedents rather than on established facts. The interpretation of hypothetical or conditional relations is in the form of “under the condition that a certain P (e.g., *Zhang San still loving Xiaoli*) occurs, then a certain Q (e.g., *Zhang San coming back*) should occur”. Please note that the P in these cases expresses something not yet realized.

(29) *Ruguo Zhangsan hai ai Xiaoli, ta hui huilai de.*

‘If Zhang San still loves Xiaoli, he will come back.’

(30) *Zhiyao Zhangsan hai ai Xiaoli, ta hui huilai de.*

‘Provided Zhang San still loves Xiaoli, he will come back.’

The relation of purpose, exemplified in (31), brings an additional instrumental meaning to the general implication relation, and thus should be interpreted as “a certain goal P (e.g., *to pursue his true love*) leads to a certain measure Q (e.g., *coming back to Xiaoli*)”. Clearly, the antecedent as a goal is still unrealized. In fact, it is what the speaker expects to happen at some time.

(31) *Weile zhuiqiu zhen ai, Zhang San jueding huidao Xiaoli shenbian.*

‘In order to pursue his true love, Zhang San decided to come back to Xiaoli.’

The second reason to exclude the hypothetical, conditional, and purpose relations concerns the organization of the lexicon of connectives, which suggests that the three relations in question should be treated as separate

⁴ Even if the antecedent P is sometimes in a future tense, the speaker still takes it to be the case. For example, “Because it is going to rain tomorrow, we shall put off the picnic”. The P “it is going to rain tomorrow” does not state something already realized. Yet, in the mind of the speaker, it is the case that it is going to rain the next day.

categories. In many languages, the conceptual reason can be expressed by a set of connectives such as *because* across content, epistemic, and speech-act domains (see Sweetser 1990), even though other connectives show preference for or are restricted to one specific domain (Dancygier 2009; Degand & Pander Maat 2003; Evers-Vermeul et al. 2011; Keller 1995; Pit 2003; Spooren et al. 2010; Stukker & Sanders 2012; Zufferey 2012). However, the same set of causal connectives can hardly serve to express hypothetical, conditional, or purpose relations.⁵ It follows that a line can be drawn between the two groups. In Chinese, a special characteristic associated with the use of causal connectives reinforces the idea that hypothetical, conditional, and purpose relations should be separated from causal relations. Chinese causal connectives can be used in a correlative way in the form of “causal connective₁ S₁, causal connective₂ S₂”.⁶ It is commonly acknowledged that there are various ways to combine two causal connectives in the Chinese discourse (Biq 1995; Wang 1982; Xing 2001; Zhang & Zhang 2011; Zhu 2009). Wang (1982) and Xing (2001), for example, have summarized several possible combinations, as listed in Table 2.

Table 2. Instances of the correlative use of causal connectives ⁷

Possible combinations (all translated as ‘because S ₁ , as a result/so S ₂ ’)	
<i>yinwei</i> S ₁ , <i>suoyi</i> / <i>yinci</i> / <i>yushi</i> S ₂ ;	<i>jiran</i> S ₁ , <i>suoyi</i> / <i>yinci</i> / <i>yushi</i> S ₂ ;
<i>youyu</i> S ₁ , <i>suoyi</i> / <i>yinci</i> / <i>yin'er</i> S ₂ ;	

Sentences (32) – (34), taken from the corpus created by the Center for Chinese Linguistics of Peking University (henceforth the CCL corpus), exemplify some of the above correlative uses.

⁵ There might be individual exceptions. For example, Quirk et al. (1985: 1108) suggest that the English *so that* (formal) and *so* are used to introduce clauses of result as well as clauses of purpose. In the former case, *we paid him immediately, so (that) he left contented*. In the latter case, *we paid him immediately, so (that) he would leave contented*. The counterpart of *so/so that* in Chinese, *cong'er*, can also be used to express purpose as well as causal relations. However, the overlap between the set of causal connectives and the set of purpose connectives is extremely small, which is why I regard them as separate sets.

⁶ When Chinese causal connectives are used correlatively, the antecedent P usually precedes the consequent Q. That is, in most cases of correlative uses in the form of “causal connective₁ S₁, causal connective₂ S₂”, the S₁ states the P and the S₂ states the Q.

⁷ The list of correlative uses given in Table 2 is not intended to be exhaustive. There are other possible combinations. For example, in corpora of Mandarin Chinese such as the CCL corpus, we can often see other combinations.

- (32) *Yinwei renmen gang deng shang zheli shi, kandao yigugu “bai yan” cong di xia mao chu, yushi jiu gei ta qu le zhe ge mingzi.*
 ‘*Yinwei* people saw clouds of white smoke from under the ground when they arrived here (the city), *yushi* they gave it a name like this (smoking city).’
- (33) *Youyu ta juli women jiao jin (jin 20 yu guangnian), yin’er kan shangqu hen liang.*
 ‘*Youyu* it (the star) is fairly close to us (only over 20 light years away), *yin’er* it looks very bright.’
- (34) *Tonghuo pengzhang jiran shi you huobi de guodu gongji yinqi de, yinci duifu tonghuo pengzhang de zuihao fangfa jiu shi kongzhi huobi gongyingliang.*
 ‘*Jiran* currency inflation is caused by oversupply of currency, *yinci* the best way to deal with inflation is to control the supply of currency.’

However, it is not possible to replace any of the causal connectives with a prototypical hypothetical, conditional, or purpose connective (e.g., *ruguo* ‘if’, *zhiyao* ‘provided that’, *yibian* ‘so as to’, *weile* ‘in order to’) in the above sentences. In fact, no correlative uses of causal connectives with connectives prototypically marking hypothetical relations, conditional relations, or purpose relations have been suggested in the literature, in the form of “hypothetical/conditional/purpose connective S₁, causal connective S₂” or in the form of “causal connective S₁, hypothetical/conditional/purpose connective S₂”. Again, this seems to indicate that there is a demarcation between hypothetical, conditional, and purpose relations on the one hand and causal relations within the three domains on the other hand.

In sum, this dissertation will put aside the hypothetical, conditional, and purpose relations, in view of the shared insights provided by classifications of coherence relations in the western and the Chinese literature, in view of the conceptual differences between these three relations and the causal relation expressed in the three domains, and in view of the organization of the lexicon of connectives in various languages. By narrowing down the scope, the issue of causality can be approached in terms of subjectivity in a restricted, but in-depth way.

1.4 The range of Chinese causal connectives to be studied

The Chinese language possesses a wide range of connectives expressing the causal relation that we have just defined. Though not meant to be exhaustive, Table 3 covers all the causal connectives that have been discussed in the most comprehensive study on Chinese complex sentences and connectives by Xing (2001).

Table 3. The list of Chinese causal connectives mentioned in Xing (2001)

Reason connectives (all translated as 'because')	Result connectives (all translated as 'so/therefore', except <i>yizhi</i> and <i>zhisuoyi</i>)	
<i>jiran</i>	<i>cong'er</i>	<i>yin'er</i>
<i>yinwei</i>	<i>kejian</i>	<i>yizhi</i> , 'with the result that'
<i>youyu</i>	<i>suoyi</i>	<i>yushi</i>
	<i>yinci</i>	<i>zhisuoyi</i> 'why there is a consequence of'

In this dissertation, a distinction is made between reason connectives and result connectives. Reason connectives, such as *because*, are used to mark the antecedent of the causal relation. Result connectives, such as *so*, are used to mark the consequent of the causal relation.

Among these causal connectives, *yushi* and *cong'er* are not restricted to causal relations (X. Li 2009; Xing 2001). In some cases, *yushi* is strictly temporal. For example, in (35) *yushi* is used merely to express the successive order of two connected events.

- (35) *Guo le na lin, chuan bian wanjin le Linggang, yushi Zhaozhuang bian zhen zai yanqian le.* (Xing 2001: 528).
 'The boat passed the woods, sailed into the Ling Harbor, **and then** Zhao Village was indeed in front of us.'

Opinions diverge regarding which type of relation *cong'er* often expresses, apart from the causal relation. According to Xing (2001), *cong'er* is often used in sentences such as (36) to express a subtype of additive relation, termed *lianguan guanxi* 'sequence relation'. In the view of X. Li (2009), however, *cong'er* is mainly used as a marker of a purpose relation, as in (37), when it is not used to express a causal relation.

- (36) *Cong tuanjie de yuanwang chufa, jingguo piping huozhe douzheng shi maodun dedao jie jue, cong'er zai xin de jichu shang dadao xin de tuanjie.* (Xing 2001: 530)
 'We should start from the desire to hold the nation together, solve problems by means of criticism and struggles, **cong'er** attain a new unity on the basis of a new (better) situation.'
- (37) *Ruguo nenggou zhaochu na yi ge shi yi na yi ge shi er, jiu bu nan zhaochu qi zuhe guilü, cong'er dadao zhu ge poyi de mudi.* (X. Li 2009: 51)
 'If we can find out which is one and which is two, then it is not difficult to figure out the regularity in these permutations, **cong'er** achieve the goal of deciphering the code.'

I am more in support of X. Li's point of view. I think *cong'er* would be better understood as a purpose connective in the above examples. This is because in both (36) and (37) *cong'er* can be perfectly replaced with the prototypical purpose connective *yibian* 'so as to'. This point of view does not necessarily reject the insights about *cong'er* that were put forward in Xing (2001). *Cong'er* is certainly associated with a sense of sequence. Meanwhile, we should keep in mind that the causal relation and the temporal relation also convey the sense of sequence. For this reason, I think the term "sequence relation" is too general, and hence less accurate in describing the usage of *cong'er*.

As far as the structural characteristics are concerned, *cong'er* 'so' seems to be distinct from the other causal connectives. It is typically used in multilayered complex sentences, in the form of "(*yinwei*) A, (*suoyi*) B, *cong'er* C" (see Xing 2001). In addition, another causal connective *yizhi* 'with the result that' is also often used in multilayered complex sentences like this, in which the result it introduces is given rise to by a previous result of some event (Xing 2001: 68). In multilayered structures in general, the first layer is the relation between A and BC, and the second layer is the relation between B and C. But sometimes there are exceptions. To avoid the complexity involved in considering and determining the relation between different layers of relations, this dissertation is limited to the complex sentences containing only one layer of coherence relations. Therefore, the usage of *cong'er* and *yizhi* is beyond the scope of the present study.

As is shown in Table 3, *zhisuoyi* 'why there is a consequence of' is also referred to as a causal connective in the literature (see also Wong 2005; Zhang 1996). However, it is always used together with *shiyinwei* or *shiyouyu* (both translated as 'be because') in a correlative way (Xing 2001). It is excluded from the present study because I would like to focus on causal connectives that can be used independently. Nevertheless, the use of *zhisuoyi* should be an interesting topic for future studies. Unlike other Chinese causal connectives, *zhisuoyi* is restricted to a causal order in which the consequent Q precedes the antecedent P.

In conclusion, the current study focuses on the reason connectives *jiran*, *yinwei*, and *youyu*, and the result connectives *kejian*, *suoyi*, *yinci*, *yin'er*, and *yushi*.

1.5 The structure of the dissertation

The main chapters (2-5) are written as independent papers. Chapter 3 has already been published in the journal of Chinese Language and Discourse (Li, Evers-Vermeul & Sanders 2013). Readers may find some overlap in the method sections and partial repetition of the literature in these chapters.

The advantage of organizing the dissertation in this way, however, is that readers can read each chapter on its own.

Chapter 2 presents a literature review, aimed at addressing research questions 1 through 3. This chapter demonstrates that many of the analytical categories used in previous studies on Chinese causal connectives can be related to subjectivity, although only two of them (Li & Liu 2004; Li 2011) explicitly refer to subjectivity as an explanation for the observed differences. This chapter also reviews the methods with which causal connectives and other linguistic expressions are dealt with in terms of subjectivity, and anticipates prospects of future development.

Chapter 3 and 4 each present a corpus-based quantitative study, for the purpose of addressing the fourth and fifth research question. Chapter 3 focuses on the result connectives *kejian*, *suoyi*, *yinci*, *yin'er*, and *yushi* (all translated as ‘so/therefore’). Chapter 4 deals with reason connectives *jiran*, *yinwei*, and *youyu* (all translated as ‘because’). In both studies, the connectives’ subjectivity profiles are analyzed with a set of variables such as domain (following Sweetser 1990), and the presence and identity of a Subject of Consciousness (Sanders & Spooren 2009). The linguistic data are randomly taken from written corpora of Mandarin Chinese, in three genres: news reports, novels, and opinion pieces. Taken together, the results of these studies show that six Mandarin causal connectives (i.e., *jiran*, *kejian*, *yin'er*, *yinwei*, *youyu*, *yushi*) display subjectivity profiles that differ from each other, but that are robust across genres. We assume that the degree of subjectivity of these ‘specified causal connectives’ is stored in the mental lexicon. The subjectivity profiles of two other, relatively frequent connectives, *suoyi* and *yinci* are genre-sensitive. We assume that the degree of subjectivity of these connectives is not determined in the lexicon. Hence, these connectives can be labeled ‘underspecified causal connectives’.

Chapter 5 presents an eye-tracking experimental study which is intended to address the sixth research question. In particular, we probe into what specified causal connectives such as *kejian* and *youyu*, add to the online interpretation of causal relations compared to the ones marked with ‘underspecified causal connectives’. It is found that the high degree of subjectivity encoded in *kejian* facilitates the reading of subjective epistemic causal relations, whereas the high degree of objectivity encoded in *yin'er* does not facilitate the processing of objective content causal relations more than the underspecified connective *suoyi* does.

Chapter 6 summarizes the results of the literature review, the corpus-based and experimental studies presented in Chapters 2 – 5, and reflects on the six research questions raised in Chapter 1. Following the summaries, topics and directions for future research in the study of Chinese causal connectives are suggested. .

Throughout the dissertation, Chinese examples, taken from the corpora under investigation or from the literature, are provided to illustrate the main points. *Pinyin* instead of characters is used to present the Chinese examples, with free English translations. Word-by-word translation is only provided when the ordering of the constituents is relevant.

CHAPTER 2

The relevance of subjectivity for the characterization of Mandarin causal connectives: A literature review

2.1 Introduction

As discussed in Chapter 1, in many languages of the world there is a division of labor between different causal connectives. Attempting to account for this division of labor in a systematic way, quite a few corpus-based analyses have been done on causal connectives, within the framework of ‘subjectivity’, in languages such as French, German, and Dutch (Degand & Pander Maat 2003; Evers-Vermeul, Degand, Fagard & Mortier 2011; Pit 2003; Spooren, Sanders, Huiskes & Degand 2010; Zufferey 2012). Findings from these studies have consistently shown that subjectivity – the degree to which the speaker is involved in the construction of the causal relation – plays an important role in explaining the speaker’s preference for one causal connective over another. Does subjectivity systematically shape the distribution of Chinese causal connectives as well as it shapes the distribution of European causal connectives? Or is the lexicon of Chinese causal connectives organized according to a principle other than subjectivity, given that Mandarin Chinese (abbr. Chinese) is typologically different from European languages?

In this chapter, I review previous research to explore the potential for subjectivity to explain the distributional patterns of causal connectives in Chinese discourse. Section 2.2 provides a review on the ways in which Chinese linguists conceptualize and deal with subjectivity as they analyze discourse elements in general. Section 2.3 zooms in on the relevant literature on Chinese causal connectives; in particular, previously-used analytical categorizations are assessed in relation to the notion of subjectivity. Section 2.4 focuses on evaluating the objects and methods of previous studies on Chinese causal connectives, while probing into the relation between genre and connective use. Section 2.5 summarizes the major findings that have so far resulted from these research questions, and makes suggestions for refining the existing approaches, especially the subjectivity approach, to studying the full range of Chinese causal connectives, based on systematic analyses into genuine corpus data, supplemented with results from processing studies.

2.2 Review of the literature on subjectivity in Mandarin Chinese

Several attempts have been made to study Chinese causal connectives in terms of subjectivity (Deng 2007; Gao 2013; Guo 2008; Li 2011; Li & Liu 2004; Xing 2001). However, as will be shown in section 2.3, only a few of

them have provided subjectivity analyses in a systematic way (Li 2011; Li & Liu 2004). One might think that the subjectivity approach has not yet received much attention from Chinese linguists, but this is not true. Studies on subjectivity and subjectification in Chinese discourse have been prolific since Shen (2001) systematically introduced to the linguistic scholarship in Mainland China the western theories of subjectivity and subjectification of Lyons (1977, 1982b), Langacker (1987, 1990), and Traugott (1995). These Chinese studies cover a wide scope of topics on linguistic phenomena at the discourse level, ranging from the use and function of single words, such as the adverb *dou* ‘all’ (Xu & Yang 2005) and discourse markers *qishi* ‘actually’ (Wang, Tsai & Yang 2010) and *hao* ‘so as to’ (Li 2005), to the discourse-pragmatics of linguistic constructions composed of multiple words or constituents such as the so-called disposal construction (i.e., *ba* construction) and the gain-and-loss construction (Shen 2002, 2009).

Although most of these studies are not directly related to the usage of causal connectives, I believe it is necessary to incorporate them into this discussion because they are relevant to the development of the subjectivity approach with which we are going to analyze the usage of causal connectives in Chinese discourse. Moreover, these studies have collaboratively demonstrated how the western-origin subjectivity theory contributes to new insights into the understanding of discourse phenomena in Mandarin Chinese.

2.2.1 Method and research questions

Due to space limitations, I will not review each individual subjectivity study on Mandarin Chinese; rather, a number of representative studies, published in major journals inside and outside China, will be presented here.

The relevant articles were collected via the in-site search engines provided by the websites CNKI (China Knowledge Resource Integrated Database) and Google Scholar, using the key words (both in Chinese and in English) *hanyu zhuguanxing* ‘Chinese linguistics, subjectivity’. The selection is restricted to the SSCI (Social Sciences Citation Index) journal papers. As a result, sixteen articles were collected. There are six review articles, three articles on grammaticalization and subjectification, and seven articles analyzing subjectivity expressed by specific Chinese words, phrases or constructions (including causal connectives). The subjectivity studies on causal connectives will be discussed separately in section 2.3.¹

¹ Also discovered were a few studies on intersubjectivity that focus on the aspects of meaning when the speaker expresses her concern about the addressee’s attitudes towards the issue at hand (see Lim 2011; Zhang & Li 2009). Intersubjectivity is salient mainly in discourse situated within natural interactive environments,

By reviewing these articles, I aim to address Research question 1 raised in Chapter 1, which can be broken down into the following questions. How do Chinese linguists define subjectivity? Which linguistic features are used in these studies to determine the degree of subjectivity?

2.2.2 Chinese linguists' views of subjectivity

Under a broad conceptualization, a subjective utterance is assumed to be one involving the speaker's expression of self (Feng 2006; Gao & Wen 2012; Lim 2011; Liu 2009; Shen 2001, 2002, 2009; Wang 2009; Xu & Yang 2005; Zeng 2005; Zhao 2010). This perspective is adopted from Lyons (1977), and is in line with Traugott (1995, 2010), who defines subjectivity in terms of speaker-relatedness. The sense of 'self-expression' is analyzed mainly in terms of stance-taking, the speaker's affect, and/or epistemic modality, which are elaborated on in the next section.

In general, subjectivity is assumed to be a matter of degree (Gao & Wen 2012; Liu 2009; Shen 2001; Wen & Huang 2009; Zhao 2010). This scalar view conforms to the views held by Lyons (1982b), Traugott (1995) and Langacker (1990). For example, Lyons (1982b: 105) has pointed out that the distinction between subjectivity and objectivity is considered to be gradual rather than absolute. Traugott and Langacker have also observed that subjectivity is not an all-or-nothing concept, and that there could be varying degrees of subjectivity.

A more refined distinction between the degrees of subjectivity is drawn by Hsieh (2008) in her treatment of the class of linguistic elements referred to as 'evidentials'. Hsieh assumes that an evaluation drawn from shared belief is experienced as less subjective than one built upon what is accessible to the speaker alone. For example, the reportative verb *tingshuo* 'be told' in (1) draws upon evidence accessible to a larger group of people than what the sensory verb *xiande* 'appear' or 'look' in (2) is based upon. For that reason, Hsieh suggests that (1) should receive an interpretation with a lower degree of subjectivity than (2).

- (1) *Tingshuo chi zhong you bei ren fangsheng de, er san bai zhi wugui.* (Hsieh 2008: 221)

It is said that there were two or three hundred turtles set free in the pond.'

therefore it is beyond the scope of the present dissertation, which is restricted to written discourse.

- (2) *Daitu zai yanyu yingdui shang, sihu xiande xiangdangde zhuanye.* (ibid)
 ‘It seemed that the gangsters **looked** quite skillful in oral response.’

This sense of subjectivity follows from Lyons’ (1977) sense of ‘speaker commitment’. As Lyons contends, the more unreliable the ground of an evaluation is, the more likely it is that the evaluation will be perceived as subjective. Within the same line of reasoning, Nuyts (2001) construes subjectivity as ‘evidential qualification’. The critical difference concerns whether the evidence is in the knowledge of the speaker alone or is shared between the speaker and the hearer.

Several Chinese review articles have highlighted Langacker’s (1990) definition of subjectivity in terms of on/off-stage conceptualization (Feng 2006; Gao & Wen 2012; Liu 2009; Shen 2001; Zhao 2010). It seems indisputable to Chinese linguists that Langacker’s view of subjectivity covers a narrower scope than the definitions by Lyons (1997) and Traugott (1995) (see also Gao & Wen 2012; Liu 2009; Zhao 2010). As introduced in Chapter 1, Langacker focuses on the specific means of ‘construal’: Does the speaker conceive of events or situations from an on-stage perspective or from an off-stage perspective? In fact, in both cases the speaker’s perspective is certainly involved. Hence, what Langacker considers to be objective construal (i.e., on-stage conceptualization) is actually not objective. Langacker does not pay attention to the truly objective cases as defined by Lyons (1977) and Traugott (1995), i.e., those cases that do not involve the speaker’s perspective at all. Perhaps it is due to this limitation that Langacker’s view of subjectivity (with regard to on/off-stage status of the speaker’s perspective) has not yet been used by Chinese linguists in any concrete analysis.

2.2.3 Linguistic features used to determine the degree of subjectivity

Studies on subjectivity and its manifestations in Chinese discourse in general have focused on the three subjectivity indicators described by Finegan (1995): speaker’s perspective, affect, and epistemic modality (for an overview, see Liu 2009, 2011; Shen 2001; Zhao 2010).

The speaker’s perspective generally refers to the aspect from which the speaker views the circumstances described in the utterance (Shen 2001). According to Shen (2001), the speaker’s perspective is often implicitly indicated. It is not linguistically realized in the utterance. In this sense, the implicit speaker perspective approximates the off-stage conceptualization in the terms of Langacker.

It is generally assumed that the speaker’s perspective often concerns the speaker’s judgment of quantity or property (Li 2005; Shen 2002, 2009; Xu &

Yang 2005). For example, Shen (2009) has observed that Chinese speakers can express a big quantity of loss in (3) by selecting the adverb *jiu*, and a small quantity of loss in (4) by using another adverb *cai*.

- (3) *Wang Mian ershi sui jiu si le fuqin.* (Shen 2009:8)
 ‘Wang Mian **jiu** lost his father when he was twenty.’
- (4) *Wang Mian ershi sui cai si le fuqin.* (ibid)
 ‘Wang Mian **cai** lost his father when he was twenty.’

The Chinese adverbs *jiu* and *cai* don’t have counterparts in English. Therefore, from the above English translations it is difficult to tell the difference between the meaning expressed by (3) and (4). Actually, (3) can be paraphrased as “Wang Mian lost his father when he was only twenty”, and (4) can be taken to mean “when his father was dead, Wang Mian was already twenty”. Sentence (3) implies that the father’s death is a big loss to Wang Mian, whereas (4) implies that the father’s death is a small loss to Wang Mian. In both cases, the speaker’s perspective (or judgment over the quantity) is involved, and the speaker’s perspective is implicit/off-stage.

The speaker’s affect is dealt with mainly in terms of empathy (Shen 2002, 2009). The definition for empathy is taken from Kuno (1987: 206): empathy refers to “the speaker’s identification, which may vary in degree, with a person/thing that participates in the event or state that he describes in a sentence”. In line with this idea, Shen discusses empathy in terms of the speaker’s sympathy with, fondness for, or dislike of a participant in the utterance. For example, Shen (2002, 2009) has observed that the *ba* construction often expresses ‘subjective disposal’ on the basis of examples such as (5).

- (5) *Yidali dui ba deguo dui ying le.* (Shen 2002:390)
 ‘The Italian team beat the German team (in the football game).’

Because there is no counterpart of *ba* in English, the translation given above cannot represent the structure accurately. If we leave *ba* in the sentence and gloss the other words following the original order, (5) would be “the Italian team *ba* the German team beat”. Syntactically speaking, the use of *ba* converts the normal SVO order into the SOV order. In terms of semantics and pragmatics, Shen (2002) suggests that this SOV structure gets encoded with a higher degree of subjectivity than the original SVO structure. As Shen argues, the speaker expresses her empathy with the loser “the German team” by means of the *ba* construction. The normal SVO structure, however, does not have this interpretation.

Another example (6) is given in Shen (2002: 391) to show how the speaker voices her dislike of the old clothes through the use of *ba*. (6) would be literally transcribed as “*ba* these old clothes quickly sell”.

- (6) *Ba zhexie jiu yifu gankuai mai le ba!*
 ‘Be quick to sell these old clothes!’

Epistemic modality also involves the speaker’s judgment/evaluation. Unlike a speaker’s perspective concerning a judgment of quantity or property, epistemic modality concerns the speaker’s judgment on the likelihood that the propositional content of the utterance is true (Shen 2001). Modal adverbs are typically used to express epistemic modality. For example, in the sentence *he must be married*, the speaker judges that it is likely that the propositional content *he is married* is true.

Wang, Tsai & Yang (2010) implicitly referred to epistemic modality as a subjectivity indicator in their study of the usage of Chinese adverbs *qishi* ‘actually’ and *shishishang* ‘in fact’ in spoken discourse. They found that both adverbs express a high degree of subjectivity. According to the authors, the two adverbs express the speaker’s judgments towards the truth of the propositions, to contrast with prior assumptions.

- (7) *Wo xiang, qishi, ta ziji bushi hen xiang dao bianyibu qu. Ruguo, yi wo dui ta de liaojie* (taken from a conversation in which two persons talk about their mutual friend who is a newspaper reporter). (Wang, Tsai & Yang 2010: 715)
 ‘I don’t think that she **actually** wants to work in the Editing and Translation department, according to my understanding of her.’
- (8) *Shishishang, wo xiang ne, zongtong ta ye shi hen jiesheng de ren oh* (taken from an interview in which the interviewee talks about an anecdote of Lee Denghui, the former President of Taiwan). (Wang, Tsai & Yang 2010: 712)
 ‘**In fact**, I think the President is also a very frugal person.’

In (7), although *she* is a newspaper reporter, the speaker expresses her judgment that actually *she* does not like the job. Contrary to one’s intuitional assumption, it is not true that *she* does such a job because *she* likes it. In (8), with the adverb *shishishang*, the speaker expresses her certainty that *the President* is a very frugal person, to contrast with the conventional assumption that *the President* or anyone in a supreme social rank is not frugal.

In addition, a diachronic study (Zeng 2005) takes the fact that a word has derived an epistemic meaning as a sign of the word being subjectified.² Zeng shows how the Chinese expressions *wo kan* ‘I see’ and *ni kan* ‘you see’ have been subjectified to be epistemic-stance markers “I/you think” (10) on the basis of their original verbal meaning ‘I/you see’ (9). In Langacker’s terms, though, (10) would be a case of objectification, i.e., on-stage conceptualization.

- (9) *Gangcai wo kan le na hu, kanchu zhe ge ren qushi shi you dian genji de, suoyi wo cai duo zhe fen zui.* (Zeng 2005: 17)
 ‘I saw that teapot, I could tell that this person was thoroughbred, so I talked so much (about him).’
- (10) *Wo kan, mei fa ban.* (ibid)
 ‘I think, there is no way out.’

In a similar way, Wang (2009) examines the subjectification of the Chinese expression *xiangshi* ‘(I) imagine’. Originally, the word *xiang* ‘imagine’ merely expressed one’s imagination. Later, the word *xiang* became an epistemic-stance marker ‘think’, and was subsequently combined with the copular verb *shi* ‘be’ to express the speaker’s prediction/inference about whether a certain event has occurred or will occur under the given circumstances. In the end, the word *xiangshi* lost the competition with its close synonyms such as *xiangbi* ‘presumably’ and *xianglai* ‘I assume’, and disappeared.

2.3 Review of the literature on Chinese causal connectives

The previous section provides a general overview of the studies on subjectivity and its manifestations in Chinese discourse. This section aims to zoom in on the studies of Chinese causal connectives, including those studies that have explicitly referred to the notion of subjectivity as an explanation for the discovered differences in use, and those that have not referred to the notion of subjectivity. Despite the fact that the latter type of study does not appear to be relevant to subjectivity at first glance, the analytical categories used in these studies and the conclusions they draw may possibly shed some light on the way in which the use of Chinese causal connectives can be approached in terms of subjectivity.

² A word is considered to have been subjectified (to some extent) when its meaning becomes more and more dependent on the speaker’s attitude towards the proposition; this process is referred to as ‘subjectification’ (see Traugott 1995, 2010).

2.3.1 Method and research questions

The relevant articles were collected via the in-site search engines provided by the websites CNKI (China Knowledge Resource Integrated Database) and Google Scholar, using the key words (both in Chinese and in English) *yinguo guanxi lianci* ‘causal connectives, Chinese’. As a result, I collected 20 articles on causal connectives: 12 synchronic studies, and 8 diachronic studies.

By reviewing these articles, I aim to address Research question 2 raised in Chapter 1: Which analytical categories have been used in these studies? Can these categories be related to analyses in terms of subjectivity, and if so, how? With these questions in mind, I intend to explore the feasibility of studying Chinese causal connectives with a subjectivity approach.

2.3.2 The analytical categories and their relevance to subjectivity

Descriptive and inferential causality and domains

Causal connectives can be analyzed on the basis of the characteristics of the causal relation they are frequently used to express. In accordance with the Chinese traditional views of causality introduced in Chapter 1, some studies have distinguished between causal connectives that are often used to mark descriptive causal relations and those that are often used to mark inferential causal relations (Deng 2007; Guo 2008; Xing 2001). On the basis of a substitution test, Xing (2001) has observed that *yinwei* ‘because’ focuses on describing or reporting causal relations whereas *jiran* ‘because, since’ carries a strong inferential flavor. Moreover, he relates this distinction to the notion of subjectivity. Examples (11) and (12), taken from Xing (2001: 73), serve to illustrate his point.

- (11) *Yinwei ta fandui, wo zhihao liu xialai.*
‘**Yinwei/Because** he objects, I have no choice but to stay.’
- (12) *Jiran ta fandui, wo zhihao liu xialai.*
‘**Jiran/Since** he objects, I have no choice but to stay.’

The propositional content of (11) and (12) are identical except for the choice of connective. According to Xing (2001), the substitution of *jiran* for *yinwei* (or vice versa) changes the interpretation of the causal relation. The speaker can use *yinwei* in (11) to inform the reader that two events have occurred, with one “I am obliged to stay” resulting from the other “he objects”. The speaker can use *jiran* in (12) to express her inference “there should be no other choice but to stay” on the basis of the situation under consideration “he objects”. Xing explains that *jiran* is encoded with a high degree of subjectivity, which is why it is taken as arguing towards causal connections

on the basis of subjective inferences; by contrast, *yinwei* carries a high degree of objectivity, which is why it is understood as reporting objective causal connections out there in the observable world.

Causal connectives in Chinese discourse can also be characterized with approaches developed from western linguistic data. In light of the ‘domain theory’ (Sweetser 1990) of western-origin, Li and Liu (2004), Li (2011) and Gao (2013) have investigated the differences between *jiran* and *youyu* ‘because’, between *yinwei* and *youyu*, and between sentence-initial *yinwei* and intersentential *yinwei* in terms of their distribution in the content, the epistemic, and the speech-act domains. Moreover, a link was established in Li (2011) and Li and Liu (2004) between the three domains and subjectivity. The content domain reflects objective reality, so it is an objective domain. The other two domains concern the speaker’s attitude towards the described events, so they are more subjective. Domain, then, is used as a variable to indicate the degree of subjectivity encoded in the connectives: causal connectives that are distributed more frequently to the content domain than to the epistemic and speech-act domains should be encoded with a lower degree of subjectivity, and vice versa.

Li (2011) provides a corpus-based subjectivity analysis of the differences between *youyu* and *yinwei*. It was observed that 63 out of 100 cases of randomly selected fragments with *yinwei* (from *Guojia yuwei pingheng yuliaoku* ‘Balanced corpus constructed by the national language committee’) are in the content domain. Fragment (13) exemplifies such a usage.

- (13) *Woguo zhuming shuxuejia Hua Luogeng, chuzhong biye hou zhi shang le yi nian zhigong xuexiao, houlai yinwei jiating shenghuo kunnan, dao yi jia xiao zahuodian dang le xuetu.* (Li 2011: 492)

‘The famous mathematician of our country Hua Luogeng, went to an apprentice school only for one year after he finished middle school. **Yinwei** his family was poor, he then went to a small grocery store as an apprentice.’

This result is then compared to the finding about *youyu* in an earlier study: 88 out of 100 cases of *youyu* (randomly taken from *1996 nian renmin ribao* ‘People’s daily newspaper in the year 1996’) are used to express the content causal relation (Li & Liu 2004). A case in point is given in (14).

- (14) **Youyu** *cuoshi de li, daotian yang yu bi shang nian kuoda 300 duo wan mu.*

‘**Youyu** the measures were effective, the rice fields used for pisciculture increased by 3 million mu (i.e., about 2000 square kilometers).’

On the basis of this disparity between *yinwei* and *youyu* with respect to their distributional proportions in the content domain, Li (2011) concludes that *yinwei* tends to be more subjectively-oriented than *youyu*.

These two corpus-based studies are particularly valuable, because they represent the very limited number of studies targeted at providing systematic analyses of synchronic differences between Chinese causal connectives in terms of subjectivity. As pioneering studies, there are inevitably certain limitations, which are discussed in section 2.4.

The defeasibility of the consequent Q

Apart from domains, Li and Liu (2004) draw our attention to the defeasibility of the consequent segment. The authors have observed that in certain contexts such as “*jiran* P, Q” where P stands for the antecedent and Q stands for the consequent, the content of Q is defeasible. This is evidenced by the widely accepted appropriateness of the usage pattern “*jiran* P, Q, but not Q”. That is, the validity of Q can be annulled by the propositional content of a subsequent clause, as in example (15).

- (15) ***Jiran*** *ta you nengli, name jiu yinggai zhong yong ta, keshi ni que bu zhong yong ta.* (Li & Liu 2004: 127)
 ‘***Jiran*** he is a capable man, we should put him into an important position, but you didn’t put him into any important position.’

In (15), the consequent clause “we should put him into an important position” expresses an inference, which is directly annulled by the subsequent clause “you didn’t put him into any important position”. Interestingly, the same consequent clause Q can become indefeasible when it is put into a different context, as in example (16).

- (16) # ***Youyu*** *ta you nengli, name jiu yinggai zhong yong ta, keshi ni que bu zhong yong ta.* (Li & Liu 2004: 128)
 ‘***Youyu*** he is a capable man, we should put him into an important position, but you didn’t put him into any important position.’

When the causal connective *youyu* is used in place of *jiran*, Q can no longer be annulled. It is perfectly acceptable to say “*youyu* he is a capable man, we should put him into an important position”, but it is impossible to add to it the third clause “but you didn’t put him into any important position” that invalidates Q. An intuitional explanation is given in Li and Liu (2004). The authors suggest that the defeasibility of Q is probably related to the degree of subjectivity encoded in the causal connective. According to the authors, *jiran*

often co-occurs with a Q that expresses subjective opinions, and subjective opinions are always defeasible, which is why it's perfectly acceptable to say "jiran P, Q; but not Q". From this, I think, we can generalize that it is subjective causal connectives that are likely to co-occur with defeasible Qs, given that subjective causal connectives generally go hand in hand with opinions. Which type of causal connectives should require an indefeasible Q? The authors suggest: even if *youyu*, which often conveys a high degree of objectivity, is occasionally used with an inferential consequent as in (16), the objectivity encoded in the connective may still positively affect the interpretation of the truthfulness of the causal relation. As a result, it is not likely to cancel out the consequent. From this, we can infer that objective causal connectives are likely to co-occur with indefeasible Qs, given that the objectivity of the connective may well enhance the objectivity of the causal relation. In addition, objective causal connectives often co-occur with observable reality, which is no doubt indefeasible. In sum, there seems to be a link between the defeasibility of Q and the subjectivity encoded in the causal connective, which deserves more academic attention in the future.

Some Qs, however, can be more elusive and complex than facts, opinions, or inferences discussed previously. These cases concern speech acts, which usually take the form of questions or imperatives. Like physical acts, speech acts cannot be annulled. By uttering a sentence in the form of a speech act, the speaker actually performs an action with a specific force, such as offering, promising, and commanding (see also Pit 2003; Stukker & Sanders 2012). See the examples below.

- (17) *Ziji you jian fangzi, zhe jiao wo juede ni tebie ke'ai. Suoyi nishuo wo zenme hui jijiao ni dui wo de taidu?* (Guo 2006: 31)

'You have your own room, and this already makes me feel you are very lovely. **Suoyi** why should I care about your attitude towards me?'

- (18) *Ta keneng you kai hui le, suoyi bie deng ta le, xian chi zhe ba.* (Guo 2006: 32)

'He probably is at a meeting, **suoyi** don't wait for him and let's eat first.'

In (17), the speaker asks a rhetorical question "why should I care about your attitude towards me?". This rhetorical question actually implicates a strong negative assertion "I don't care about your attitude towards me". In (18), the Q "don't wait for him and let's eat first" is obviously an imperative. To be more precise, it is a suggestion. As the utterance is made, the speaker has performed a speech act of making a suggestion.

The speech act as a realized act is indefeasible, but the content of the speech act as an assertion or a suggestion is defeasible. This is what makes

speech acts appear elusive and complex. Nevertheless, we cannot leave out the speech act as an independent analytical category for the analysis of causal connectives. In particular, the importance of the speech-act category can be seen from the observation in the literature that certain causal connectives are distinguishable from each other on the basis of whether or not they can co-occur with rhetorical questions and imperatives (Guo 2006). Guo (2006) has pointed out that *suoyi* ‘so’ co-occurs with both rhetorical questions and imperatives (see (17) and (18)), whereas *yushi* ‘so’ does not. Suppose that a causal connective frequently co-occurs with speech-act Qs. What does that imply about the connective’s subjectivity? Guo (2006) does not discuss such questions. However, I think the answer is readily available. Given that the speech-act domain (in which the speaker provides a justification for the speech act that he has performed) is considered as a subjective domain (Li 2004; Li & Liu 2011), the frequent co-occurrences of causal connectives with speech-act consequents in the speech act domain are likely to indicate that the connective expresses a high degree of subjectivity.

The modality of the consequent Q

Another factor that seems to affect the speaker’s choice of causal connectives has been shown to be the modality of Q (Guo 2006, 2008; X. Li 2009; Zhao 2003). Consequent Qs are deemed to exist in one of two modalities: the static modality versus the dynamic modality. The Q is considered static when it describes the state of something, as in (19), or judges upon the state of something, as in (20). The Q is considered dynamic when it involves action (‘process’ in Guo’s (2008) terms), such as (21), or change of state (‘achievements’ in Guo’s terms), as in (22).

- (19) *Dong kou zhang man le shumu, suoyi dong li heihuhu de, shenme ye kan bu jian.* (Guo 2008: 24)
 ‘The mouth of the cave is fully covered with woods, *suoyi* it is very dark inside, one can hardly see anything.’
- (20) *Wo dui mianqian de hutaomu canzhuo gan xingqu, dian le ta yi ba, faxian ta tai zhong, suoyi shi zhong hecheng cailiao, bushi zhen hutaomu de.* (ibid)
 ‘I was interested in the walnut table in front of me, weighed it in the hand, and found it too heavy, *suoyi* it was made of synthetic materials rather than of walnut.’
- (21) *Jiran ta shenwufenwen, jingcha jiu fang ta zou le.* (ibid)
 ‘*Jiran* he was penniless, the policemen released him.’
- (22) *Hui jia de lushang xia yu le, suoyi dao jia de shihou ta de yifu quan shi le.* (ibid)
 ‘It rained on his way back home, *suoyi* his clothes were all wet when he got home.’

Some causal connectives show a clear preference of one modality over another. For example, it is found in Guo (2006) that 70% of the *yushi* cases co-occur with Qs in the dynamic modality, describing actions, whereas 74% of the *suoyi* cases co-occur with Qs in the static modality.

Can we interpret these findings in terms of subjectivity? It has been suggested in the western literature that the observability of the resulting event often reflects the degree of the subjectivity of causality (Pit 2003). In Pit's view, objective causality always concerns resulting events or states that are physically observable, whereas subjective causality always involves a result that is internal to the speaker's reasoning process. Accordingly, the findings presented above imply at least that *yushi* is relatively objective, since the dynamic modality concerns observable reality, while the static modality seems to be a mixture. In order to analyze the degree of subjectivity associated with the static modality, we have to split it into two parts: those involving states and those involving judgments upon states. The former, involving states, should be objective, and the latter, involving judgments, should be subjective.

Linearization: forward order versus backward order

Causality can be expressed via two sequences (Su 2002; Tsai 1996; Y. Wang 2002, 2006). One is termed the *zhengxu* 'forward order/linking' (Biq 1995; Zhang & Zhang 2011; see also 'basic order' in Sanders, Spooren & Noordman 1992), in which the antecedent P precedes the consequent Q. The other is termed the *daoxu* 'backward order/linking' (Biq 1995; Zhang & Zhang 2011; see also 'non-basic order' in Sanders et al. 1992), in which the antecedent P follows the consequent Q. Causal connectives can be analyzed in terms of the order in which they are often used to express causality. For example, Li and Liu (2004) suggest that *youyu* strictly follows the forward order (see (23)), whereas *jiran* shows flexibility in sequencing causal events (see (24)).

- (23) a. **Youyu** *mai huo mei zhuantou, shangdian zhihao guanmen.*
 'Youyu it made no profits, the shop had to be closed.'
 b. **Shangdian zhihao guanmen, youyu mai huo mei zhuantou.*
 'The shop had to be closed, youyu it made no profits.'
 (Li & Liu 2004: 125)
- (24) a. **Jiran** *mai huo mei zhuantou, shangdian zhihao guanmen.*
 'Jiran it made no profits, the shop had to be closed.'
 b. *Shangdian zhihao guanmen, jiran mai huo mei zhuantou.*
 'The shop had to be closed, jiran it made no profits.'
 (Li & Liu 2004: 126)

In view of another finding in Li and Liu (2004), that *youyu* is used more often in the objective content domain (88%) whereas *jiran* is frequently used in the

subjective epistemic and speech-act domains (98%), the above observations seem to imply that degrees of subjectivity encoded in connectives impose constraints on the linearization of causality.

Two other important findings are reported in Gao (2013) and Zhang and Zhang (2011), respectively. According to the former study, sentence-initial *yinwei* (used in a forward order) mainly expresses content relations whereas a considerable number of *yinwei* are used in intersentential positions (in backward order) to express epistemic relations. The latter study reports a similar result: when *yinwei* is used in backward order, the consequent is always a conclusion. That is to say, intersentential-*yinwei* is always used in causal relations holding between premises and conclusions, i.e., the epistemic domain. The findings from these two studies are introspective, without quantitative analyses. Still, they can be taken to presumptively suggest that intersentential *yinwei* expresses a higher degree of subjectivity than sentence-initial *yinwei*, and hence that the notion of subjectivity is relevant for classifying Chinese causal connectives.

The variation in linearization can also be dealt with in terms of communicative functions. This is in line with the idea in the field of functional linguistics that structural variants are often motivated and constrained by functional factors (Dik 1989; Halliday 1994). Studies (Biq 1995; Song & Tao 2009) have shown that the two causal sequences “*yinwei* P, Q” and “Q, *yinwei* P” perform different functions: ‘ideational function’ versus ‘interpersonal function’, – two categorizations created by Halliday in his ‘systematic functional grammar’. Halliday and Hasan (1976: 27) distinguish the two functional components in terms of the speaker’s role in the utterance: the ideational component represents the speaker in her role as observer, while the interpersonal component represents the speaker in her role as intruder. In their view, the interpersonal component concerns the speaker’s angle: her attitudes and judgments, and her motive in saying anything at all (ibid). From their interpretations of these two components, we can tell that the distinction between the ideational function and the interpersonal function actually mirrors the one between the content domain and the epistemic/speech-act domain in Sweetser’s terms, and hence can be related to the degree of subjectivity of the utterance. That is, expressions serving the interpersonal function are associated with a higher degree of subjectivity than expressions that serve the ideational function.

Biq (1995) has found that the written corpus (of domestic news articles) shows a balanced distribution between the two sequences (45%: 55%). In the spoken corpus (of conversations), however, only 12% of the *yinwei* tokens are used in the “*yinwei* P, Q” order, and all of them express the ideationally determined causal relation. Two implications might be derived from these findings. First, interpersonal factors inherent in conversations seem to constrain the occurrence of the “*yinwei* P, Q” order. Second, there seems to be

a tendency that the “*yinwei* P, Q” order is tied to the ideational function. Given that the ideational function has been linked to the objective content domain, we can presume that the “*yinwei* P, Q” order tends to be objectively-oriented.

Similar and more straightforward results were obtained in Song and Tao (2009). On the basis of a written corpus LCMC (Lancaster Corpus of Mandarin Chinese) (McEnery & Xiao 2004) and a spoken corpus CallFriend (Canavan & Zipperlen 1996), this study has also provided a functional analysis of sentence-initial *yinwei* versus intersentential *yinwei*. They conclude that, in both conversation and writing, these two structures are associated with separate discourse functions: the information-sharing and the interactional function.³ It was found that sentence-initial *yinwei* is used to provide background information for interpreting the event described in the following clause. In this sense, it is informative in nature. Or, it can be termed ideational in accordance with Biq (1995). From this we can presume further that sentence-initial *yinwei* should occur frequently in the content domain, and hence is likely to be objectively-oriented. Results also show that intersentential *yinwei* is interactional/interpersonal, as a follow-up justification to address “potential enquiries, disagreement, suspicion, or general unexpectedness”. In this sense, I think the usage of intersentential *yinwei* can be considered subjective. If the speaker wants to react to potential objections and defend her own point of view, her utterances tend to be subjective. It is likely that she needs the very subjective form to express herself in the utterance.

Variation in the type of subject

Causal connectives vary with respect to the type of subject they can co-occur with. Guo (2006) drew a distinction between animate subjects and inanimate subjects. The study reports that, out of 50 cases of *yushi*, 94% of the antecedent Ps have animate subjects and 96% of the consequent Qs have animate subjects. For the connective *suoyi*, 52% of the antecedent Ps have animate subjects and 76% of the consequent Qs have animate subjects. The author explains that *yushi* is typically used in narration, which is why it shows a high proportion of animate subjects.

³ Among the studies collected for this Chapter, this is the only one that has calculated statistical significance for some of the observed differences. However, the statistical significance test was only applied to the distributional difference between the two sequences (i.e., sentence-initial *yinwei* and intersentential *yinwei*) in the corpora. It was found that final causal clauses with *yinwei* (i.e., intersentential *yinwei*) occur significantly more frequently than initial causal clauses with *yinwei* (i.e., sentence-initial *yinwei*) in both the corpora under investigation.

Is the distinction between animate and inanimate sentence subjects related to subjectivity? I think this is not very likely. Although inanimate subjects often correlate with objectivity and animate subjects often correlate with subjectivity, the animacy of the subject is not the only relevant factor (cf. the animacy of the ‘causally primary participant’ in Pit 2003). Even if the sentential subjects are inanimate, the relation can still be subjective. For example, in (25), the subjects *it* and *the football match* are both inanimate, but the sentence is associated with a high degree of subjectivity. The animate speaker is implicitly involved in arguing towards the conclusion “the football match will probably be canceled” on the basis of the supporting evidence “it is going to rain tomorrow”.

(25) *Jiran mingtian hui xiayu, yinci zuqiusai hen keneng hui bei quxiao.*

‘*Jiran* it is going to rain tomorrow, *yinci* the football match will probably be canceled.’

(26) *Wo cong zixingche shang shuai le xia lai, suoyi wode zuo tui shou shang le.*

‘I fell off my bike, *suoyi* I was hurt in the left leg.’

On the other hand, even if the subject is animate, the relation is not necessarily subjective. This is because the animate subject is not necessarily involved in the subjective construal of the causal relation. That is, it could be the case that the animate subject is merely involved as an object of description. In (26), for example, the subject of both P and Q is the speaker *I*, but the animacy of the speaker does not increase the subjectivity of the causal relation. The sentence is still in the content domain: the fact of falling off the bike leads to the fact of hurting the leg. Hence, the animacy of the subject is not a direct indicator of the degree of subjectivity associated with the causal relation or connective.

While the animacy of the subject does not seem to be relevant to subjectivity, the distinction between *juzi zhuyu* ‘sentential subject’ and *yanzhe zhuru* ‘speaker subject’ does. These two terms for the types of subject are taken from the literature on the theory of subjectivity (Shen 2001). The sentential subject is straightforward: it is the syntactic subject of the sentence. The speaker subject (also called *yanshuo zhuyu* ‘utterance subject’) can be taken to be a conscious speaker. For example in the sentence *the highway runs from the valley floor to the mountain ridge*, there is an implicit speaker subject *I* besides the explicit sentential subject *the highway*. The highway does not run in the real world, but the speaker *I* mentally scans the highway from the valley to the ridge in his mind. An utterance is considered subjective when there is a speaker subject in it (Shen 2001).

The speaker subject can be identified with the ‘off-stage speaker’ proposed by Langacker (1990). Both have to be referred to for the interpretation of the utterance, but neither is linguistically realized as the source of the subjective construal in the utterance. Furthermore, both have been related to the notion of subjectivity. Langacker considers the off-stage conceptualization maximally subjective on the grounds that the speaker is so absorbed in the perceptual experience that she loses all awareness of self (1990: 7). This idea coincides with Shen’s (2001) view on the subjectivity of speaker subjects.

We cannot simply equate sentential subject with ‘on-stage speaker’, even if the subject happens to be in the form of *I*. Very often the on-stage speaker is the subject of one or both of the connected clauses, as in (27). In some cases, as in (26), the sentence subject is in the form of *I*, but *I* does not count as an on-stage speaker. This is because *I* does not function as a conceptualizer there.

(27) I think he still loves Mary, **because** he has come back.

Chinese causal connectives have not yet been studied systematically in terms of on-stage or off-stage conceptualization, but one study has suggested that causal connectives may differ with respect to whether they co-occur with speaker subjects. On the basis of introspection, Li and Liu (2004) propose that *jiran*, as a subjective connective, should commonly have both ‘sentential subjects’ and ‘speaker subjects’ whereas *youyu* as an objective connective should have sentential subjects but no speaker subjects. The ideas expressed in Li and Liu (2004) can be illustrated with examples (28) and (29).

- (28) *Ta jiran bu xiang qu, ni jiu yi ge ren qu ba.* (Li & Liu 2004: 127)
 ‘He **jiran** does not want to go (there), you should go there alone.’
- (29) *Youyu tianqi qi le bianhua, feiji mei anshi qifei.* (Li & Liu 2004: 126)
 ‘**Youyu** the weather conditions changed, the plane did not take off as scheduled.’

Sentence (28) contains an implicit speaker subject *I*: *I* know he does not want to go (there), *I* suggest you go there alone. According to the authors, this typically represents the context in which *jiran* is used; by contrast, *youyu* is typically used to express causal relations such as (29), which does not involve the speaker as a conscious subject, apart from the fact that the speaker/writer reports events and circumstances in the utterance.

What Li and Liu (2004) observe about the co-occurrence of *jiran* and *youyu* with sentential subjects and speaker subjects is simply a suggestion. Corpus-based quantitative analysis is needed to testify to their suggestion,

and check whether co-occurrence with a ‘speaker subject’, or an ‘off-stage speaker’ in Langacker’s terms, can be regarded as an indicator of the connective’s degree of subjectivity. In the meantime, the category of ‘on-stage subject’ also deserves more attention, given that Langacker considers it to be an indicator of a lower degree of subjectivity.

Finally, some studies attempt to distinguish between causal connectives in terms of whether or not the connected clauses share the same subject (X. Li 2009; Zhao 2003; Deng 2007). This does not seem to be relevant to subjectivity, so I do not introduce these studies here in detail.

Tense and aspect: future versus perfect

Zhao (2003) attempts to account for the difference between causal connectives in terms of the tense and aspect of the consequent clause Q. The study observes that *yushi* ‘so’ is restricted to the perfect aspect and is never used in the future tense; whereas *cong’er* ‘so’ and *yinci* ‘so’ are used in both the future tense and the perfect aspect.

Can we relate the above results to subjectivity? Zhao herself does not address this issue, but it seems that the answer should be in the positive. Western studies have already explicitly related tense and aspect to subjectivity (Fleischman 1990; Pit 2003). According to these studies, both the future tense and the perfect aspect are highly subjective, because they are ‘speaker-based’ (Pit 2003:150). Furthermore, Pit (2003) observes that the future tense carries a higher degree of subjectivity than the perfect aspect. In her view, the more remotely the event is situated (in the past) from the time of speaking, the harder it is (for the author and the reader) to adopt the speaker’s perspective. Moreover, she observes that the future tense is often pragmatically endowed with modality, which creates a subjective mode.

2.4 Objects, methods, and genre

As introduced in Chapter 1, European studies in the use of causal connectives with respect to subjectivity are developing rapidly in breadth and depth. Regarding the objects of these studies, some explore diachronically the evolution of the subjective uses of causal connectives (i.e., subjectification), including both monolingual analyses (Keller 1995) and cross-linguistic comparisons (Evers-Vermeul, Degand, Fagard & Mortier 2011); some focus on the synchronic differences between different causal connectives within a single language (Pander Maat & Sanders 2001; Pit 2003; Spooren, Sanders, Huiskes & Degand 2010; Zufferey 2012), or between causal connectives across different languages (Dancygier 2009; Degand & Pander Maat 2003; Stukker & Sanders 2012). In addition, several acquisition studies have been conducted to inquire into the development of causal relations/connectives in child language (Evers-Vermeul & Sanders 2009,

2011; Spooren & Sanders 2008; van Veen 2011). Results show that content relations are acquired before epistemic relations (Evers-Vermeul 2005; Evers-Vermeul & Sanders 2011; Spooren & Sanders 2008; van Veen 2011). Moreover, psycholinguists show their interest in the role of subjectivity in online reading. Several self-paced reading experiments and eye-tracking experiments have already shown that subjective causal relations/connectives take more time to process than objective causal relations/connectives (Canestrelli, Mak & Sanders 2013; Traxler, Bybee & Pickering 1997a; Traxler, Sanford, Aked & Moxey 1997b). The results from these processing studies have lent support to the idea that the notion of subjectivity is related to general human cognition (see Noordman & De Blijzer 2000; Sanders 2005).

In terms of method, many of these studies use written corpora (Degand & Pander Maat 2003; Evers-Vermeul et al. 2011; Pander Maat & Sanders 2001; Pit 2003; Zufferey 2012), or spoken corpora (Evers-Vermeul & Sanders 2009; Spooren, Sanders, Huiskes & Degand 2010; van Veen 2011), and employ quantitative methods and statistical techniques to work empirically with the linguistic data drawn from the corpora. A few studies are not corpus-based and do not aim for quantification (Dancygier 2009; Keller 1995; Sweetser 1990). Still, they have offered a rich set of categories for classification, and have provided detailed qualitative analyses of the use of causal connectives with the identified categories.

This section provides an evaluation of the objects and methods of previous studies on Chinese causal connectives, in comparison with those on European causal connectives (as discussed above). The aim of this section is to address Research question 3 raised in Chapter 1, which can be specified as follows: Do Chinese and European studies address the issue of subjectivity in causality with similar methods? Do they target a scope of objects in similar breadth?

Furthermore, this section is intended to inquire into the impact of genre on the use of causal connectives. It has been observed that medium type has an influence on the type of relations to be found. For example, speech-act causal relations are reported to occur more often in the spoken discourse than in the written discourse (Sanders & Spooren submitted; Spooren, Sanders, Huiskes & Degand 2010). In spontaneously produced discourse, the medium influences the use of Dutch causal connectives *omdat* and *want* (both translated as 'because'): the instances of *want* constructions that express speech-act relations are more frequently observed in chat than in conversations (Sanders & Spooren 2009).

Within the written modality, we might also expect genre or text type to affect the use of causal relations/connectives. Studies on French, German, and Dutch causal connectives have already shown that the distribution of connectives over objective or subjective causal categories seems to vary in relation to the context (Degand & Pander Maat 2003; Frohning 2007; Pit

2003; Zufferey 2012). Subjective causal connectives such as French *car*, German *denn*, and Dutch *want*, which are all roughly translated as ‘because’, display consistent usage patterns across text types such as newspapers, novels, and periodicals. By contrast, usage patterns of their objective counterparts, French *parce que*, German *weil*, and Dutch *omdat*, are less consistent across these genres (Stukker & Sanders 2012).

2.4.1 Objects of study

A large proportion of previous studies focus on the synchronic usage of Chinese causal connectives (Biq 1995; Deng 2007; Gao 2013; Guo 2006, 2008; Li 2011; X. Li 2009; Li & Liu 2004; Li & Wang 2013; Qu 2002; Song & Tao 2009; Wang 1999; Xing 2002; Zhang & Zhang 2011; Zhao 2003). Among them, some focus on polysemic uses of a particular connective, and some aim for describing similarities and differences between causal connectives. As a common practice, the number of causal connectives studied at a time is around one or two, and hardly exceeds three. Consequently, no study has presented an overall view of the distributional differences of the whole set of causal connectives that are commonly used in Chinese discourse. This can be seen as a limitation of these synchronic studies in general.

Another limitation regarding the object of study is that the less frequent causal connectives such as *kejian* ‘so’ have not yet been given due attention. For studies dealing with lexical polysemy, the most frequent causal connective *yinwei* ‘because’ seems to have attracted the most research interest (Biq 1995; Gao 2013; Song & Tao 2009; Wang 1999; Zhang & Zhang 2011). As to the comparative studies, the center of attention also seems to be around *yinwei*. Li and Wang (2013) have compared *yinwei* with the result connective *suoyi* ‘so’. A couple of studies have compared between *yinwei* and *youyu* ‘because’ (Li 2011; Qu 2002; Xing 2002). For studies dealing with result connectives (all translated as ‘so/therefore’), again they seem to center around the most frequent ones *suoyi* and *yinci*, comparing between *suoyi* and *yushi* or between *yinci* and *yin'er* (Deng 2007; Guo 2006, 2008; Li 2011; Li & Liu 2004), or between *yinci*, *yushi*, and *cong'er* (Zhao 2003).

Linguists with expertise in Ancient Chinese focus on the diachronic development of causal connectives (Jiang 2010; Li 2012; W. Li 2009; X. Li 2009; Li & Tang 2007; Qiu 2006; W. Wang 2002; Zhang 2012). All these studies aim to examine the process of grammaticalization and subjectification (or subjectivisation in the terms of Finegan 1995), using historical data across different periods of time. In addition, some diachronic studies also show interest in the factors that have motivated grammaticalization, such as frequency and repetition (Li 2012; Zhang 2012), or the mechanisms of grammaticalization, such as metaphor, pragmatic inference, and reanalysis (Jiang 2010; Li 2012; X. Li 2009; Li & Tang 2007;

Zhang 2012). As introduced in the first chapter, this dissertation will focus on the synchronic differences between Chinese causal connectives. Therefore, the diachronic studies are not introduced in detail. Nevertheless, I am in agreement with the view that the historical origin of a specific causal connective can constrain its current usage (W. Li 2009; X. Li 2009). Therefore, in the upcoming chapters I refer to the findings about the connectives' historical origins as a basis for predicting their current uses.

Of all the literature collected for this chapter, no studies were found that aim at examining how the order of acquisition is affected by the degree of subjectivity which each Chinese causal connective expresses. Neither have experimental studies been observed that seek to investigate the impact of subjectivity on the online processing of Chinese discourse. In the limited number of subjectivity studies on the synchronic differences between Chinese causal connectives (Deng 2007; Gao 2013; Guo 2008; Li 2011; Li & Liu 2004; Xing 2001), no cross-linguistic comparisons have been attempted. In view of the development of studies on subjectivity in European causal connectives, it seems that Chinese causal connectives have yet to receive a more extensive study in terms of subjectivity, from the perspective of different linguistic disciplines.

2.4.2 Methods

Some of the synchronic studies on Chinese causal connectives reviewed here are corpus-based and quantitative. Either written corpora (Guo 2006; Li 2011; Li & Liu 2004; Li & Wang 2013; Qu 2002; Xing 2002; Zhang & Zhang 2011), or spoken corpora (Song & Tao 2009), or both (Biq 1995) are used. Others stick to the traditional qualitative method (Deng 2007; Gao 2013; Guo 2008; Wang 1999; Xing 2001; Zhao 2003). Among all the synchronic studies collected for this chapter, only six were found that refer to the notion of subjectivity as an explanation for the discovered differences in use between the various causal connectives (Deng 2007; Gao 2013; Guo 2008; Li 2011; Li & Liu 2004; Xing 2001). Four of them are qualitative studies (Deng 2007; Gao 2013; Guo 2008; Xing 2001) and two of them are corpus-based quantitative studies (Li 2011; Li & Liu 2004).

One advantage of corpus-based studies is that they examine naturally-occurring linguistic data rather than examples made up by the scholar(s) actually conducting the work. This can secure the validity of the individual examples as supporting evidence. Another advantage concerns the fact that corpus-based studies seek balance and representativeness of the data at hand for the whole language population under investigation (McEnery & Hardie 2012). Representativeness refers to the ability to fully capture the variability of a language (Biber 1993), which is commonly achieved by using an unbiased (e.g., randomized) subsample of the corpus. This kind of

representativeness allows for findings to be generalizable to a larger language population.

A major limitation is apparent when we look at the statistical methodology used in these corpus-based studies. They tend to avoid carrying out quantitative analyses beyond the simplest of descriptive statistics: percentages. Percentages, as the most basic descriptive techniques, are particularly useful for summarizing data (Cantos 2009). However, percentages simply describe the data, and do not test for significance (McEnery & Hardie 2012). Hence, from percentages alone, it is difficult to tell how likely it is that a particular result occurs due simply to chance. Those conducting future research should be encouraged to consider using more advanced statistical tests to sort out significant differences from non-significant differences. Moreover, inferential statistics that allow us to test significance provide a distinct advantage to corpus linguistics because they enable linguists to generalize conclusions beyond actual observations (see Núñez 2007). It is not always possible or practical to study a whole data set. Corpus-based studies usually take a representative portion (sample) of the whole corpus. Inferential statistics provide the bridge between the properties of a sample and the population it has been selected from.

Other studies are purely qualitative in the sense that they use data only for describing features of language use. No attempt is made to assign frequencies, percentages and the like to the linguistic features identified in the data. Nevertheless, these qualitative studies have offered rich and detailed perspectives on the meaning and use of Chinese causal connectives. They contribute to identifying and describing categories involved in connective use (which are discussed in detail in section 2.3). In the future, we can seek to supplement these qualitative studies with quantitative analysis. As Schmied (1993) has observed, a stage of qualitative research is often a precursor for quantitative analysis, because the categories for classification must first be identified before they are counted (see also McEnery & Wilson 2001).

Furthermore, considering the concrete subjectivity analyses provided by the corpus-based quantitative studies (Li 2011; Li & Liu 2004) in particular, I believe a somewhat more refined approach is in order. Domain alone has been explicitly used as an indicator of degrees of subjectivity encoded in causal connectives in these studies.⁴ Consequently, previous analyses have only coarsely divided meanings into objective and subjective. As is well acknowledged in western as well as in Chinese scholarship, the distinction

⁴ As introduced in section 2.3, Li (2011) and Li & Liu (2004) have attempted to explain the usage of causal connectives in terms of causal sequencing and the defeasibility of Q. However, these two factors were not used as subjectivity indicators as domain was, and were not used in quantitative analyses.

between objectivity and subjectivity is not straightforward. Subjectivity is often taken to be a matter of degree. Therefore, more subjectivity indicators have yet to be pinpointed and used for a finer distinction between different degrees of subjectivity. For example, Langacker's distinction between on-stage and off-stage conceptualization can be used to rate one utterance (concerning off-stage conceptualization) as being even more subjective than another utterance (concerning on-stage conceptualization). The previously-discussed causal categories, such as the defeasibility of Q, can also be converted into subjectivity indicators. Indications of the speaker's perspective, affect, and epistemic modality (which have frequently been used as subjectivity indicators by previous studies of expressions of subjectivity) should continue to catch the attention of future studies on subjectivity. When the range of subjectivity indicators have been decided on, we need to design a model to hold them. The analytical model as a shared standard can enable different coders to analyze the use of causal connectives in a consistent way. We do not always make accurate introspective judgments regarding the use of language. That is why different coders should be involved. Basically, an analytical model enables us to calculate the inter-coder reliability and thus ensure that the results are valid and replicable.

2.4.3 Genre

In the Chinese literature, we have not observed results or claims pertinent to the relationship between genre and degree of subjectivity encoded in causal connectives, although some synchronic studies did report that certain causal connectives prefer one genre over another. For example, Guo (2006) analyzes genre at an abstract level in terms of 'narrative', 'expository/descriptive', and 'argumentative' genres. This way of categorizing texts is in accordance with the communicative functions of the text (e.g., expression, information, persuasion). Based on results from a substitution test, Guo (2006) observes that *yushi* is mainly used in narrative texts (30), and that *suoyi* is mainly used in expository/descriptive texts (31) or argumentative texts (32).

- (30) *Yi qingzao women ting jiazhang pa dao hou cang de dingpeng, jie qi yi zhang dingpeng, ba dengban ge zai liang bian cang bi shang. Yi jiao deng shangqu, shou ba zhu duo. Yushi?suoyi qianmian de raofu jiu xia haozi, ba chuan cheng kai le.* (Guo 2006: 29)

'In the early morning, we heard that the master climbed to the canopy at the back of the boat, opened the canopy, and put pedal boards on two sides of the boat. He stepped on the pedal boards, and held the rudder with his hands. *Yushi?Suoyi* the

boatman at the front put the pole deep in the water, and shoved the boat off.’

- (31) *Ta de pifu hen xi, danshi you dian fa lü, suo yi/*yushi bing bu xian zhe zirun.* (ibid)

‘Her skin is fine and smooth, but is a little green, *suo yi/*yushi* it does not look soft.’

- (32) *Kao le 600 fen bucuo le, suo yi/*yushi ni yinggai gaoxing.* (ibid)

‘It is not bad to get a score of 600, *suo yi/*yushi* you should be glad.’

In (30), though, the genre does not in fact seem to play a crucial role in the process of the speaker choosing between connectives. Instead, the nature of the relation does. The relation between the event “the master holding the rudder” and the event “the boatman putting the pole in water” is strictly sequential. The two events occur one after another, with scarce causality in between. It follows that the connective *yushi* functions purely as a temporal connective in this case.⁵ The strictly causal connective *suo yi* does not fit the temporal relation, and is thus inappropriate for this context. As to cases such as (31) and (32), I am in agreement with the author’s intuitions. *Yushi* is generally not acceptable in these cases. Although these observations from the substitution test are only provisional, they hint at the impact of genre-specific properties on the use of connectives.

Xing (2002) has conducted a small-scale corpus-based study into novels and texts on political theory. The finding is that *youyu* is mainly used in argumentative texts, and is seldom used in narrative texts.

These two Chinese studies introduced above suggest that certain causal connectives are used more often in a specific genre than other connectives are, or than they are used in other genres. But the question is: when the connective is used in its less typical context, does it express different relations that are associated with different degrees of subjectivity? A systematic analysis is needed.

2.5 General discussion and conclusion

The major goal of this chapter was to inquire into the connection between the notion of subjectivity and the use of Chinese causal connectives. More specifically, the question addressed was: is the subjectivity account likely to provide a unified explanation for the differences between Chinese causal connectives observed in the literature? Furthermore, I explored the ways in which previous approaches can be refined so as to give a precise

⁵ As is introduced in chapter 1, the use of *yushi* is not restricted to causal relations. In some cases, *yushi* is strictly temporal (X. Li 2009; Xing 2001).

characterization of the subjectivity profiles of all major Chinese causal connectives across genres.

2.5.1 The relevance of subjectivity to the usage patterns of Chinese causal connectives

From previous studies on Chinese causal connectives, it can be seen that the Chinese lexicon of causality has been dealt with in terms of different conceptual and syntactic characteristics. Consequently, different outcomes have been produced as to the way in which one causal connective differs from another. However, an account that can explain the suggested differences between various causal connectives within a single and uniform framework is still lacking. The results from the present literature review show that the subjectivity approach is promising, as many of the analytical categories adopted in previous studies can be related to subjectivity. In fact, two analytical categories have already been linked to the notion of subjectivity. One of them concerns the distinction between the content, epistemic, and speech-act domains (Li 2011; Li & Liu 2004), and the other concerns the distinction between descriptive and inferential causality (Gao 2013; Guo 2008; Deng 2007). Apart from these distinctions, most previously-used analytical categories are, in essence, also relevant to the notion of subjectivity, although those studies have not set forth the link explicitly. These categories include the defeasibility of Q (Guo 2006; Li & Liu 2004), the modality of Q (Guo 2006, 2008; X. Li 2009; Zhao 2003), the distinction between ideational and interpersonal function (Biq 1995; Song & Tao 2009), causal sequencing (i.e., forward order vs. backward order) (Gao 2013; Guo 2008; Li 2011; Li & Liu 2004; Qu 2002; Zhang & Zhang 2011), the presence/absence of the so-called speaker subject (Li & Liu 2004), and tense and aspect (Zhao 2003).

2.5.2 Refining existing approaches to the issue of subjectivity in causal connectives

From previous studies on subjectivity and its manifestations in Mandarin Chinese, it is clear that Chinese linguists' views on subjectivity mainly derive from western theories of subjectivity, represented by Lyons (1977), Langacker (1990), and Traugott (1995, 2010). Subjectivity is generally defined in terms of the expression of self, speaker-relatedness, and on/off-stage conceptualization. Moreover, we have seen that the western-origin subjectivity approach has been applied to various linguistic phenomena in Chinese discourse. These studies have mainly focused their analyses on three aspects: the speaker's perspective, the speaker's affect, and epistemic modality.

Among these subjectivity studies, only a limited number focus on Chinese causal connectives (Deng 2007; Gao 2013; Guo 2008; Li 2011; Li & Liu 2004; Xing 2001). They have provided insightful suggestions as well as concrete corpus-based results. However, as pioneering studies of the type, the approaches they have adopted still need some refinement in the following directions. First, these studies usually make comparisons between only two causal connectives at a time. For an overall view of the distributional patterns of causal connectives in the Chinese discourse, more causal connectives should be included in the investigation. Therefore, we will investigate whether the subjectivity approach can offer a unified account of the distributional differences among the full range of causal connectives that commonly occur in the Chinese discourse.

Second, most subjectivity studies on causal connectives use qualitative approaches. They use examples made up by the authors or casually selected corpus data in an inductive way to identify and describe characteristics of language use in terms of subjectivity. The introspective remarks put forward in these qualitative studies are provisional, but have paved the way for further quantitative analyses of actual language use. With quantitative approaches and related statistical techniques, we can count the instances of subjectivity features identified by qualitative studies, and even more complex statistical models can be constructed to assess the significance of observed differences and to estimate how reliably the current findings, derived from a randomly selected sample, can be generalized to the language population from which the sample was drawn.

Third, we need to consider involving more subjectivity indicators, apart from just domain, in quantitative analyses, in order to specify more refined distinctions between different degrees of subjectivity. For example, the previously-discussed causal categories, such as the defeasibility of Q and the modality of Q, can be converted into subjectivity indicators. The distinction between on-stage and off-stage conceptualization, introduced by the review articles on subjectivity, can also be considered useful for real analysis.

Finally, we need to create a detailed analytical model with precise specifications and decision rules for calculating degrees of subjectivity, in order to ensure high inter-coder reliability and replication.

2.5.3 Genre as an independent factor

Western studies on connectives have shown the relevance of taking genre into account (Sanders & Spooren 2009, submitted; Stukker & Sanders 2012). Results from some Chinese studies also seem to imply that genre might play a role in shaping the meaning and use of causal connectives. Guo (2006) and Xing (2002) have reported that certain causal connectives are more frequently distributed in a specific text category than other causal

connectives in Chinese discourse. It is found that *yushi* is often used in narrative texts whereas *suoyi* usually occurs in expository/descriptive texts and argumentative texts (Guo 2006), and that *youyu* occurs more often in argumentative texts than in narrative texts (Xing 2002). This could be related to the subjectivity issue. So far, genre has not been explicitly taken into account as a factor, yet it might affect subjectivity in causal connectives. Future research should be expected to pay explicit attention to variations of the observed subjectivity profiles across different genres.

2.5.4 A multi-angle approach: connectives, language use and cognition

In European studies on the usage of causal connectives, insights have been obtained from different perspectives including purely introspective reflections, empirical investigations of naturally-occurring language materials (i.e., corpus-based studies), and areas such as discourse processing and language acquisition. It was found that causal connectives are often distinguishable in terms of subjectivity. Furthermore, causal connectives that differ in degrees of subjectivity show different patterns in language acquisition and discourse processing. Taken together, these studies seem to reveal that subjectivity is cognitively relevant. As Sanders and Spooren (2009) suggest, subjectivity is likely to be one of the cognitive principles that organize our knowledge of linguistic use of causal connectives.

If subjectivity is really cognitively relevant, we should expect that it is systematically encoded in Chinese discourse as well, and that the observed subjectivity categories manifest themselves in some way in cognitive activities, such as those involved in online discourse processing. As various corpora of written Chinese have recently become accessible, a few studies on Chinese causal connectives have set out to empirically put introspective reflections (proposed by previous qualitative studies) to the test. However, differences in subjectivity (or domain) have been used as descriptive entities, but to the best of our knowledge no attempts have been made to explore whether these categories also play a role during language processing. In this book, I endeavor to integrate corpus-based linguistic analysis with psycholinguistic work, deploying eye-tracking techniques that reveal internal cognitive processes.⁶ This dissertation consists of studies on the full range of Chinese causal connectives, based on systematic analysis into genuine corpus data, supplemented with an experimental eye-tracking study probing into the role of subjectivity in online discourse processing. I expect that results from corpus-based analyses, and results from the online processing study can provide converging evidence on the role of subjectivity in shaping

⁶ Although language acquisition also provides a window on human cognition, this book focuses on looking at the cognitive relevance of subjectivity in discourse processing.

the meaning and use of Chinese causal connectives, and will hopefully contribute to relating the use of language to general human cognition.

CHAPTER 3

Subjectivity and result marking in Mandarin: A corpus-based investigation¹

3.1 Introduction

With the rise of functional linguistics, pragmatics, and cognitive grammar, the concept of ‘subjectivity’ has gained increasing prominence (see De Smet & Verstraete 2006; Langacker 1985; Lyons 1977, 1982a, 1995; Shen 2001; Traugott 1989, 1995, 2010; Zhang 1994). According to Lyons (1977), the linguistic notion of subjectivity refers to the speaker’s self-expression in the utterance. Speakers often express their attitude and emotion towards the propositions in an utterance. For instance, speakers sometimes explicitly express their attitude by using evaluative adjectives, such as *wonderful* in the utterance *It is a wonderful stamp*. Alternatively, speakers can judge a situation by means of modal adverbs such as *probably* in the utterance *She probably fell ill*. Speakers can also express themselves by using other subjectivity markers such as prepositions (Langacker 1985) and connectives (Sanders & Sweetser 2009; Traugott 1995). The theme of this study is to examine the role of subjectivity in the use of result connectives in Mandarin Chinese.

Causal connectives are explicit linguistic markers of causal relations between adjacent segments (Murray 1997; Stukker & Sanders 2012). They can be used to mark the reason segment or the result segment of the causally related clauses. Accordingly, they divide into reason connectives and result connectives. In English, the result connective *so* can be used to mark the cause-consequence relation in (1), and the argument-claim relations in (2) and (3).

- (1) It was a very hot day, **so** Bob went swimming.
- (2) John is ill, **so** he probably won’t come to the meeting.
- (3) The lights are out in the house, **so** nobody is at home.

In (1), *so* marks the fact that *Bob went swimming* as the consequence of the situation that *it was a very hot day*. In (2) and (3), *so* is used to present a conclusion or an inference on the basis of a fact or an observation. English speakers also have other linguistic markers at their disposal that seem specialized in expressing specific types of causal relations. For instance, cause-consequence relations, such as (1), are often marked by the linguistic cue phrase *that’s why*, and the connective *therefore* is typically used to mark causal relations at an epistemic or argumentative level, as in (2).

In Mandarin Chinese, there are at least five connectives that can be considered to be the counterparts of the English *so/therefore*: *kejian*, *suoyi*, *yinci*, *yin'er*, and *yushi* (see Chinese Academy of Social Sciences 2002). Just as *so/therefore*, the semantics of the five Mandarin connectives denote that the segment containing them expresses the consequent of the causally related events or states. That is why they are called result connectives. In terms of their syntactic status, they are as highly grammaticalized as *so/therefore*: they are acknowledged conjunctions.² The usage of the five Mandarin connectives is illustrated in (1') – (3').

- (1') *Tianqi feichang de re, yushi/suoyi Bob qu youyong le.*
 'It was a very hot day, **so** Bob went swimming.'
- (2') *John bingle, yinci/yin'er ta keneng buhui lai kaihui le.*
 'John is ill, **so** he probably won't come to the meeting.'
- (3') *Fangzi li de deng heizhe, kejian meiren zai jia.*
 'The lights are out in the house, **so** nobody is at home.'

Although *kejian*, *suoyi*, *yinci*, *yin'er*, and *yushi* share great similarity in terms of semantics and syntax, they are only partially interchangeable in these contexts. For example, *yinci* can be substituted for *yushi* in (1') to express the causal connection between observed situations; however, *yushi* cannot be substituted for *yinci* in (2') to link a conclusion with its argument. The fact that Mandarin has at least five conjunctions to mark the consequent of causal relations and that these conjunctions are only partially interchangeable suggests a division of labor between these closely related connectives. How can this division of labor be characterized?

In this study, we will characterize these five Mandarin connectives in terms of subjectivity. One way of looking at subjectivity is to examine the type of causality marked by each connective. In the literature about English and other European languages, dichotomous distinctions have been made between external/internal causality (Halliday & Hasan 1976; Martin 1992), semantic/pragmatic causality (Moeschler 1989; Sanders 1997; Sanders, Spooren & Noordman 1992; van Dijk 1979), and subject matter/presentational causality (Mann & Thompson 1986, 1988). A more fine-grained categorization is proposed by Sweetser (1990), who distinguishes between three domains of causal relations: the content domain (which corresponds to the external/semantic/subject matter type of causality), the epistemic domain, and the speech-act domain. According to Sweetser (1990:77), the content domain concerns "real-world causality" holding between events in the described world; in the epistemic domain the speaker's knowledge is involved as the basis for a logical conclusion; and in the speech-act domain the reason clause gives the cause of the speech act embodied in the main clause.

Shen (2008) has applied the Domain Theory to the interpretation of complex sentences in Chinese discourse. He observes that the division between three domains helps to account for the semantic relations in causally related sentences such as (4) – (6). Sentences (4) – (6) can actually serve to exemplify the content, epistemic, and speech-act domain, respectively.

- (4) *Zhang Gang huilai le, yinwei ta hai ai Xiaoli.*
‘Zhang Gang has come back **because** he still loves Xiaoli.’
- (5) *Zhang Gang hai ai Xiaoli, yinwei ta huilai le.*
‘Zhang Gang still loves Xiaoli, **because** he has come back.’
- (6) *Zhang Gang hai huilai ma? Yinwei Xiaoli zai deng ta.*
‘Will Zhang Gang come back still? **Because** Xiaoli is waiting for him.’

Apparently, *yinwei*, just like *because*, can be used across domains. Now, it is interesting to ask whether or not there are causal connectives in Mandarin that specialize in specific domains, and whether the theory of subjectivity can help categorize causal connectives in Mandarin discourse. In fact, a number of studies have shown that causal connectives in European languages such as French, German, and Dutch can be categorized in terms of subjectivity (see Sanders & Sweetser 2009 for an overview). It is shown, for example, that subjectivity determines the choice of Dutch result connectives *dus* ‘so’ versus *daarom* ‘that’s why’ (Pander Maat & Sanders 2001) and Dutch reason connectives *omdat* ‘because’ versus *want* ‘because’ in a variety of discourse modes (Degand & Pander Maat 2003; Pit 2003; Sanders & Spooren 2009, submitted; Spooren, Sanders, Huiskes & Degand 2010). Investigations have also been made into other European languages such as French, German, and Polish (see Dancygier 2009; Evers-Vermeul, Degand, Fagard & Mortier 2011; Keller 1995). The results show that causal connectives of these languages can also be distinguished in terms of the degree of subjectivity they encode. For instance, French *car* and *puisque*, and German *denn* mark higher degrees of subjectivity in the sense that they are typically used to express epistemic causal relations that are constructed with higher speaker-involvement (see (2)). In contrast, French *parce que*, and German *weil* mark lower degrees of subjectivity in the sense that they typically involve lower speaker-involvement in constructing content causal relations as manifested in (1) (see Stukker & Sanders 2010b, 2012).³

The studies mentioned above show that subjectivity plays an important role in categorizing causal connectives in major European languages. In fact, the theory of subjectivity is derived mainly from Germanic data. If we take seriously the proposals that the linguistic categories apparent in people’s everyday language use provide us with insights into the working of the mind

(Lakoff 1987; Lakoff & Johnson 1999) and that the notion of subjectivity is likely to be a general cognitive mechanism underlying the categorization of causal connectives across different languages (Sanders & Sweetser 2009), then it would be reasonable to assume that subjectivity as a general factor should manifest in other language families as well. Therefore, we hypothesize subjectivity to be relevant to the categorization of causal connectives in Mandarin, in spite of the fact that Mandarin is typologically different from most of the major European languages.

Indeed, in the literature on Mandarin causal connectives there have been theoretical as well as corpus-based studies that suggest the possibility of categorizing Mandarin causal connectives with respect to subjectivity (Li 2011; Li & Liu 2004; Shen 2008; Xing 2001; see also Y. Zhang 2012 for the categorization of reason connectives in terms of intersubjectivity). Referring to examples (4) – (6), Shen (2008) suggests that causal relations in Mandarin discourse can be categorized into different domains of causality, and therefore reflect different levels of subjectivity. Furthermore, Xing (2001) demonstrates systematic differences among a number of reason connectives in Mandarin discourse, as exemplified in (7) and (8).

- (7) *Yinwei* laoshi you shi, xiaowu women zixi.
 ‘**Because** the teacher has some other business to do, this afternoon we study by ourselves.’
- (8) *Jiran* laoshi you shi, xiaowu women zixi.
 ‘**Since** the teacher has some other business to do, this afternoon we study by ourselves.’

According to him, *yinwei* as used in (7) is to mark the ‘descriptive causality’ on the basis of facts, whereas *jiran* ‘since’ as used in (8) typically expresses an ‘inferential causality’ on the basis of rationality. We believe Xing’s concepts of descriptive causality and inferential causality to mirror the distinction between the content and the epistemic domains. In (7), the content causal relation holds between two observable situations, and there is no speaker involvement. In the epistemic relation in (8), by contrast, the speaker is directly involved in linking an argument with a conclusion, and thus we consider (8) to be more subjective than (7).

The examples provided by Xing (2001) only concern reason connectives (i.e., causal connectives that are used to mark the reason segment). Although the author does not provide examples to illustrate the systematic difference between result connectives, he does remark that the result connectives *suoyi*, *yinci*, and *yin’er* typically occur in descriptive causality, whereas *kejian* typically marks inferential causality (Xing 2001:40). On the basis of this claim, we should expect to find more content use of the first three result connectives, and more epistemic relations in sentences connected by *kejian*.

Xing (2001:527) claims that in clauses connected with *yushi*, the second clause often expresses a natural consequence of the state of affairs, judgment, or some kind of feeling expressed in the first clause. This seems to suggest that *yushi* occurs more often in the content domain to describe cause-consequence relations existing in the physical world. This is an issue we will investigate in this study.

To sum up, previous studies touched upon the issue of subjectivity in causal marking, reason marking in particular, in Chinese language and discourse. Concerning result marking, however, no systematic corpus-based investigations have been undertaken so far to explore conceptual differences among result connectives in terms of subjectivity. In this study, we conducted a corpus-based analysis of the five Mandarin result connectives that are apparent equivalents in terms of both syntactic and semantic properties – *kejian*, *suoyi*, *yinci*, *yin'er*, and *yushi* – to study whether and how subjectivity categorizes causal connectives in a non-Indo-European language such as Mandarin Chinese. Responding to Research question 4 and 5 raised in Chapter 1, this chapter particularly seeks to address the following questions:

1. Are previous claims (e.g., the ones in Xing 2001) about the usage of Mandarin result connectives supported by the corpus study?
2. Do the five result connectives *kejian*, *suoyi*, *yinci*, *yin'er*, and *yushi* show systematic variation in terms of subjectivity?
3. Are the observed subjectivity profiles sensitive to text genres?

Studies on French, German, and Dutch causal connectives have already shown that the distribution of connectives over causal categories seems to vary in relation to the context (Degand & Pander Maat 2003; Frohning 2007; Pit 2003; Zufferey 2012). Subjective causal connectives such as French *car*, German *denn*, and Dutch *want*, which are all roughly translated as ‘because’, display consistent usage patterns across text types such as newspapers, novels, and periodicals. By contrast, usage patterns of their objective counterparts, French *parce que*, German *weil*, and Dutch *omdat*, are less consistent across these text genres (Stukker & Sanders 2012). These results raise the question whether the subjectivity profiles are really part of the inherent semantic characteristics of the connectives themselves, or whether they are (partially) determined by the pragmatics of the context.

The paper is organized as follows. In Section 3.2, theories about the linguistic notion of subjectivity are discussed in detail as the basis for the operationalization in our corpus-based analysis. In Section 3.3, we introduce our research method. In Section 3.4, the results of the corpus analysis are reported, showing how exactly the five causal connectives display differences along the dimension of subjectivity. In Section 3.5, the results are discussed

further and suggestions are made for future studies.

3.2 Subjectivity defined

In order to give a precise characterization of the degrees of subjectivity expressed by the five result connectives, we adopt an integrative approach to the issue of subjectivity. In spite of various differences in their formulations, most definitions refer to subjectivity as a specific property of the utterance: speaker-relatedness. For example, Traugott (2010:30) defines subjectivity as the “relationship to the speaker and the speaker’s beliefs and attitudes”. In the same vein, De Smet and Verstraete (2006:365) claim that subjectivity “covers the fact that a particular element or construction requires reference to the speaker in its interpretation”. While focusing on discourse, Sanders and Spooren (2009, submitted) argue that speaker involvement is an important property of subjectivity in causally related sentences as well as in isolated sentences. Examples (9) and (10) illustrate their idea.

- (9) *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 planted with a hormone that can stimulate the production of blue pigments, **as a result**, the petals are blue.’
- (10) *Zhe zhong yao hanyou weiliang de youdu wuzhi, yinci changqi fuyong hen keneng dui jiankang buli.*
 ‘This kind of medicine contains a tiny amount of poisonous substance, **so** long-term use (of the medicine) is probably detrimental to the health.’

In (9), the relation between the antecedent and the consequent is objective, because it concerns a relation in an objective reality, which does not depend on the speaker’s reasoning or intentionality. In contrast, (10) is subjective because it involves a subjective construction of the causal connection, which is not apparent in the physical world but exists in the speaker’s mind. For cases like this, we need to refer to the speaker’s attitude for its interpretation. Therefore, on the basis of the property of speaker-relatedness we can make a division between subjectively constructed and objectively described causal relations.

In order to track any differences in the degree of subjectivity of the Mandarin result connectives, we go beyond the absence versus presence of speaker-relatedness. On the basis of earlier work on causal connectives (e.g., Pander Maat & Sanders 2000), subjectivity has also been defined in terms of the presence of an active ‘subject of consciousness’ (henceforth SoC) (see Sanders & Spooren submitted; Sanders & Sweetser 2009). An SoC is “an

animate subject, a person, whose intentionality is conceptualized as the ultimate source of the causal event, be it an act of reasoning or some real-world activity” (Pander Maat & Sanders 2001:251). The SoC may be the actual speaker or the character(s) in the utterance. Accordingly, the notion of subjectivity has been extended beyond speaker-relatedness: it could be character-related as well. To give an illustration, (11) is considered subjective because it involves a conclusion drawn by a speaker SoC, while (12) can also be called subjective because it involves the reasoning of a character SoC.

(11) The lights are out in the house, so I think nobody is at home.

(12) The lights are out in the house, so John thinks nobody is at home.

Both (11) and (12) contain a subjectively constructed causal relation because they are both SoC-related. They involve the SoC’s point of view towards the state of affairs in the world. The author/speaker can be considered the first voice in the discourse, who has constant access to her feelings and thoughts. She does not have access to the feelings and thoughts of a third person. As a result, *I think Utrecht is nice* can be a direct report of an inner feeling, whereas *He thinks Utrecht is nice* is a description of an evaluation. As a consequence, first person evaluations are more subjective than third person evaluations. What distinguishes the speaker/author from a character as SoC is that the speaker concerns a first voice, which is grounded in the Deictic Centre of Communication (Sanders, Sanders & Sweetser 2009). This reflects Traugott’s (1989, 1995) view on subjectivity as closeness to the communicative “here and now”: the speaker/author here and now asserts that a particular state of affairs holds. By contrast, the character type concerns a third person in the discourse, who is more distant from the Deictic Centre of Communication. The notion of SoC, and the distinction we make here between speaker/author and character SoC, is related to the notion of *perspective* in literary studies (Fludernik 1993) and *voice* in linguistics (Ducrot 1980): the speaker/author is the first voice in the discourse, and the character is another person whose inner thoughts and evaluations can be accessible in a narrative (Sanders 2010).

Closely related to the concept of SoC is the Domain Theory, which we have briefly mentioned in the Introduction. Many previous studies show that the division between causal domains is an effective way to operationalize subjectivity in causally related sentences (see Degand & Pander Maat 2003; Evers-Vermeul et al. 2011; Sanders & Spooren 2009). These and other studies rank causal domains in terms of subjectivity according to the degree to which they necessarily imply the subjective involvement of an SoC. In the content domain, the speaker describes a causal relation in the physical world, as in (9). This content domain can be subdivided into two subtypes (Stukker,

Sanders & Verhagen 2008): the volitional content domain always involves an SoC who performs an intentional act (see (13)), whereas the non-volitional content domain does not have an SoC at all (see (14)).⁴

(13) It was a very hot day; that's why Bob went swimming.

(14) Bob fell off the bike. As a result, he got hurt in the left leg.

Due to the absence of an SoC, the non-volitional content domain is considered the least subjective domain: it concerns causal processes in which human intentions do not play a role. Volitional relations carry a higher degree of subjectivity than non-volitional ones, because human volition or motivation is a prerequisite for performing acts. Epistemic relations are of still higher subjectivity because they directly involve speakers' or characters' opinions, beliefs or inferences, as in (10) – (12). Speech-act causal relations are also highly subjective because performing a speech act (making a promise, issuing a command, or raising a question, as in (6)) is bound to a situation in which the speaker is present.

As a final part of our integrative approach to subjectivity, we adopt Langacker's (1990) interpretation of subjectivity in terms of the explicitness of a ground element, an element that refers to the speech event, its participants and its immediate circumstances. From Langacker's point of view, the defining standard of subjectivity is not only speaker-relatedness, but also whether or not the speaker remains implicit in the utterance, without any formal linguistic encoding. An utterance is objectified when the speaker is put "on stage" and becomes observable (see Langacker 1985, 1990). For example, despite the fact that the causal relation expressed in (10) (repeated here as (15) for convenience) is identical to the one in (15'), we consider (15') as less subjective than (15) because the speaker is on stage, as an explicit reference point.

(15) *Zhe zhong yao hanyou weiliang de youdu wuzhi, yinci changqi fuyong hen keneng dui jiankang buli.*

'This kind of medicine contains a tiny amount of poisonous substance, **so** long-term use (of the medicine) is probably detrimental to health.'

(15') *Zhe zhong yao hanyou weiliang de youdu wuzhi, yinci wo renwei changqi fuyong hen keneng dui jiankang buli.*

'This kind of medicine contains a tiny amount of poisonous substance, **so** I think long-term use (of the medicine) is probably detrimental to health.'

3.3 Corpus-based analysis

3.3.1 Mandarin result connectives

In the present study, we aim to investigate the subjectivity profiles of five Mandarin result connectives in a quantitative corpus-based analysis. We focus on single words, highly grammaticalized *guanxici* ‘connectives’ that are used to mark the consequent clause: *kejian*, *suoyi*, *yinci*, *yin'er*, and *yushi*.⁵ The frequencies of these result connectives are different (see Table 1). *Yinci* is the most frequent of the five, followed in turn by *suoyi*, *yushi* and *yin'er*. The least frequent one is *kejian*.

Table 1. Frequency of the five result connectives (based on *the Lancaster Corpus of Mandarin Chinese*).⁶

Connective	Absolute frequency	Frequency per 10,000 words
<i>yinci</i>	439	4.39
<i>suoyi</i>	393	3.93
<i>yushi</i>	274	2.74
<i>yin'er</i>	150	1.50
<i>kejian</i>	59	0.59

In the *Contemporary Chinese Dictionary* (Chinese Academy of Social Sciences 2002), *yushi* is annotated as indicating that the latter event immediately follows the former, and that the latter is often caused by the former (see also Lü 1999). On the grounds that it describes causal connections between physical events, we may predict that *yushi* expresses causality within the content domain. This prediction is also in line with Xing’s interpretation of the word (as discussed in Section 3.1). It also conforms to *yushi*’s closeness to the descriptive temporal usage (Xing 2001; Y. Zhang 2008). Xing (2001) labels *yushi* as a ‘cross-relation marker’, because it is used to mark a relation of temporal succession as well as causality. For instance, in (16) *yushi* is strictly temporal; it merely expresses the successive order of two connected events (see Xing 2001:528). Still, *yushi* can also express relations that are solely causal. For example, in (17), *yushi* connects a conclusion with its argument, and temporal ordering is irrelevant. Lu (2000) claims that *yushi* cannot express inferences or conclusions, but (17) already serves as a counterexample. We will testify this further with the corpus data.

- (16) *Guole na lin, chuan bian wanjinle Linggang, yushi Zhaozhuang bian zhen zai yanqian le.*
 ‘The boat passed the woods, sailed into the Ling Harbor, **and then** Zhao Village was indeed in front of us.’

- (17) *Ji meiyou youxiao de jingji zeren shenji, you quefa biyao de juece shiwu zhuicheng, yushi chaobiao jianzhu, “lanweigongcheng” de chuxian, ye jiu bu lingrenchayi le.*

‘There is neither effective auditing of financial responsibility nor essential penalty for mistakes on a decision, so it is not surprising that super-luxurious buildings and wasted real estate projects appeared.’

The dictionary annotation of *kejian* presents us a totally different picture. *Kejian* is translated as ‘it is thus obvious that, it shows that, that proves, so’. Clearly, its English translations are all closely related to opinions, beliefs, or conclusions, on the basis of which we may predict that *kejian* expresses causality within the epistemic domain. This prediction conforms to the word’s lexical origin. The conjunction *kejian* was grammaticalized from a verbal phrase *ke jian* (Liu & Yao 2011; Q. Zhang 2012), which can be roughly glossed as ‘can see’. Given this root meaning, it seems logical to expect that *kejian* is used in the subjective epistemic domain, expressing the speaker’s conclusion. In particular, the word *jian* ‘see’ has been demonstrated as a subjectivity indicator, which encodes perspective, hence the increased involvement of the speaker/writer (see Tao 2007 for evidence of *jian* indicating subjectivity in the context of existential/presentative constructions). In fact, this prediction about the use of *kejian* is also consistent with Xing’s (2001) claim that *kejian* typically expresses inferential causality on the basis of rationality.

The dictionary annotation does not give us a clear clue on the profile of *yinci*, *yin'er*, or *suoyi*, because only the English translation *so/therefore* is given in the dictionary and no further explanation can be found there. However, Xing (2001) has observed that these three connectives typically express descriptive causality on the basis of facts. Accordingly, we should expect them to occur more often in the objective content domain.

3.3.2 Corpus and sampling method

In order to study the impact of text genre on the connective meaning and use, i.e., the third research question, we collect the samples in a balanced way from three genres. We choose to analyze these three categories because they are generally assumed to be associated with different communicative purposes – information, expression, and persuasion – which are likely to have an impact on the degree of the subjectivity of the text. For example, opinion pieces typically express the writer’s point of view and aim to convince the reader by presenting sound reasoning and arguments, and therefore are likely to display an overall higher degree of subjectivity than novels and news reports (see Sanders 1997). As for news reports, the main goal is to

describe and inform something unknown and unfamiliar to the readers, and hence we should expect to read more objective information in such a genre. Novels are different from the other two text categories in that they tell stories and narrate events that characters perform, and hence are related to human intensions/motivations. By taking these three distinct text categories into analysis, this dissertation aims to investigate whether and how genre shapes or constrains the use of causal connectives, specifically with respect to subjectivity.

This sampling method is designed also to minimize any contextual bias towards discourse interpretation; as earlier studies have shown that in canonical cases contextual factors play a smaller role in determining the coherence relation, but that context can have a major impact on the interpretation when the relation is ambiguous (Sanders 1997). Language users show a systematic preference to interpret ambiguous cases as semantic in descriptive contexts, and a tendency to interpret them as pragmatic in argumentative contexts.

The news reports and opinion pieces are taken from *People's Daily Online* (abbr. PPD), which is a comprehensive and influential website (http://search.people.com.cn/rmw/GB/rmwsearch/gj_search_pd.jsp) claiming the largest daily amount of news releases in Mainland China. As it has a database of 20 billion characters, it is an excellent data source for analyzing newspaper texts. News reports are taken from the category of *News on Technology*, and the opinion pieces are taken from the category of *Opinion*.

Although *People's Daily Online* is a large database for linguistic investigations into Mandarin Chinese, it is limited to newspaper articles. Thus, we look to another corpus for novels: the CCL Corpus. This corpus, created by the Center for Chinese Linguistics of Peking University, consists of 477 million characters. The contemporary Chinese part of the corpus is divided into various subject areas, such as *Novel*, *Prose*, and *Biography*.

For each genre, seventy-five occurrences of each connective were randomly selected. See Table 2.

Table 2. Number and nature of connective fragments in the sample

Connective	News report (PPD)	Opinion piece (PPD)	Novel (CCL)	Total
<i>kejian</i>	75	75	75	225
<i>suoyi</i>	75	75	75	225
<i>yinci</i>	75	75	75	225
<i>yin'er</i>	75	75	75	225
<i>yushi</i>	75	75	75	225
Total	375	375	375	1125

For the 1125 fragments, special care has been taken to make sure that the connectives are used as causal connectives in every instance. For example,

the temporal usage of *yushi* as in (16) is not included in our collection of fragments. In our final sample, four cases of *yushi* are discarded due to its temporal usage. In addition, cases such as (18) are also excluded, because *ke jian* is used as a predicate ‘may be seen’ rather than as a connective. In total, six occurrences of verbal *ke jian* are replaced with causal uses.

- (18) *Baifa yi qingxi ke jian.*
 ‘Grey hair *can* already **be seen** clearly.’

3.3.3 Analysis

In order to systematically measure the degree of subjectivity each causal connective expresses, we use an analytical model to operationalize the notion of subjectivity. The analytical model is created based on the integrative approach to subjectivity described in Section 3.2. It contains all the important components of subjectivity: domain, the presence of an SoC, and the identity of the SoC. Apart from these three factors, we take a fourth variable into consideration: propositional attitude of Q (i.e., the consequent). We decide to add this variable because propositional attitude has been shown to be an important indicator of subjectivity in previous analyses (Sanders & Spooren 2009). Empirically, it also proved to be effective for several Germanic languages (Degand & Pander Maat 2003; Pander Maat & Degand 2001; Spooren et al. 2010).

Our analytical model is presented in Table 3. It shows the above-mentioned four variables together with their subjectivity values, to be elaborated on in the following subsections.

Table 3. Model for subjectivity analysis with variables and their values

Variable	+subjectivity value -
Domain	Speech-act / Epistemic Volitional content Non-volitional content
Propositional attitude	Speech-act / Judgment Mental fact Physical fact
The presence of SoC	Implicit Explicit Absent
The identity of SoC	Author..... Current speaker..... Character

In the literature, the distinction between subjectivity and objectivity is considered to be gradual rather than absolute (Lyons 1982b:105). Therefore, we do not calculate an exact value of subjectivity for the connectives. Instead, we compare the relative degrees of subjectivity that are encoded in different connectives. For each connective, we use the four variables to examine the degree to which the speaker or SoC is responsible for the construction of the causal coherence relation. A connective is considered to signal a higher

degree of subjectivity than another if there is at least one variable that discriminates between the connectives in the more subjective direction (e.g., more epistemic cases) and if none of the other variables shows a preference in the more objective direction (e.g., more physical facts). The fragments were coded independently by the author and two other native speakers of Mandarin, who were working in the field of discourse studies at the Utrecht Institute of Linguistics OTS. The level of inter-rater agreement is quite high (for domain: Kappa = 0.93; for propositional attitude of Q: Kappa = 0.94; for presence of SoC: Kappa = 0.90; for identity of SoC: Kappa = 0.88). The team discussed discrepancies among their analyses until they reached an agreement.

Domain

As discussed in Section 3.2, we distinguish between four causal domains in the present study: the non-volitional content, volitional content, epistemic, and speech-act domain. In order to interpret the domains accurately, we use a paraphrase test. The paraphrase test is presented in Table 4, in which P and Q correspond to the antecedent and the consequent of the causal relation, respectively.

Table 4. The paraphrase test used in the domain analysis

Domain	Paraphrase
Non-volitional content	P leads to the physical fact / mental fact that Q, and no intention is involved in Q
Volitional content	P leads to intentional physical act / mental act that Q
Epistemic	P leads to claim / decision / inference / conclusion that Q
Speech-act	P leads to question / advice / command / promise that Q

However, in natural discourse the division between domains is not always clear-cut. In cases where ambiguity emerges, we opt for the more objective interpretation. For example,

- (19) *Youyu tamen de zhongzi cunzaizhe xiumianqi, luo di hou bu liji mengfa, ershi yao dengdao cinian chuntian lailin cai faya, yin'er nenggou duo guo dongji yanhan.*

‘Their seeds have a period of dormancy. They don’t sprout immediately after falling to the ground, but will wait till the next spring. **As a result/Therefore** they can survive the severe cold during the winter.’

The ambiguity of the causal relation expressed in (19) springs from the lexical ambiguity of the word *nenggou* ‘can’, which may denote capability as well as possibility. Consequently, we can plausibly arrive at two interpretations of (19). One way to interpret it is to treat the word *nenggou*

‘can’ as expressing a kind of possibility: the fact that the seeds have a period of dormancy leads to a conclusion that it is possible for them to survive the severe cold during the winter. This can be regarded as a subjective epistemic relation. The other interpretation is to construe the *yin’er* segment as expressing an observable fact: the fact that the seeds have a period of dormancy leads to the fact that they are able to survive the severe cold in winter. This can be seen as denoting an objective content relation. In the analysis, we interpret the fragment in the latter way, that is, we have adopted an objectivity bias, for the sake of consistency. In fact, an ambiguity like this is often disambiguated by the context. For example, if we precede the sentence with the phrase “In his view,” (see (20)), the second interpretation is immediately ruled out.

- (20) *Zai ta kanlai, youyu tamen de zhongzi cunzaizhe xiujianqi, luo di hou bu liji mengfa, ershi yao dengdao cinian chuntian lailin cai faya, yin’er nenggou duo guo dongji yanhan.*

‘In his view, their seeds have a period of dormancy. They don’t sprout immediately after falling to the ground, but will wait till the next spring. **Therefore** they can survive the severe cold during the winter.’

We have encountered a special type of causal relation during the analysis, which we term ‘chain causality’. In the so-called chain causality, readers need to add an inferential link to the causal relation to comprehend the sentence. For example,

- (21) *(Baishikele Gongsi) zai qishui pinggai shang yinyou shuzi haoma, dajiang you yibai wan bisuo. Yushi Baishikele daweichangxiao.*

‘(The Pepsi-Cola Company) printed numbers on the bottle-caps; the biggest bonus was one million peso. **So** Pepsi was sold very well.’

In this case readers need to add the following inferential link: many customers intentionally went to buy Pepsi-Cola in order to win the big bonus. The next question to consider is: what is the causal domain of (21)? On the one hand, there is an implied SoC in it: the customers. On the other hand, the best paraphrase is “the fact that the Pepsi-Cola Company offered a big bonus leads to the fact that it was sold very well”, which suggests that the domain is non-volitional content. During the analysis, we always choose the non-volitional content interpretation, to be consistent with the way in which we deal with ambiguous relations. That is also the case here.

Propositional attitude

We distinguish four values of propositional attitude. We code the result segment as expressing a speech act, a judgment, an observable physical fact, or a mental fact. A segment is coded as a speech act if it is in the form of a general question, a rhetorical question, or an imperative. A segment is considered to express a judgment if it presents opinions, decisions, conclusions, or inferences. A segment is coded as a physical fact if it describes events or states that take place or exist in the observed world, and as a mental fact if it depicts mental states such as personal feelings, mental processes, or psychological activities. For example, we code the consequent in (21) as a physical fact, and the consequent in (15) (repeated here as (22) for convenience) as a judgment. (23) is considered to have a speech-act consequent, and (24) a consequent representing a mental fact.

- (22) *Zhe zhong yao hanyou weiliang de youdu wuzhi, yinci changqi fuyong hen keneng dui jiankang buli.*

'This kind of medicine contains a tiny amount of poisonous substance, **so** long-term use (of the medicine) is probably detrimental to health.'

- (23) *Ta de tizhi hen buhao, zhe shi yiding de. Suoyi yao kuai, women zou ba.*

'Her physique is very bad, this is true for sure. **So** be quick, let's go.'

- (24) *Ta zheng nuli huixiangzhe dangchu lingchen yi jiao cai jin equan shi de qingjing, yushi ta fangfu you tingdaole equan yinwei jinghuang fachu de jiaosheng.*

'He is trying to recall the situation in which he stamped into a flock of geese in the early morning, **so** he seems to hear the calls made by the flock of geese out of panic.'

Mental facts and judgments are located in the inner world. They are not directly observable in the external world. By contrast, physical facts exist or take place in the external world. The distinction is clear-cut. However, the boundary between mental facts and judgments is less clear, since they both involve mental processes. To make the distinction, we examined whether an opinion is formed or a decision is made during these mental activities. For example, in (24) he felt as if he heard the calls of a flock of geese, but no opinion crossed his mind. It is a description of a mental illusion. By contrast, a typical judgment is formed in (22). Sometimes, a judgment is not as typical as the one in (22). For instance, we consider the second segment of the following sentence (25) as a judgment. Some analysts would probably argue that it is a description of a fact. In our view, one needs to make judgments

before making decisions. Decisions are closely connected to judgments, and thus should be analyzed as judgments.

- (25) *Xia yu le, yinci tamen jue ding qu xiao yecan huodong.*
 'It started to rain, **so** they decided to cancel the picnic.'

Physical facts are considered to be less subjective than mental facts, judgments and speech acts because physical facts do not involve mental processes. Mental facts are less subjective than judgments and speech acts since more cognitive efforts are involved on the part of the SoC to form an opinion, make a conclusion, or ask a question.

The presence of SoC

We distinguish between utterances with no SoC, those with an implicit SoC, and those with an explicit SoC. There is no SoC in utterances where the causal relation exists independent of volition. For example,

- (26) *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.'

In (26), the fact of living in the same area automatically leads to the fact of encountering each other often. The causal connection has nothing to do with human volition: *we* did not intentionally meet each other. There is no SoC in it. When there is an SoC, we make a distinction between implicit and explicit SoCs. If the SoC is explicitly referred to by linguistic elements in the consequent, we will consider the SoC to be explicit (as in (27)). If the SoC is not linguistically realized, we will take it to be implicit (see (28)).

- (27) *Ta tingdao men wai you dongjing, yushi ta tui kai men qu kan ge jiu jing.*
 'He heard some sounds outside the door, **so** he pushed open the door to see what actually happened.'
- (28) *Xiaodunzi an kai dianshiji, chuxian de huamian yanse te dan, kejian xianxiangguan yiran laohua.*
 'Xiaodunzi pressed on the television, the color of the picture was particularly light, **so** the kinescope has already been aging.'

Importantly, when *I* or *me* occur in a clause, we need to determine whether it functions as the SoC. Take (29), for example. The *me* in the consequent clause does not function as an SoC. We can easily replace *me* with *him* or *her*, and meanwhile the SoC will still be the speaker. Hence, we

label the SoC implicit for (29). The explicit counterpart is provided in (30), in which the concluder is explicitly realized by the pronoun *I*. Again, we may use the substitution test to check whether *I* is indeed the SoC. If we replace *I* with *Xiaoli* (see (31)), the SoC is changed to *Xiaoli* simultaneously. So, *I* is considered functioning as the SoC in (30), and therefore there is an explicit SoC.

- (29) *Ta shuohua shi bing bu kan wo, **suoyi** ta bing fei tong wo shuohua.*
 'He did not look at me when he spoke, **so** he was not speaking to me.'
- (30) *Ta shuohua shi bing bu kan wo, **suoyi** wo renwei ta bing fei tong wo shuohua.*
 'He did not look at me when he spoke, **so** I think he was not speaking to me.'
- (31) *Ta shuohua shi bing bu kan wo, **suoyi** Xiaoli renwei ta bing fei tong wo shuohua.*
 'He did not look at me when he spoke, **so** Xiaoli thinks he was not speaking to me.'

We are aware of the fact that in Mandarin sometimes an SoC is implicit due to the pervasive phenomenon of subject drop (see (32)). This is not typically possible in European languages such as Dutch, French, and German. As a pro-drop language, Mandarin allows subjects to be omitted when they are in some sense recoverable. In (32), the SoC is *he*, who intentionally stayed to take care of his father. However, the SoC is omitted via subject drop in the consequent. We consider such cases as containing an implicit SoC.

- (32) *Ta danxin fuqin de shenti, **yin'er** Ø liu xia lai zhaogu fuqin.*
 'He worried about his father's health, **therefore** (he) stayed to look after his father.'

Since the SoC is the ultimate source of the subjective construction of the causal relation, a relation with an SoC is more subjective than one with no SoC. In addition, following Langacker's (1990) suggestion that an explicit reference to the SoC objectifies the SoC by making himself/herself observable, fragments with an implicit SoC are taken to be more subjective than fragments with an explicit SoC.

The identity of SoC

In utterances where SoCs are present, a distinction is often made between speaker SoCs, and character SoCs. There is a speaker SoC in (33), and a

character SoC in (34).

- (33) *Nongmin zui danxin de shi zhengce you bian, **suoyi** dang zai nongcun de ge xiang jiben zhengce de wendingxing he lianxuxing zhiguanzhongyao.*

‘What farmers worry about the most is change of policy, **so** it is critical for the Party to maintain stability and continuity of policies in the rural areas.’

- (34) *Suiran zhongguo muqian hai bu shi qianguo, danshi weilai zhongguo you gengduo de jihui, gengda de fazhan, **suoyi** tamen yuanyi liu zai zhongguo.*

‘Although China is not a great power yet, it will have more opportunities and bigger development in future. **So** they (Chinese people) are willing to stay in China.’

In (33) it is the speaker/author who draws the conclusion that it is critical to maintain the stability and continuity of policies in the rural areas, whereas in (34) it is the character *they* who volitionally wants to stay in China. Speaker SoC is considered more subjective than character SoC. After all, the linguistic term of subjectivity refers to the property of being speaker-related.

However, in natural discourse we may encounter sentences in which the author is quoting a causal relation constructed by another person. For example, the causally related sentence within the quotation marks in (35) is constructed by a third person *he* rather than the author. Nevertheless, *he* functions as a speaker, who is responsible for the content of the utterance – or the subjectively constructed causal relation.

- (35) *Ta shuo: “tonghuo-pengzhang renguan cunzai, **yinci** wuyue hui jia xi, shenzhi liuyue hai hui jia xi.”*

‘He said: “Inflation still exists, **so** the interest-rate will be raised in May and even in June.”’

To deal with such cases, we have made a distinction within speaker SoCs: the author versus the current/quoted speaker. The author generally agrees with the quoted speaker while quoting his words, but the author is not directly responsible for the causal relation. Thus, we consider current speaker SoCs to be less subjective than author SoCs.

3.4 Results

Having spelled out the definitions of the key coding categories, we now analyze the use of each of the five connectives in terms of the four subjectivity factors: domain, propositional attitude of the consequent, and

presence and identity of the SoC. For each subjectivity factor, we have conducted General Log-linear Analysis to find the model with the lowest amount of parameters (such as connective and genre) and the best fit to the observed data. Statistical details about the models for each subjectivity factor can be found in Appendices 1 to 4.

3.4.1 Domain

Figure 1 shows the distribution of the five connectives over domains. Clearly, connective use in the speech-act domain is rare (only eleven cases in total). To avoid the loss of test power, we have collapsed the speech-act and the epistemic domain, the two most subjective categories.⁷ And for this reason, we will not discuss in this paper the behavior of the five Mandarin result connectives in the speech-act domain.

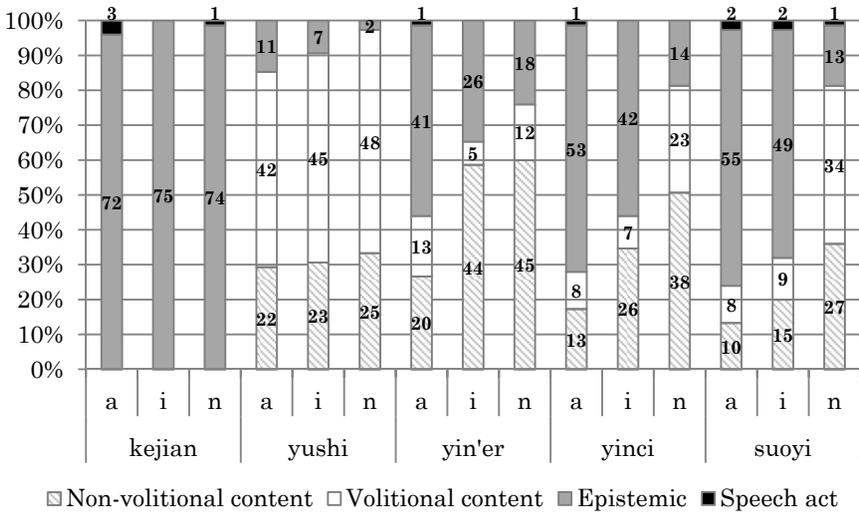


Figure 1. The distribution of connectives over domains per genre (a = argumentative opinion pieces; i = informative news reports; n = narrative novels)

As can be seen in Figure 1, the distribution over domains varies across connective tokens ($\chi^2 [8] = 502.57, p < 0.05$). Moreover, the distribution of connectives over domains differs across genres as well ($\chi^2 [24] = 66.37, p < 0.05$). Below we will translate these findings into domain profiles for each connective.

The data show that language users have a strong preference to use *kejian* in the epistemic domain across genres ($z = 4.16, p < 0.001$). Fragment (36) is a typical example.

- (36) *Cong zhe xiang weiqi shi nian de yanjiu laikan, zhengmian qingxu youzhu yufang guanxinbing, **kejian** baohu xinzang xuyao zengjia zhengmian qingxu.*

‘According to the ten-year long study, positive moods help to prevent the occurrence of coronary heart disease, **so** one needs to increase his positive moods to protect the heart.’

Kejian is typically used to connect clauses such as those in (36) to express an epistemic relation that is subjectively constructed by the speaker: the speaker is drawing a conclusion that one needs to increase his/her positive moods to protect the heart, on the basis of the research finding that positive moods help to guard against coronary diseases.

The profile of *yin'er* is rather different from that of *kejian*. *Yin'er* exhibits a preference for the non-volitional content domain: it occurs less often in the volitional content domain ($z = -5.31, p < 0.001$) and at the same it adheres to the overall main effect of having fewer occurrences in the epistemic domain ($z = -2.72, p = 0.006$). This profile is rather stable across genres. Example (37) serves to illustrate that *yin'er* is typically used to express the non-volitional content domain in which one fact (i.e., the film-shaped solar battery uses only a small amount of silicon) leads to another (i.e., the price is low) in the observed world.

- (37) *Bomoxing taiyangneng dianchi zhi xuyao shiyong shaoliang de yuanliao gui, **yin'er** jiage di.*

‘The film-shaped solar battery uses only a small amount of silicon, **as a result**, its price is low.’

Yinci and *suoyi* seem to resemble each other. They exhibit a general preference for the epistemic domain (for *yinci*: $z = 3.32, p = 0.001$; for *suoyi*: $z = 4.60, p < 0.001$), and are less often used in the volitional content domain (for *yinci*: $z = -3.96, p < 0.001$; for *suoyi*: $z = -2.38, p = 0.02$). However, unlike the other three connectives, the profile of *yinci* and *suoyi* is not fully stable across genres: *yinci* and *suoyi* show an increase of their volitional use in novels (for *yinci*: $z = 2.31, p = 0.02$; for *suoyi*: $z = 2.92, p = 0.004$). The fragments (38) and (39) illustrate the typical causal domain that *yinci* and *suoyi* generally prefer to express.

- (38) *Lishi shi wanzheng de, **yinci** wenhua ye yinggai shi wanzheng de.*

‘History is continuative, **so** culture should also be continuous.’

- (39) *Ta shou shang daile hao ji ge jiezh, **suoyi** wo yiwei ta jiehun le.*

‘He wore quite a few rings on the hand, **so** I thought he was married.’

Both (38) and (39) involve an SoC's epistemic reasoning: the speaker makes a judgment (in (38): *culture should also be continuous*; in (39): *I thought he was married*) on the basis of the propositional fact expressed in the antecedent (in (38): *history is continuative*; in (39): *he wore quite a few rings on the hand*).

The final connective, *yushi*, adheres to the overall main effects of domains: it has a clear preference for use in the volitional content domain ($z = 2.60, p = 0.009$), and hardly occurs in the epistemic domain ($z = -2.72, p = 0.006$). Just like *kejian* and *yin'er*, *yushi* shows a stable pattern across genres. The typical use of *yushi* is exemplified in (40): the fact that he heard some knocking sounds on the seabed causes his intentional act of swimming there to see what happened.

- (40) *Ta tingdaole haichuang shang qiaoda de shengyin, yushi ta you guoqu xiang kan ge jiuqing.*
 'He heard some sounds produced by knocks on the seabed, **so** he swam over to see what happened.'

3.4.2 Propositional attitude

Figure 2 shows the distribution of the five connectives over the different propositional attitudes. Similar to what we have found for the speech-act domain, the speech-act propositional attitude is rarely observed (only eleven cases in total). This is why we collapse the speech-act propositional attitude with another highly subjective category: judgment.

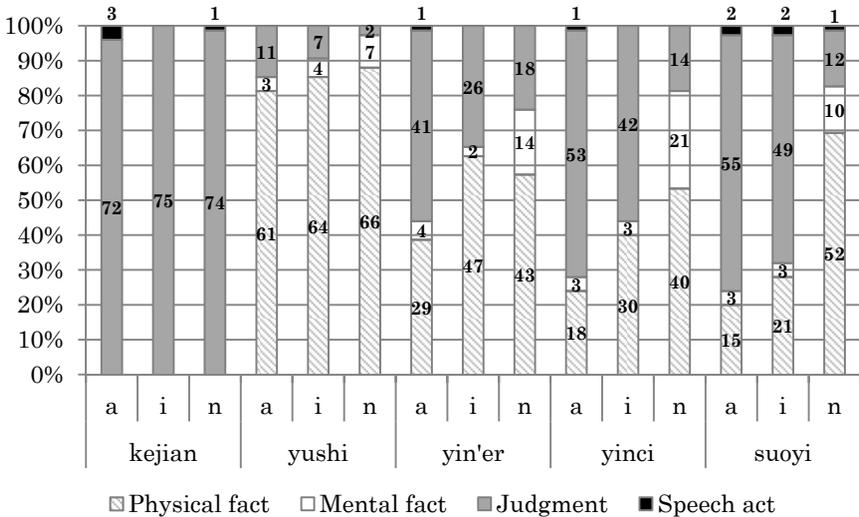


Figure 2. The distribution of connectives over propositional attitude per genre (a = argumentative opinion pieces; i = informative news reports; n = narrative novels)

As can be seen in Figure 2, these connectives present different patterns of distribution over propositional attitude (χ^2 [8] = 429.61, $p < 0.001$). Furthermore, the distribution patterns vary by genre (χ^2 [24] = 61.09, $p < 0.001$). Below we will translate these findings into profiles of propositional attitude for each connective.

The profile of *kejian* in terms of propositional attitude appears to be salient and stable. Across genres *kejian* has a great chance to co-occur with judgments ($z = 2.87$, $p = 0.004$). While *yushi*, *yinci*, and *suoyi* show a decrease of judgments in novels (cf. the interaction effect of genre and modality: $z = -5.60$, $p < 0.001$), *kejian* maintains its strong preference for judgments even in novels ($z = 3.87$, $p < 0.001$). Fragment (36), given in the previous section, also exemplifies the typical usage of *kejian* in this respect. The consequent clause *one needs to increase his positive moods to protect the heart* is clearly opinion-oriented, so it is considered a judgment.

Yinci and *suoyi* are very similar to each other with respect to propositional attitude. In news reports and opinion pieces, they both adhere to the main effect: they prefer to co-occur with judgments ($z = 3.40$, $p = 0.001$), and appear less often with mental facts ($z = -3.15$, $p = 0.002$). In novels, however, *yinci* and *suoyi* show an increase of physical facts and a decrease of judgments (cf. the interaction between genre and propositional attitude: $z = -5.60$, $p < 0.001$).⁸ To illustrate the typical propositional attitude that *yinci* and *suoyi* co-occur with, we may look at fragments (38) and (39) again. The consequent clause in (38) *culture should also be continuous* is conclusive in nature, and the consequent in (39) *I thought he was married* is a personal opinion. These are both typical examples of judgments, the type of propositional attitude *yinci* and *suoyi* prefer to co-occur with.

Yin'er and *yushi* are quite similar to one another in terms of propositional attitude. They exhibit a preference for physical facts: they appear less often with judgments (for *yin'er*: $z = -4.13$, $p < 0.001$; for *yushi*: $z = -6.53$, $p < 0.001$) and mental facts (i.e., they adhere to the main effect of propositional attitude: $z = -3.15$, $p = 0.002$). Their preference for physical facts is rather constant across genres (see Endnote 8). But *yushi* is also sensitive to the genre: it has less chance to co-occur with judgments in novels than in the other two genres (i.e., it adheres to the interaction between genre and propositional attitude: $z = -5.60$, $p < 0.001$). The two connectives' preference for physical facts is illustrated in fragments (37) and (40). The consequent clause in (37) *the price is low* is an observable fact, rather than an SoC's conclusion. Similarly, the consequent in (40) *he swam over to see what happened* is a physical act that can be observed in the outside world.

3.4.3 The presence of SoC

Figure 3 shows the distribution of the five connectives over the kinds of SoC they can co-occur with. The connectives show different distribution profiles ($\chi^2 [8] = 383.60, p < 0.001$). Moreover, the distribution of connectives over the kinds of SoC varies by genre ($\chi^2 [24] = 44.73, p = 0.006$). Below we will interpret these findings in detail.

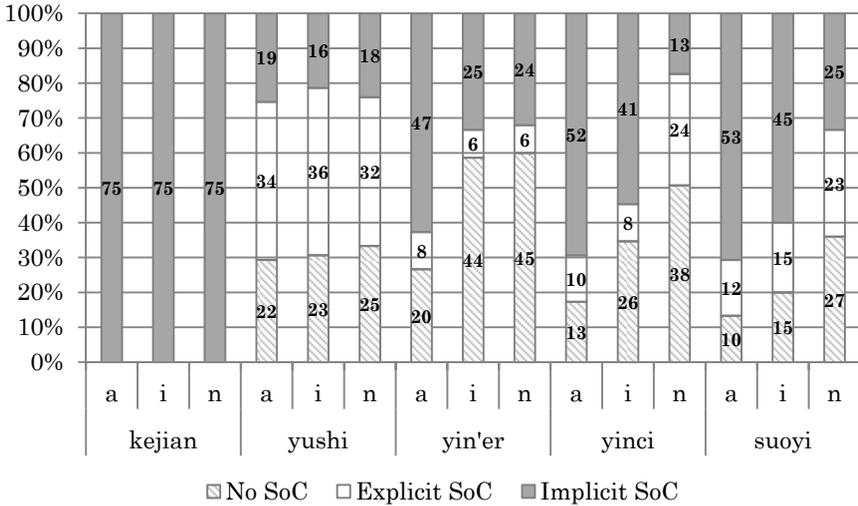


Figure 3. The distribution of connectives over presence of SoC per genre (a = argumentative opinion pieces; i = informative news reports; n = narrative novels)

Kejian shows a strong preference for implicit SoCs ($z = 3.69, p < 0.001$). This profile appears to be stable across genres. The preference for an implicit SoC is illustrated clearly in fragment (36), where a conclusion is drawn by an SoC, *one needs to increase his positive moods to protect the heart*, and meanwhile the conclusion maker is not explicitly referred to by any linguistic element. *Suoyi* also shows a preference for implicit SoCs across genres ($z = 3.29, p = 0.001$) (see (33)).

As to *yinci*, in news reports and opinion pieces it exhibits a preference for implicit SoCs ($z = 1.98, p = 0.048$) and a dislike of explicit SoCs ($z = -3.32, p = 0.001$), whereas in novels it shows an increase of explicit SoCs ($z = 2.52, p = 0.01$) and a decrease of implicit SoCs ($z = -2.68, p = 0.007$). Fragment (38) exemplifies the kind of SoC *yinci* generally co-occurs with. The person who makes the judgment *culture should also be continuous* is inferable from the context, but it is implicit in the consequent.

The connective *yin'er* is often used in causal relations with no SoC: speakers do not prefer to use *yin'er* together with explicit SoCs ($z = -4.76, p <$

0.001), and this lower number of explicit SoCs does not result in a higher number in the category of implicit SoC ($z = -0.50$, $p = 0.62$).⁹ The profile of *yin'er* is rather constant across genres. Fragment (37) shows the typical usage of *yin'er*: the causal connection involved does not depend on intentionality. The result segment expresses the natural consequence of the economic law: when the cost of production is low, the price for the product is low.

Compared to the connectives *yin'er*, *yinci* and *kejian*, the connective *yushi* co-occurs more often with explicit SoCs (*yin'er*: fewer explicit SoCs, $z = -4.76$, $p < 0.001$; *yinci*: fewer explicit SoCs, $z = -3.32$, $p = 0.001$; *kejian*: no instances of explicit SoCs were observed). Fragment (40) exemplifies this prototypical use of *yushi*, in which the source of the intentional act of *swimming there* is explicitly realized by the pronoun *he*.

In the analysis, we have observed 79 instances of implicit SoCs in volitional content relations. As introduced earlier (see (32) in Section 3.3.3), this type of implicit SoC occurs as a result of subject drop. It occupies the subject position of the result segment and functions as the agent of an intentional act. It is different from the type of implicit SoC commonly found in epistemic relations, which is outside the syntactic structure of the sentence. On the basis of their semantic roles, we name the former type of implicit SoC ‘implicit agent SoC’, and the latter type ‘implicit concluder SoC’. Table 5 shows the distribution of the connectives over the two types of implicit SoC.

Table 5. The distribution of the connectives over two types of implicit SoC

Connective	Type of implicit SoC		Total
	Implicit agent SoC	Implicit concluder SoC	
<i>kejian</i>	0 (.0%)	225 (100.0%)	225 (100.0%)
<i>yushi</i>	41 (77.4%)	12 (22.6%)	53 (100.0%)
<i>yinci</i>	3 (2.8%)	103 (97.2%)	106 (100.0%)
<i>yin'er</i>	18 (18.8%)	78 (81.3%)	96 (100.0%)
<i>suoyi</i>	17 (13.8%)	106 (86.2%)	123 (100.0%)
Total	79 (13.1%)	524 (86.9%)	603 (100.0%)

In contrast to the other four connectives, *yushi* co-occurs more often with the implicit agent SoC resulting from subject drop. An example is given in (41) to illustrate the special usage of *yushi*. We will discuss the two types of implicit SoC in depth in Section 3.5.

(41) *Ta ganjuedao nanbian fuza, yushi Ø laidao Beiping.*

‘He felt that the situation in the South was complicated, so (he) came to Beijing (the old name of Beijing).’

3.4.4 The identity of the SoC

Figure 4 shows the distribution of the five connectives over the types of SoC they can co-occur with. The connectives' distributional patterns differ ($\chi^2 [8] = 374.80, p < 0.001$), and vary with genre ($\chi^2 [24] = 82.36, p < 0.001$). Next, we will interpret these findings in more detail.

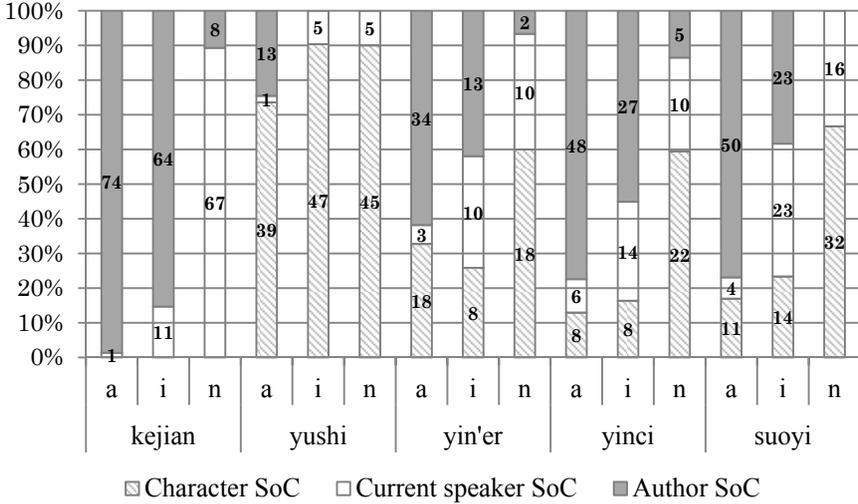


Figure 4. The distribution of connectives over identity of SoC per genre (a = argumentative opinion pieces; i = informative news reports; n = narrative novels)

Kejian exhibits an overall preference for author SoCs ($z = 2.50, p = 0.01$), except that in novels there is an increase of current speaker SoCs ($z = 4.06, p < 0.001$). Fragment (36) illustrates the typical use of *kejian* with an author SoC: in (36) the author is the ultimate source of the judgment *one needs to increase his positive moods to protect the heart*. The typical use of *kejian* in novels is shown in the following sentence (42): it is used to express causal relations constructed by a first person character in the novel rather than the author. It is not the writer of the novel who draws the conclusion about the picture, but the first person character “I”, which we code as a case of current speaker SoC.

- (42) *Wo zhidao Zhang Zeduan buhui fan zhe zhong cuowu, kejian zhe bushi Zhang Zeduan suo hua.*
 ‘I know that Zhang Zeduan won’t make this type of error, **so** the picture is not painted by Zhang Zeduan.’

In contrast with *kejian*, the connective *yushi* prefers character SoCs: it co-occurs less often with current speaker ($z = -4.31, p < 0.001$) and author SoCs ($z = -3.88, p < 0.001$). Its preference for character SoCs is quite stable across genres. In general, there is an increase of character SoCs in novels (cf. the interaction between genre and identity of SoC: fewer cases of current speaker SoC, $z = -2.27, p = 0.02$; fewer cases of author SoC, $z = -4.19, p < 0.001$), but *yushi* does not show such an increase of character SoCs ($z = -2.24, p = 0.03$). The preference for character SoCs is illustrated in fragments (40) and (41). Although the character SoC *he* is implicit in (41), it can be recovered from the context: it is the character *he* who did the intentional act of coming to the city of Beijing.

The other three connectives (i.e., *suoyi*, *yinci*, and *yin'er*) also appear more often with author SoCs (cf. the main effect of the identity of SoC: $z = 2.99, p = 0.003$), except in novels, where the three connectives show an increase of character SoCs (cf. the afore-mentioned interaction between genre and identity of SoC). Fragments (38) and (39) have exemplified the typical use of *yinci* and *suoyi*, respectively. Furthermore, we have already seen that there is often no SoC in the *yin'er* fragments. However, *yin'er* can also express causal relations that are subjectively constructed by an SoC. In such cases, the SoC is most often the author, as shown in the following fragment (43). Here it is the author who draws the conclusion that *changes must be made*.

(43) *Huangjinzhou de fumian houguo ye riye xianxian chulai, yin'er huangjinzhou de gaibian ye jiu shizai-bixing.*

'The negative effects of Golden Week holidays have revealed themselves gradually, **therefore** changes must be made.'

3.5 Discussion

The current study has produced several interesting findings and topics for discussion. Below we will reflect on the subjectivity profiles of the connectives (Section 3.5.1), as well as their genre-sensitivity (3.5.2). Furthermore, we will discuss the distributional patterns and frequencies of use in terms of speaker economy (3.5.3), and make some remarks on the analytical model we used in this paper (3.5.4).

3.5.1 Subjectivity profiles identified

Table 6 summarizes the subjectivity profiles of the five connectives in question. Three robust subjectivity profiles across genres can be identified for our data: epistemic *kejian*, volitional *yushi*, and non-volitional *yin'er*.

Table 6. Overview of the subjectivity profiles per connective

Connective	Domain	Modality	Presence of SoC	Identity of SoC
<i>kejian</i>	+epistemic	+judgment	+implicit	+author * (current speaker)
<i>yin'er</i>	+non-vol. content	+physical fact	+no SoC	+author * (character)
<i>yushi</i>	+vol. content	+physical fact	+explicit	+character
<i>yinci</i>	+epistemic * (vol. content)	+judgment * (physical fact)	+implicit * (explicit)	+author * (character)
<i>suoyi</i>	+epistemic * (vol. content)	+judgment * (physical fact)	+implicit	+author * (character)

Note: * indicates that effects are genre-sensitive; within brackets we mention the category that increases in novels.

Kejian, a highly subjective causal connective, specializes in expressing epistemic relations and prefers to occur in highly subjective contexts involving judgments and implicit speaker SoCs. By contrast, *yin'er* has a rather objective profile. It prefers to occur within the non-volitional content domain to describe causal connections between physical facts in observable reality. *Yushi* also exhibits a clear preference for the objective content domain and for the objective contexts involving physical facts. It should be noted that *yushi* can occasionally express epistemic relations, which counters Lu's (2000) claim that *yushi* does not introduce inferences and conclusions at all. Unlike *yin'er*, *yushi* prefers to express content relations in which human intentions play a role: it is often used to express volitional content relations in which an SoC can be identified as the source of the causal process. Moreover, the SoC is often a character explicitly realized in segments containing *yushi*. Thus, *yushi* may be considered to be more subjective than *yin'er*, and less subjective than *kejian*.

These results are basically in line with observations in previous publications and our hypotheses formulated in Section 3.3.1. In particular, the objective profile of *yushi* correlates with its closeness to a descriptive temporal use; the subjective profile of *kejian* conforms to the epistemic nature of its root meaning, 'can see'; and *yin'er* expresses objective causal relations – descriptive causality – precisely as Xing (2001) predicts.

Generally speaking, *yinci* and *suoyi* are more subjective than *yin'er* and *yushi*. Speakers prefer to use *yinci* and *suoyi* in the epistemic domain, linking arguments with conclusions, where the conclusion maker (i.e., the SoC) is often the author and is often not linguistically realized (i.e., implicit).

Yinci and *suoyi* are also used in the objective content domain; it is on this basis we conclude that they are less subjective than *kejian*, which specializes in expressing epistemic relations and never occurs in the objective content domain.

Except for *yushi*, each of the connectives under discussion has one or two instances of speech-act relations. The number of observed speech-act relations is rather low in our samples: only 11 cases in total. This is because the present study focuses on the written discourse. Previous studies have already shown that speech-act relations rarely occur in the written discourse. For example, Sanders & Spooren (submitted) find that speech-act relations occur relatively more often in chat (27.8%) and spoken language (10.1%) than in written language (5.0%). On the basis of the present study, we cannot draw a conclusion on the behavior of the five Mandarin result connectives in the speech-act domain. Spoken discourse should be investigated for that purpose.

3.5.2 Genre-sensitivity: semantics vs. pragmatics

Given the fact that *kejian* and *yin'er* show stable usage patterns across genres in terms of domain, modality, and presence of SoC, why does it seem that their usage varies with genre when the fourth variable, identity of SoC, is concerned? As shown in Table 6, in novels there is an increase of current speaker SoCs in the *kejian* fragments and an increase of character SoCs in the *yin'er* fragments. Do these variations affect our conclusion about the subjectivity profiles of *kejian* and *yin'er*? The answer is negative. In fact, in spite of the observed variations in novels, *kejian* and *yin'er* can still be argued to have stable subjectivity profiles across genres. For *kejian*, the variation from author SoC to current speaker SoC does not affect the nature of its subjectivity profile because both of these categories are highly subjective. For *yin'er*, the increase in character SoCs does not affect our conclusion that *yin'er* has a rather objective profile across genres, because the fragments with an SoC make up only 52% of the *yin'er* cases. This means that almost half of the fragments is of the most objective type: without any SoCs.

While *kejian*, *yushi*, and *yin'er* have robust profiles across genres, the subjectivity profile of *yinci* and *suoyi* are genre-sensitive. In news reports and opinion pieces, they are most often used in the epistemic domain, but exhibit less subjective patterns in novels: in novels they co-occur more often with physical facts, volitional content relations, and character SoCs. The observed preference of *suoyi* and *yinci* for the epistemic domain seems to diverge from the findings in Xing (2001), who finds that these connectives typically mark descriptive causality (i.e., in the content domain). This apparent disagreement disappears when we take the genre-sensitivity of *suoyi* and

yinci into account: in novels, the only genre Xing (2001) focused on, these connectives are indeed more often used to express objective relations in the content domains.

The observed genre-sensitivity conforms to previous findings in the literature that the distribution of connectives over causal categories can vary across contexts. For example, the subjectively oriented French *car*, German *denn*, and Dutch *want* have a strong preference for subjective causal relations across text types, whereas the profiles of objectively oriented connectives such as French *parce que*, German *weil*, and Dutch *omdat* are inconsistent in newspaper corpora (for an overview, see Stukker & Sanders 2012).

The observed genre-sensitivity makes us wonder whether the subjectivity profiles are really part of the inherent semantic characteristics of the connectives themselves or whether they are (partially) determined by the pragmatics of the context. Neither option is sufficient to account for the entire set of data. Therefore, we propose a third option. If we assume that subjectivity is a semantic feature expressed by all causal connectives, we may propose that for *kejian*, *yushi*, and *jin'er*, the feature of subjectivity is specified in the lexicon, resulting in a subjectivity profile that is constant across genres. For *yinci* and *suoyi*, however, the subjectivity feature is semantically underspecified. Consequently, they depend on the pragmatics of the context for the exact value of subjectivity. For example, in novels authors prefer to describe observable situations and relate characters' physical activities to their motivations. Hence in novels *yinci* and *suoyi* show an increase of volitional use.

3.5.3 Speaker economy

The current study reveals that Mandarin has a specific result connective, *kejian*, which is restricted to expressing highly subjective epistemic relations. In comparison with the other four connectives, *kejian* is the least frequently used in the Mandarin written discourse (see Table 1). The other four connectives are quite general across causal domains: they can be used to mark different domains of causality, although they all have their own prototypical usage. Among them, the connectives *yinci* and *suoyi* are the most general ones because they are able to serve not only as markers for subjective epistemic relations (in news reports and opinion pieces) but also as markers for objective volitional content relations (in novels). These two connectives are also the most frequently used causal connectives in the Mandarin written discourse. Taken together, we can see that general causal connectives have a higher frequency of usage than specific ones do. The findings seem to suggest that speakers prefer to choose the most general connectives to express a relation, which is in line with a speaker economy

account (Knott & Sanders 1998).

3.5.4 Remarks on the analytical model

The four variables in the analytical model have provided consistent information regarding the degrees of subjectivity that each result connective expresses (see Table 6). For example, whenever a connective is used in the non-volitional content domain, there is no SoC involved and the consequent always involves facts. This begs the questions as to why four variables instead of one would be needed and whether there is redundancy in positing multiple variables in characterizing subjectivity. We would like to stress that the integrative system in our analytic model is needed, as all four of the variables are needed to quantify degrees of subjectivity from different angles. Therefore, it is expected that the four variables can serve to complement each other while depicting a connective's degree of subjectivity. For example, in case the domain factor fails to distinguish one connective from another, other variables may provide complementary information to set them apart. In principle, we are able to distinguish less objective from more objective non-volitional content causal relations with the variable propositional attitude (physical fact versus mental fact). The model also makes it possible to distinguish less subjective from more subjective epistemic relations with variables such as the presence and identity of SoC.

As reported in Section 3.4.3, two types of implicit SoC have been observed in Mandarin Chinese. This is a notable difference between Mandarin Chinese, a pro-drop language, and the Germanic languages, on whose basis studies on the subjectivity of connectives have been conducted. In Germanic languages such as Dutch and German in which subject drop is impossible, implicit SoCs are rarely found in the volitional content causal domain: the agent SoCs (in the volitional content domain) are always linguistically realized, except in passive structures such as *I/He did a good job, that's why I/he was promoted*, in which the agent of promotion is avoided. That is to say, in Germanic languages, implicit SoCs always occur in the epistemic domain, functioning as the source of a conclusion (see also Sanders, Sanders & Sweetser 2009, 2012). For this reason, previous studies based on non-pro-drop languages have not made a distinction between the implicit agent SoC (resulting from subject drop) and the implicit concluder SoC. The two types of implicit SoC can be illustrated with fragment (41) (repeated here as (44)) and fragment (43) (repeated here as (45)), respectively.

(44) *Ta ganjuedao nanbian fuza, yushi Ø laidao Beiping.*

'He felt that the situation in the South was complicated, **so** (he) came to Beijing (the old name of Beijing).'

- (45) *Huangjinzhou de fujian houguo ye riye xianxian chulai, yin'er huangjinzhou de gaibian ye jiu shizai-bixing.*

'The negative effects of Golden Week holidays have revealed themselves gradually, **therefore** changes must be made.'

We think the two types of SoC can be considered as expressing different degrees of subjectivity. For example, the agent SoC in (44) *he* can be recovered at the syntactic level (e.g., by feature-checking), that is, *he* is actually present in the deep structure of the sentence. By comparison, the concluder SoC in (45), the speaker, cannot be recovered through syntactic mechanisms. It is more implicit in the sense that it is completely outside the structure of the sentence. Therefore, in order to get a clear picture of the subjectivity profiles of connectives in subject drop languages such as Mandarin, we should distinguish the two types of implicit SoC. A connective preferring implicit agent SoCs should be considered less subjective than a connective preferring implicit concluder SoCs.

3.6 Conclusion

In this paper, we have performed a corpus-based analysis on five Mandarin result connectives (i.e., *kejian*, *suoyi*, *yinci*, *yin'er*, and *yushi*) in terms of four subjectivity indicators: propositional attitude, domain, and the presence and identity of an SoC. This analysis has uncovered distinct subjectivity profiles for result connectives. Three result connectives display robust profiles in terms of subjectivity across genres, which can be arranged in a hierarchical order from more to less subjective: *kejian* > *yushi* > *yin'er*. Genre-sensitivity has been observed for the connectives *yinci* and *suoyi*. Therefore, we propose that for connectives with a robust subjectivity profile (*kejian*, *yushi*, and *yin'er*), the feature of subjectivity is specified in the lexicon, whereas for genre-sensitive connectives (*yinci* and *suoyi*), the subjectivity feature is semantically underspecified, leaving room for the pragmatics of the context to determine their exact value of subjectivity.

Furthermore, we have confirmed existing claims (e.g., the ones in Xing 2001) about the ways of result marking in Mandarin discourse with evidence from a systematic corpus-based study. Differences between Xing's and our findings could be ascribed to the genres that were studied: Xing only studied novels, while we included news reports and opinion pieces as well. The ostensible inconsistency between studies urges us to take the effect of genre into account as we discuss the meaning and usage of causal connectives and other discourse phenomena.

The present study has shown the relevance of the notion of subjectivity to the description of causal connectives in Mandarin Chinese – a language that is typologically very different from the Germanic and Romance

languages that were so far the focus of studies on subjectivity and connectives. The results on Mandarin connectives suggest that the theory of subjectivity can be generalized across various languages. At the same time, we have seen differences that are – at least to a certain extent – related to the systematic differences between the languages. For instance, two types of implicit SoC (i.e., implicit agent SoC and implicit concluder SoC) come into our view in Mandarin Chinese, a pro-drop language, of which the first one is lacking in Germanic languages that have been studied from a subjectivity perspective so far. It is likely that implicit SoCs also occur in Romance languages which allow pro-drop as well, such as Spanish, Italian, and Portuguese (Haegeman 1994). Further research needs to reveal the ways in which the phenomenon of implicit agent SoC occurs in Romance languages and other pro-drop languages. With the present study, we hope to have shown how a theory-driven and corpus-based approach to the categorization of causal connectives produces interesting results for further cross-linguistic comparison.

Notes

1. An almost identical version of this chapter appeared as Li, F., J. Evers-Vermeul & T. J. M. Sanders. 2013. Subjectivity and result marking in Mandarin: A corpus-based investigation. *Chinese Language and Discourse* 4(1): 74-119. It is included in this dissertation with kind permission from John Benjamins Publishing Company, Amsterdam / Philadelphia. [www.Benjamins.com]
2. According to Chao (1968), *yin'er*, *yinci*, and *suoyi* are adverbial conjunctions because they can be put either in front of or immediately after the subject (which is the position for adverbs). However, many other authors disagree with this definition. For example, Lü (1979) and Zhang (1996) suggest that a lexeme is a conjunction if it can appear in both the pre- and post-subject position, and an adverb if it can only appear in the post-subject position.
3. But see Keller (1995) for recent developments in the German connective system.
4. In (14), there is an animate participant, Bob. However, Bob's intentionality is not the ultimate source of the causal event of getting hurt in the leg as Bob did not intentionally get himself hurt in the leg. The causal relation in (14) exists independent of Bob's intentionality. Neither does it depend on the speaker's intentionality. That is, there is no SoC at all in (14).
5. *Guanxici* or *lianjiexi* are defined as words or expressions that have the function of connecting clauses and simultaneously marking the coherence relation between the connected clauses (Xing 1996:321; Xing 2001:26).
6. The LCMC Corpus (McEnery & Xiao 2004) is a balanced corpus representing modern Mandarin Chinese from Mainland China. It consists of five hundred 2,000-

word samples taken from a wide range of text-types, including news reports, novels, and opinion pieces.

7. The Log-linear analysis requires that there should be no expected counts less than 1 in the contingency table (Field 2009:710). However, in our study there are ten cells with expected counts of 0.4 due to the rare occurrence of the speech-act category. This is why we decide to collapse speech-act with another highly subjective category.

8. As clearly shown in Figure 2, *suoyi* has (at least) twice as many physical facts in novels than in news reports and opinion pieces. *Yinci* does not significantly differ from *suoyi* in this respect. The statistics in Appendix 2B only show this effect indirectly, because we take *suoyi* as the reference category. *Suoyi*'s (and *yinci*'s) relatively high frequency in fragments with physical facts in novels can be derived from the significantly lower frequency of this category in the use of *yin'er* and *yushi* in novels (for *yin'er*: $z = -2.96$, $p = 0.003$; for *yushi*: $z = -2.78$, $p = 0.005$).

9. Statistically speaking, *yin'er*'s preference for "no SoC" can only be shown indirectly, because "no SoC" serves as the reference category in our statistical analyses.

Appendix 1. The analysis of *domain*

A. Results of the General Log-linear analysis of *domain*

Model	χ^2 (model)	<i>df</i>	<i>p</i> (model)	χ^2 (factor)	<i>df</i>	<i>p</i> (factor)
1. constant+conn.	813.04	40	<.001			
+ 2. genre	813.04	38	<.001	0	2	> .05
+ 3. domain	643.37	36	<.001	169.67	2	< .001
+ 4. domain*conn.	140.80	28	<.001	502.57	8	< .001
+ 5. domain*genre	66.37	24	<.001	74.43	4	< .001
+ 6. domain*conn.*genre	.00	.00	1.000	66.37	24	< .001

B. Parameter estimates *domain* for Model 6

Parameter	Estimate	Std. Error	<i>z</i>	<i>p</i>
constant	3.157	.206	15.304	< .001
[conn. = kejian]	-3.850	1.429	-2.694	.007
[conn. = suoyi]	-.416	.327	-1.272	.203
[conn. = yinci]	.120	.283	.424	.672
[conn. = yin'er]	.638	.255	2.504	.012
[genre = argu.]	-.043	.295	-.147	.883
[genre = narra.]	.082	.286	.286	.775
[domain = volitional content]	.661	.254	2.601	.009
[domain = epistemic/speech-act]	-1.142	.419	-2.723	.006
[conn. = kejian] * [domain = vol. content]	-.661	2.016	-.328	.743
[conn. = kejian] * [domain = epistemic]	6.159	1.480	4.163	< .001
[conn. = suoyi] * [domain = vol. content]	-1.150	.484	-2.376	.017
[conn. = suoyi] * [domain = epistemic]	2.343	.510	4.596	< .001
[conn. = yinci] * [domain = vol. content]	-1.923	.485	-3.962	< .001
[conn. = yinci] * [domain = epistemic]	1.614	.487	3.315	.001
[conn. = yin'er] * [domain = vol. content]	-2.751	.518	-5.307	< .001
[conn. = yin'er] * [domain = epistemic]	.624	.486	1.284	.199
[genre = argu.] * [domain = vol. content]	-.025	.364	-.068	.946
[genre = argu.] * [domain = epistemic]	.471	.554	.850	.396
[genre = narra.] * [domain = vol. content]	-.018	.353	-.051	.960
[genre = narra.] * [domain = epistemic]	-1.180	.784	-1.505	.132
[conn. = kejian] * [genre = argu.] * [domain = vol.]	.068	2.011	.034	.973
[conn. = kejian] * [genre = argu.] * [domain = epistemic]	-.427	.497	-.860	.390
[conn. = kejian] * [genre = argu.] * [domain = non-vol.]	.043	2.022	.022	.983
[conn. = kejian] * [genre = narra.] * [domain = vol.]	-.064	2.011	-.032	.975
[conn. = kejian] * [genre = narra.] * [domain = epistemic]	1.099	.748	1.468	.142
[conn. = kejian] * [genre = narra.] * [domain = non-vol.]	-.082	2.020	-.040	.968
[conn. = suoyi] * [genre = argu.] * [domain = vol.]	-.043	.518	-.083	.934

[conn. = suoyi] * [genre = argu.] * [domain = epistemic]	-.317	.507	-.626	.532
[conn. = suoyi] * [genre = argu.] * [domain = non-vol.]	-.346	.497	-.696	.486
[conn. = suoyi] * [genre = narra.] * [domain = vol.]	1.226	.421	2.915	.004
[conn. = suoyi] * [genre = narra.] * [domain = epistemic]	-.169	.788	-.214	.830
[conn. = suoyi] * [genre = narra.] * [domain = non-vol.]	.492	.427	1.150	.250
[conn. = yinci] * [genre = argu.] * [domain = vol.]	.193	.545	.355	.722
[conn. = yinci] * [genre = argu.] * [domain = epistemic]	-.179	.512	-.349	.727
[conn. = yinci] * [genre = argu.] * [domain = non-vol.]	-.631	.446	-1.415	.157
[conn. = yinci] * [genre = narra.] * [domain = vol.]	1.078	.467	2.307	.021
[conn. = yinci] * [genre = narra.] * [domain = epistemic]	.023	.791	.029	.977
[conn. = yinci] * [genre = narra.] * [domain = non-vol.]	.292	.381	.765	.444
[conn. = yin'er] * [genre = argu.] * [domain = vol.]	.966	.549	1.760	.078
[conn. = yin'er] * [genre = argu.] * [domain = epistemic]	.045	.531	.085	.933
[conn. = yin'er] * [genre = argu.] * [domain = non-vol.]	-.732	.398	-1.839	.066
[conn. = yin'er] * [genre = narra.] * [domain = vol.]	.757	.552	1.372	.170
[conn. = yin'er] * [genre = narra.] * [domain = epistemic]	.739	.791	.935	.350
[conn. = yin'er] * [genre = narra.] * [domain = non-vol.]	-.059	.355	-.167	.867

Appendix 2. The analysis of *propositional attitude*

A. Results of the General Log-linear analysis of *propositional attitude* (of the consequent, Q)

Model	χ^2 (model)	df	p (model)	χ^2 (factor)	df	p (factor)
1. constant+conn.	1068.64	40	< .001			
+ 2. genre	1068.64	38	< .001	0	2	> .05
+ 3. Q	581.59	36	< .001	487.05	2	< 0.001
+ 4. Q*conn.	151.98	28	< .001	429.61	8	< 0.001
+ 5. Q*genre	61.09	24	<.001	90.89	4	< 0.001
+ 6. Q*conn.*genre	.00	.00	1.000	61.09	24	< 0.001

B. Parameter estimates *propositional attitude* for Model 6

Parameter	Estimate	Std. Error	z	p
constant	3.068	.216	14.226	< .001
[conn. = kejian]	-3.761	1.431	-2.629	.009
[conn. = yushi]	1.099	.249	4.412	< .001
[conn. = yinci]	.350	.282	1.242	.214
[conn. = yin'er]	.793	.260	3.050	.002
[genre = argu.]	-.327	.333	-.982	.326
[genre = narra.]	.893	.256	3.487	< .001
[Q = mental fact]	-1.815	.576	-3.149	.002
[Q = judgment]	.874	.257	3.402	.001
[conn. = kejian] * [Q = mental fact]	1.815	2.081	.872	.383
[conn. = kejian] * [Q = judgment]	4.144	1.442	2.874	.004
[conn. = yushi] * [Q = mental fact]	-.847	.755	-1.122	.262
[conn. = yushi] * [Q = judgment]	-3.025	.463	-6.528	< .001
[conn. = yinci] * [Q = mental fact]	-.350	.807	-.433	.665
[conn. = yinci] * [Q = judgment]	-.542	.350	-1.549	.121
[conn. = yin'er] * [Q = mental fact]	-1.129	.868	-1.301	.193
[conn. = yin'er] * [Q = judgment]	-1.457	.353	-4.126	< .001
[genre = argu.] * [Q = mental fact]	.327	.826	.396	.692
[genre = argu.] * [Q = judgment]	.437	.384	1.138	.255
[genre = narra.] * [Q = mental fact]	.206	.668	.308	.758
[genre = narra.] * [Q = judgment]	-2.232	.399	-5.596	< .001
[conn. = kejian] * [genre = argu.] * [Q = mental]	1.816E-16	2.138	.000	1.000
[conn. = kejian] * [genre = argu.] * [Q = judgment]	-.110	.252	-.438	.661
[conn. = kejian] * [genre = argu.] * [Q = physical]	.327	2.028	.161	.872
[conn. = kejian] * [genre = narra.] * [Q = mental]	-1.099	2.093	-.525	.600
[conn. = kejian] * [genre = narra.] * [Q = judgment]	1.339	.346	3.865	< .001
[conn. = kejian] * [genre = narra.] * [Q = physical]	-.893	2.016	-.443	.658

[conn. = yushi] * [genre = argu.] * [Q = mental]	-.251	1.039	-.242	.809
[conn. = yushi] * [genre = argu.] * [Q = judgment]	.317	.507	.626	.532
[conn. = yushi] * [genre = argu.] * [Q = physical]	.280	.378	.740	.459
[conn. = yushi] * [genre = narra.] * [Q = mental]	-.588	.858	-.685	.493
[conn. = yushi] * [genre = narra.] * [Q = judgment]	.240	.792	.303	.762
[conn. = yushi] * [genre = narra.] * [Q = physical]	-.862	.310	-2.781	.005
[conn. = yinci] * [genre = argu.] * [Q = mental]	-4.225E-16	1.069	.000	1.000
[conn. = yinci] * [genre = argu.] * [Q = judgment]	.138	.281	.494	.622
[conn. = yinci] * [genre = argu.] * [Q = physical]	-.173	.445	-.388	.698
[conn. = yinci] * [genre = narra.] * [Q = mental]	.717	.844	.849	.396
[conn. = yinci] * [genre = narra.] * [Q = judgment]	.264	.431	.611	.541
[conn. = yinci] * [genre = narra.] * [Q = physical]	-.609	.351	-1.737	.082
[conn. = yin'er] * [genre = argu.] * [Q = mental]	.588	1.093	.538	.591
[conn. = yin'er] * [genre = argu.] * [Q = judgment]	.362	.313	1.156	.248
[conn. = yin'er] * [genre = argu.] * [Q = physical]	-.149	.407	-.366	.714
[conn. = yin'er] * [genre = narra.] * [Q = mental]	.659	.922	.715	.475
[conn. = yin'er] * [genre = narra.] * [Q = judgment]	.980	.430	2.276	.023
[conn. = yin'er] * [genre = narra.] * [Q = physical]	-.981	.331	-2.962	.003

Appendix 3. The analysis of *presence of SoC*

A. Results of the General Log-linear analysis of *presence of SoC*

Model	χ^2 (model)	<i>df</i>	<i>p</i> (model)	χ^2 (factor)	<i>df</i>	<i>p</i> (factor)
1. constant+conn.	708.24	40	< .001			
+ 2. genre	708.24	38	< .001	0	2	> .05
+ 3. presence	475.05	36	<.001	233.19	2	< 0.001
+ 4. presence*conn.	91.45	28	<.001	383.60	8	< 0.001
+ 5. presence*genre	44.73	24	.006	46.72	4	< 0.001
+ 6. presence*conn.*genre	.00	.00	1.000	44.73	24	= 0.006

B. Parameter estimates *presence of SoC* for Model 6

Parameter	Estimate	Std. Error	<i>z</i>	<i>p</i>
constant	3.157	.206	15.304	< .001
[conn. = kejian]	-3.850	1.429	-2.694	.007
[conn. = suoyi]	-.416	.327	-1.272	.203
[conn. = yinci]	.120	.283	.424	.672
[conn. = yin'er]	.638	.255	2.504	.012
[genre = argu.]	-.043	.295	-.147	.883
[genre = narra.]	.082	.286	.286	.775
[presence_SoC = explicit]	.440	.264	1.665	.096
[presence_SoC = implicit]	-.354	.321	-1.101	.271
[conn. = kejian] * [presence = explicit]	-.440	2.017	-.218	.827
[conn. = kejian] * [presence = implicit]	5.371	1.455	3.692	< .001
[conn. = suoyi] * [presence = explicit]	-.440	.446	-.987	.324
[conn. = suoyi] * [presence = implicit]	1.431	.435	3.285	.001
[conn. = yinci] * [presence = explicit]	-1.577	.475	-3.323	.001
[conn. = yinci] * [presence = implicit]	.802	.406	1.975	.048
[conn. = yin'er] * [presence = explicit]	-2.364	.496	-4.764	< .001
[conn. = yin'er] * [presence = implicit]	-.203	.406	-.500	.617
[genre = argu.] * [presence = explicit]	-.013	.379	-.034	.973
[genre = argu.] * [presence = implicit]	.211	.446	.472	.637
[genre = narra.] * [presence = explicit]	-.198	.374	-.529	.597
[genre = narra.] * [presence = implicit]	.033	.443	.074	.941
[conn. = kejian] * [genre = argu.] * [presence = explicit]	.056	2.014	.028	.978
[conn. = kejian] * [genre = argu.] * [presence = implicit]	-.167	.372	-.449	.653
[conn. = kejian] * [genre = argu.] * [presence = no SoC]	.043	2.022	.022	.983
[conn. = kejian] * [genre = narra.] * [presence = explicit]	.116	2.014	.058	.954
[conn. = kejian] * [genre = narra.] * [presence = implicit]	-.114	.376	-.305	.761
[conn. = kejian] * [genre = narra.] * [presence = no SoC]	-.082	2.020	-.040	.968
[conn. = suoyi] * [genre = argu.] * [presence = explicit]	-.159	.448	-.354	.723

[conn. = suoyi] * [genre = argu.] * [presence = implicit]	-.005	.391	-.013	.990
[conn. = suoyi] * [genre = argu.] * [presence = no SoC]	-.346	.497	-.696	.486
[conn. = suoyi] * [genre = narra.] * [presence = explicit]	.532	.406	1.309	.190
[conn. = suoyi] * [genre = narra.] * [presence = implicit]	-.693	.419	-1.654	.098
[conn. = suoyi] * [genre = narra.] * [presence = no SoC]	.492	.427	1.150	.250
[conn. = yinci] * [genre = argu.] * [presence = explicit]	.268	.519	.516	.606
[conn. = yinci] * [genre = argu.] * [presence = implicit]	.068	.394	.173	.863
[conn. = yinci] * [genre = argu.] * [presence = no SoC]	-.631	.446	-1.415	.157
[conn. = yinci] * [genre = narra.] * [presence = explicit]	1.175	.465	2.524	.012
[conn. = yinci] * [genre = narra.] * [presence = implicit]	-1.237	.461	-2.682	.007
[conn. = yinci] * [genre = narra.] * [presence = no SoC]	.292	.381	.765	.444
[conn. = yin'er] * [genre = argu.] * [presence = explicit]	.325	.573	.567	.571
[conn. = yin'er] * [genre = argu.] * [presence = implicit]	.455	.415	1.097	.273
[conn. = yin'er] * [genre = argu.] * [presence = no SoC]	-.732	.398	-1.839	.066
[conn. = yin'er] * [genre = narra.] * [presence = explicit]	.116	.605	.192	.848
[conn. = yin'er] * [genre = narra.] * [presence = implicit]	-.154	.441	-.350	.726
[conn. = yin'er] * [genre = narra.] * [presence = no SoC]	-.059	.355	-.167	.867

Appendix 4. The analysis of *identity of SoC*

A. Results of the General Log-linear analysis of *identity of SoC*

Model	χ^2 (model)	df	p	χ^2 (factor)	df	p (factor)
1. constant+conn.	767.73	40	< .001			
+ 2. genre	767.73	38	< .001	0	2	> .05
+ 2. identity	718.12	36	< .001	49.61	2	< 0.001
+ 3. identity*conn.	343.32	28	< .001	374.80	8	< 0.001
+ 4. identity*genre	82.36	24	< .001	260.96	4	< 0.001
+ 5. identity*conn.*genre	.00	.00	1.000	82.36	24	< 0.001

B. Parameter estimates *identity of SoC* for Model 5

Parameter	Estimate	Std. Error	z	p
constant	2.140	.343	6.239	< .001
[conn. = kejian]	-2.833	1.455	-1.947	.052
[conn. = suoyi]	.534	.432	1.236	.216
[conn. = yushi]	1.721	.372	4.620	< .001
[conn. = yin'er]	-9.725E-17	.485	.000	1.000
[genre = argu.]	-3.353E-18	.485	.000	1.000
[genre = narra.]	.973	.403	2.418	.016
[identity_SoC = current]	.534	.432	1.236	.216
[identity_SoC = author]	1.174	.392	2.992	.003
[conn. = kejian] * [identity = current]	2.601	1.508	1.725	.084
[conn. = kejian] * [identity = author]	3.686	1.473	2.502	.012
[conn. = suoyi] * [identity = current]	-.051	.546	-.094	.925
[conn. = suoyi] * [identity = author]	-.691	.515	-1.341	.180
[conn. = yushi] * [identity = current]	-2.690	.624	-4.310	< .001
[conn. = yushi] * [identity = author]	-5.728	1.475	-3.884	< .001
[conn. = yin'er] * [identity = current]	-.323	.632	-.511	.610
[conn. = yin'er] * [identity = author]	-.711	.588	-1.210	.226
[genre = argu.] * [identity = current]	-.802	.677	-1.185	.236
[genre = argu.] * [identity = author]	.567	.541	1.049	.294
[genre = narra.] * [identity = current]	-1.296	.571	-2.269	.023
[genre = narra.] * [identity = author]	-2.583	.617	-4.188	< .001
[conn. = kejian] * [genre = argu.] * [identity = current]	-1.235	.988	-1.249	.212
[conn. = kejian] * [genre = argu.] * [identity = author]	-.423	.293	-1.444	.149
[conn. = kejian] * [genre = argu.] * [identity = character]	9.924E-17	2.058	.000	1.000
[conn. = kejian] * [genre = narra.] * [identity = current]	2.093	.516	4.057	< .001
[conn. = kejian] * [genre = narra.] * [identity = author]	-.417	.593	-.704	.482
[conn. = kejian] * [genre = narra.] * [identity = character]	-.973	2.040	-.477	.633
[conn. = suoyi] * [genre = argu.] * [identity = current]	-.851	.698	-1.218	.223

[conn. = suoyi] * [genre = argu.] * [identity = author]	.198	.345	.572	.567
[conn. = suoyi] * [genre = argu.] * [identity = character]	-.232	.625	-.371	.711
[conn. = suoyi] * [genre = narra.] * [identity = current]	-.031	.517	-.060	.952
[conn. = suoyi] * [genre = narra.] * [identity = author]	-2.241	1.504	-1.490	.136
[conn. = suoyi] * [genre = narra.] * [identity = character]	-.166	.512	-.325	.745
[conn. = yushi] * [genre = argu.] * [identity = current]	-.497	1.035	-.480	.631
[conn. = yushi] * [genre = argu.] * [identity = author]	2.728	1.460	1.869	.062
[conn. = yushi] * [genre = argu.] * [identity = character]	-.184	.531	-.348	.728
[conn. = yushi] * [genre = narra.] * [identity = current]	.323	.727	.444	.657
[conn. = yushi] * [genre = narra.] * [identity = author]	1.609	2.054	.784	.433
[conn. = yushi] * [genre = narra.] * [identity = character]	-1.016	.453	-2.244	.025
[conn. = yin'er] * [genre = argu.] * [identity = current]	-.296	.777	-.381	.703
[conn. = yin'er] * [genre = argu.] * [identity = author]	.371	.400	.927	.354
[conn. = yin'er] * [genre = argu.] * [identity = character]	.778	.638	1.219	.223
[conn. = yin'er] * [genre = narra.] * [identity = current]	.323	.596	.542	.588
[conn. = yin'er] * [genre = narra.] * [identity = author]	-.077	.832	-.092	.926
[conn. = yin'er] * [genre = narra.] * [identity = character]	-.196	.578	-.339	.735

On the subjectivity of Mandarin reason connectives: Robust profiles or genre-sensitivity

4.1 Introduction

Like the English connective *because*, Mandarin reason connectives have the properties of linking pairs of clauses within discourse, expressing the relation that holds between them, and specifying which is the antecedent of the causally related events. Example (1) demonstrates the ways in which various reason connectives are used in Mandarin discourse. It also exemplifies the so-called *forward linking*, with the antecedent preceding the consequent, which is considered to be the canonical usage for most Mandarin connectives expressing causal coherence relations in both spoken and written Chinese, including *yinwei* ‘because’ and *suoyi* ‘so, therefore’ (see Biq 1995).

- (1) *Jiran/ Youyu/ Yinwei laoshi you shi, xiaowu women zixi.*
(based on Xing 2001)
‘**Because** the teacher has other engagements, this afternoon we study by ourselves.’

Apparently, *jiran* ‘because/since’, *yinwei* ‘because’, and *youyu* ‘because’ are (partially) synonymous in the above sentence in the sense that they all express the meaning of causality, and specify that the segment containing them is the causal antecedent for the upcoming information. However, it is intuitively obvious for native speakers of Mandarin that the type of causal relation can vary as one reason connective is substituted for another. *Yinwei* and *youyu* denote that there is a cause and effect relation between the facts expressed by connected segments: the fact that the teacher has other engagements leads to the fact that *we study by ourselves*. The type of causal relation is changed when *jiran* is used instead (see Xing’s (2001) descriptive vs. inferential causality). The relation marked with *jiran* can be paraphrased as the fact that the teacher is not available leads to one’s inference or conclusion that *this afternoon we have to study by ourselves*. This type of causal relation is termed an epistemic causal relation.

These different interpretations of (1) make us wonder whether the three reason connectives really express identical meanings. In other languages, it seems that causal connectives display differences with regard to *subjectivity*: the degree to which the speaker is involved in the construction of the causal relation (see Sanders & Sweetser 2009 for an overview). For example, in Dutch, *want* ‘because’ marks a higher degree of subjectivity than *omdat*

'because' (Degand & Pander Maat 2003; Sanders & Spooren 2009; Spooren, Sanders, Huiskes & Degand 2010). Sentence (2) exemplifies the type of subjective causal relations *want* typically expresses: the speaker is highly involved in justifying his conclusion that *Peter must be ill* with the argument that *he looks pale*. In contrast, sentence (3) exemplifies the typical usage of *omdat* in an objective causal relation with lower speaker-involvement. In (3), the speaker objectively describes the causal connection between observable facts, rather than subjectively arguing towards a conclusion.

- (2) *Peter zal wel ziek zijn, want hij ziet bleek.*
 'Peter must be ill, **because** he looks pale.'
- (3) *Peter blijft thuis omdat hij ziek is.*
 'Peter is staying at home **because** he is ill.'

Similar differences in subjectivity have been attested in corpus-based studies on reason and result connectives in several European languages (Evers-Vermeul, Degand, Fagard, & Mortier 2011; Pander Maat & Sanders 2001; Pit 2003; Zufferey 2012). In a corpus-based study on Mandarin Chinese (Chapter 3), the subjectivity profiles of the result connectives *kejian*, *suoyi*, *yinci*, *yin'er*, and *yushi*, which can all be translated as 'so/therefore', were systematically analyzed in terms of domain and three other subjectivity indicators. Four causal domains were distinguished on the basis of Sweetser (1990), and Stukker, Sanders and Verhagen (2008). These different domains are illustrated in (4) – (7).

- (4) *Youyu yong li guo meng, wumen bei qiangbi fantan huilai.*
 '**Because** she pushed too violently, the door was bounced back by the wall.'
- (5) *Shang shiji ba jiu shi niandai, yinwei Ø xiang duo zheng xie qian, yixie nongmin likai xiangcun dao chengzhen dagong.*
 'In the 1980s and 1990s, **because** (they) want to earn more money, some farmers left the villages to look for job opportunities in the towns.'
- (6) *Jiran zhufang shi minsheng bixupin, jiu bu neng renping muqian yuan fei jianquan de shichang qu zhipei zhufang.*
 '**Since** housing is a basic need, the government should not allow the market, which is far from robust so far, to allocate housing.'
- (7) *Jiran ni zhidao le, hai wen wo gan ma?*
 '**Since** you already know (about it), why are you asking me?'

The non-volitional content domain concerns objective causal processes in observable reality (see (4)). The volitional content domain also describes the causal relation between observable facts, but human intentions play a role in

it (see (5)). The epistemic domain involves subjective reasoning processes on the basis of rationality (see (6)). The speech-act domain involves imperative or interrogative utterances and gives the cause to the speech act underlying the imperative or interrogative forms, for example, the speech act of asking a question, as in (7). Different causal domains are associated with different degrees of subjectivity. There is considered to be a stepwise increase of subjectivity from the non-volitional content domain (4) to the volitional content domain (5), and to the epistemic and speech-act domains, (6) and (7). The other three subjectivity indicators that were used in the study presented in Chapter 3 are: propositional attitude, and the presence and identity of a Subject of Consciousness (henceforth SoC), which refers to the person responsible for constructing the causal relation. These three factors will be discussed in depth in section 4.2.2.

The corpus-based study on Mandarin result connectives (Chapter 3) has yielded three robust profiles across genres: the subjective *kejian*, which is restricted to the epistemic domain, the less subjective *yushi*, which typically expresses causal relations in the volitional content domain, and the objective *yin'er*, which is typically used to mark causal relations in the non-volitional content domain. Importantly, the study also suggests that the subjectivity account, which is derived from Germanic data, is also relevant to accounting for the meaning differences between Mandarin causal connectives.

This raises the question as to whether the subjectivity account is also relevant to the differences between the reason connectives *jiran*, *yinwei*, and *youyu* in (1), and whether genre plays a role in determining the subjectivity values of these connectives. Several remarks are available in the literature that hint at the relevance of subjectivity for accounting for the distribution of reason connectives in Chinese discourse. Xing (2001: 73) has put forward a suggestion that *jiran* is more subjective than *yinwei* because it emphasizes inferences. He claims that *yinwei* is rarely followed by questions (see (8), in which the asterisk indicates the sentence is semantically or logically incomprehensible despite the fact that it is grammatical, structurally speaking), whereas *jiran* often co-occurs with a consequent in the form of a question (see (9)). The frequent co-occurrence with questions suggests that *jiran* often appears in the speech-act domain, a highly subjective causal category. Sweetser (1990: 78) wrote: “If an utterance is imperative or interrogative in form, then it cannot reasonably be causally conjoined to another utterance except at the speech-act level.”

- (8) * *Yinwei ta fandui, nimen weishenme hai yao gan?*
 ‘Because he objects to it, why are you still going to do it?’
- (9) *Jiran ta fandui, nimen weishenme hai yao gan?*
 ‘Since he objects to it, why are you still going to do it?’

Xing supports his claims about *yinwei* and *jiran* by providing anecdotal examples. Empirical evidence for the relevance of the notion of subjectivity comes from corpus-based studies conducted by Li and Liu (2004) and Li (2011). Using a newspaper corpus (i.e., People's daily newspaper in the year 1996), Li and Liu (2004) provide a subjectivity analysis on the differences between *jiran* and *youyu*. They conclude that *jiran* is more subjective than *youyu* on the grounds of their distributional proportions across domains of use. It is found that 98% of the cases of *jiran* are used in subjective domains (i.e., epistemic and speech-act domains), whereas only 12% of the cases of *youyu* are used in subjective domains. In addition, Li and Liu (2004) have pointed out that a considerable number of cases with *jiran* (i.e., 32%) are used in the speech-act domain.

Li (2011) compares between *youyu* and *yinwei* in terms of subjectivity. It was observed that 63 out of 100 cases of randomly selected fragments with *yinwei* (from *Guojia yuwei pingheng yuliaoku* 'Balanced corpus constructed by the national language committee') are in the objective content domain. This result is then compared to the finding about *youyu* in her earlier study, Li and Liu (2004): 88 out of 100 cases of *youyu* are used to express the objective content causal relation. On the basis of this disparity with respect to domains of use, Li (2011) concludes that *yinwei* tends to be more subjectively-oriented than *youyu*.

The two corpus-based studies mentioned above are valuable, because they have demonstrated how systematic subjectivity analyses contribute to increasing our understanding of *jiran*, *yinwei*, and *youyu* in actual language use. As pioneering studies of the type, however, there are inevitably certain limitations. They only compare pairs of reason connectives rather than the full range of reason connectives simultaneously. In addition, these studies tend to avoid carrying out quantitative analyses beyond the simplest of descriptive statistics: percentages. Percentages simply describe the data, and do not test for significance as inferential statistics do (McEnery & Hardie 2012). Hence, from percentages alone, it is difficult to tell how likely it is that a particular result occurs simply due to chance. For corpus-based quantitative analyses, inferential statistics should be more informative. This is particularly because inferential statistics enable linguists to generalize to conclusions beyond actual observations (see Núñez 2007). Furthermore, these two corpus-based studies have not investigated the impact of genre on the meaning and use of connectives. Studies on French, German, and Dutch causal connectives have shown that the distribution of connectives over objective or subjective causal categories seems to vary in relation to the context (Degand & Pander Maat 2003; Frohning 2007; Pit 2003; Zufferey 2012). Subjective causal connectives such as French *car*, German *denn*, and Dutch *want*, which are all roughly translated as 'because', display consistent usage patterns across text types such as newspapers, novels, and periodicals.

By contrast, usage patterns of their objective counterparts, French *parce que*, German *weil*, and Dutch *omdat*, are less consistent across these text genres (Stukker & Sanders 2012). These results raise the question whether the subjectivity profiles are really part of the inherent semantic characteristics of the connectives themselves, or whether they are (partially) determined by the pragmatics of the context.

The corpus-based study on result connectives in Mandarin Chinese (Chapter 3) has also drawn our attention to the genre-sensitivity of connectives' subjectivity values. Results have shown that some Chinese result connectives (i.e., *kejian*, *yin'er*, and *yushi*, which are all translated as 'so/therefore') display robust subjectivity profiles across genres, whereas the subjectivity meanings of other relatively frequent result connectives (i.e., *suoyi* and *yinci*, both translated as 'so/therefore') vary across genres.

In view of the limitations presented above, this paper aims to provide a more thorough and comprehensive subjectivity analysis of the complete set of reason connectives in present-day Chinese written discourse, i.e., *jiran*, *yinwei*, and *youyu* (all translated as 'because'). Unlike previous studies on Chinese reason connectives (e.g., Li 2011; Li & Liu 2004), we adopt an integrative approach to the issue of subjectivity by looking at several indicators of subjectivity instead of one. Furthermore, we systematically take the impact of genre on the meaning and use of causal connectives into account. We focus on detecting variations across three genres, i.e., news reports, novels and opinion pieces. Our first two research questions are the following (cf. Research question 4 & 5 raised in Chapter 1).

1. What are the subjectivity profiles of *jiran*, *yinwei*, and *youyu*?
2. Are the observed subjectivity profiles sensitive to text genre?

We are also interested in the frequently used reason connective *yinwei* (with the frequency of 2.09 per 10,000 words),¹ in its specific property of being equally felicitous in sentence-initial and inter-sentential positions to express forward linking as well as backward linking (cf. (1) and (10)).

(10) *Xiawu women zixi, yinwei laoshi you shi.*

'This afternoon we study by ourselves, **because** the teacher has other engagements.'

Traditional grammars of Mandarin Chinese have treated forward linking as the default sequence order for subordinate clauses, including causal clauses (Chao 1968). But numerous corpus-based studies have shown that backward linking (with the consequent preceding the antecedent) is also very prominent, at least for *yinwei* (Biq 1995; Su 2002; Tsai 1996; Y. Wang 2002, 2006). Song and Tao (2009) have conducted a corpus-based study to

analyze the differences between sentence-initial and inter-sentential *yinwei*. According to their study, in both conversation and written corpora inter-sentential *yinwei* often appears after a negation, a comparison, a strong assertion with degree words such as *feichang* ‘extremely’ and *juedui* ‘absolutely’, or a statement with a modality marker such as *keneng* ‘maybe, probably, possibly’ and *bixu* ‘must’. In these cases, inter-sentential *yinwei* functions as a follow-up justification to address “potential enquiries, disagreement, suspicion, or general unexpectedness” (Song & Tao 2009: 87). By contrast, sentence-initial *yinwei* is informative in nature. It provides background information for interpreting the event described in the following clause, especially in narrative texts.

Song and Tao (2009) analyze these differences from a functional point of view, and conclude that inter-sentential and sentence-initial *yinwei* clauses are two different linguistic constructions functioning as an interactional device, and an information-sharing device, respectively. It is interesting to find out whether these differences can be related to the degrees of subjectivity each sequence order expresses. This seems highly probable. For instance, the fact that inter-sentential *yinwei* often co-occurs with evaluative words *feichang* ‘extremely’ and *juedui* ‘absolutely’, and modality markers *keneng* ‘maybe, probably, possibly’ and *bixu* ‘must’ seems to suggest that it is closely related to conclusions and inferences (see Zhang & Zhang 2011 for a similar claim on the basis of anecdotal examples that inter-sentential *yinwei* often co-occurs with judgments). Therefore, it is likely that inter-sentential *yinwei* marks a high degree of subjectivity. Meanwhile, the fact that sentence-initial *yinwei* is more information-oriented seems to suggest that it is more objective (see also Gao 2013 for the introspective reflection that sentence-initial *yinwei* is mainly used in the content domain). In order to test our hypotheses concerning the two sequence orders, we provide separate subjectivity analyses for inter-sentential *yinwei* and sentence-initial *yinwei*, respectively. Our third research question is as follows.

3. Do sentence-initial *yinwei* and inter-sentential *yinwei* express different degrees of subjectivity?

In sum, this paper provides a subjectivity analysis of Mandarin reason connectives *jiran*, *yinwei*, and *youyu*, as they occur in natural, connected, written Mandarin discourse. By providing corpus-based analyses of actual connective use in different genres, we expect to depict a subjectivity profile for each reason connective, so as to provide an account for the intuition that the type of causal relation in (1) varies as one reason connective is substituted for another. Furthermore, we investigate the genre-sensitivity of the observed subjectivity profiles, and explore whether sentence-initial and inter-sentential *yinwei* display different preferences in terms of subjectivity.

4.2 Method

4.2.1 Corpora

In our previous study on Mandarin result connectives (Chapter 3), it was found that a connective's subjectivity profile can vary across genres. In order to depict the subjectivity profiles of the reason connectives thoroughly, we collected the samples evenly from three text categories: news reports, opinion pieces, and novels. Samples of the former two categories were taken from *People's Daily Online* (hereafter PPD),² and fragments of novels were taken from *the CCL Corpus*.³ Apart from being large in size, another important reason for selecting PPD and CCL is that it is possible to keep track of wide contexts with the two corpora. Table 1 shows the number and nature of connective fragments taken from these corpora.

Table 1. Number and nature of connective fragments in the sample

Connective	News report (PPD)	Opinion piece (PPD)	Novel (CCL)	Total
<i>jiran</i>	75	75	75	225
<i>yinwei</i>	75	75	75	225
<i>youyu</i>	75	75	75	225
Total	225	225	225	675

From each genre, we randomly collected seventy-five occurrences of *jiran*, seventy-five occurrences of *yinwei*, and seventy-five occurrences of *youyu*. In total, we arrived at 675 fragments. The prepositional use of *youyu* 'due to' (see (11)) was not included in our collection of fragments. Altogether, 42 occurrences of prepositional *youyu* were discarded and were replaced with fragments where *youyu* is used as a connective. Similarly, 20 occurrences of prepositional *yinwei* 'because of' (such as (12)) were discarded and were replaced with the connective usage.

(11) *Youyu* *mou zhong yuanyin, ta mei you zhijie hui jia.*

'Due to some reason, he didn't go home directly.'

(12) *Yinwei pijuan, ta pa zai zhuozi shang shuizhao le.*

'Because of tiredness, he bent over the desk and fell asleep.'

4.2.2 Analysis

In order to give a precise characterization of the degrees of subjectivity expressed by the three reason connectives, an analytical model (see Table 2) was created to incorporate all the major textual subjectivity indicators suggested by previous theoretical and corpus-based studies (Degand &

Pander Maat 2003; Fagard & Degand 2010; Langacker 1985, 1990; Li 2011; Li & Liu 2004; Sanders & Spooren 2009; Spooren et al. 2010; Sweetser 1990). These subjectivity indicators include domain, propositional attitude, the presence of an SoC, and the identity of the SoC.

Table 2. The analytical model with variables and degrees of subjectivity

Variable	Degree of subjectivity
Domain	Non-vol content < Vol content < Speech act / Epistemic
Propositional attitude	Physical fact < Mental fact < Speech act / Judgment
The presence of SoC	No SoC < Explicit SoC < Implicit agent SoC < Implicit concluder SoC
The identity of SoC	Character SoC < Speaker SoC

Along these four dimensions, we compared the relative degrees of subjectivity that are encoded in each connective.⁴ The fragments were coded independently by the author and two visiting scholars from Zhejiang University, who were temporarily doing their research at the Utrecht Institute of Linguistics OTS. The inter-coder agreement is quite high with Kappa between 0.75 and 0.83 (for propositional attitude: Kappa = 0.83; for domain: Kappa = 0.82; for presence of SoC: Kappa = 0.80; for identity of SoC: Kappa = 0.75). Discrepancies among the coders' analyses were discussed until agreement was reached.

Domain

On the basis of the Domain Theory (Sweetser 1990) and the notion of volitionality (Stukker et al. 2008), we distinguished between four causal domains in the current study: the non-volitional content, volitional content, epistemic, and speech-act domains. In order to interpret the domains accurately, we used a paraphrase test. The paraphrase test is presented in Table 3, in which P and Q correspond to the antecedent and the consequent of the causal relation, respectively.

Table 3. The paraphrase test used in the domain analysis

Domain	Paraphrase
Non-volitional content	P leads to the fact that Q, and no intention is involved in Q
Volitional content	P leads to the intentional act that Q
Epistemic	P leads to the claim / decision / inference / conclusion that Q
Speech act	P leads to the question / advice / command / promise that Q

Volitional content relations are considered to be more subjective than non-volitional relations because they involve human intentions. Epistemic relations are of still higher subjectivity because they directly involve opinions, beliefs, or inferences. Speech-act causal relations are also highly subjective

because performing a speech act is bound to a situation in which the speaker is present.

In natural discourse the division between domains is occasionally not clear-cut. In cases of ambiguity, we opted for the more objective interpretation. For example, see (13).

- (13) *Lao Wang huilai le, yinwei ta de bao zai zher.* (Shen 2003: 198)
'Lao Wang came/must have come back, **because** his bag was/is here.'

The ambiguity arises in (13) because tense is not explicitly marked on the verbs or any other words in Mandarin clauses. When no context or time adverbials (e.g., *yesterday* and *tomorrow*) are available, we could interpret (13) as having either a past tense or a present tense. In the former case, there is a content causal relation holding between the cause that *Lao Wang's bag was here* and the consequence that *he came back (to take his bag)*. In the latter case, the epistemic interpretation arises: a conclusion (i.e., *Lao Wang must have come back*) is made on the basis of an observable fact (i.e., *his bag is here*). In the analysis, we interpreted the fragment in the former way; that is, we adopted an objectivity bias, for the sake of consistency.

During the analysis, we saw a few cases of *chain causality*, in which readers need to add an inferential link to the causal relation to represent the sentence completely. For example, in (14), readers need to add the following inferential link between P and Q: people like to go to that restaurant. Chain causality is special, because it is hard to decide whether it is volitional or non-volitional. On the one hand, human intention is implied (through the inferential link). On the other hand, it is not directly involved in the consequent Q. We chose to interpret chain causality as being non-volitional content, to be consistent with the objectivity bias we adopted while analyzing ambiguous cases.

- (14) *Youyu fancai kekou wumeijialian fuwuzhoudao, "yipinzhai" tiantian mentingruoshi, zuowuxuxi.*
'**Because** the food is nice as well as inexpensive, and the service is good, "Yipinzhai" (restaurant) is as crowded as the market everyday, and there are no empty seats.'

Propositional attitude

While the domain analysis focused on the subjectivity of the coherence relation as a whole, the analysis of propositional attitude focused on the subjectivity of the result segment (Q). We distinguished four values of propositional attitude. We coded the result segment as expressing a speech act, a judgment, an observable physical fact, or a mental fact. A segment is

coded as a speech act if it is in the form of a general question, a rhetorical question, or an imperative. A segment expresses a judgment if it presents opinions, decisions, conclusions, or inferences. A segment expresses a physical fact if it describes events or circumstances in the observed world. A segment expresses a mental fact if it depicts mental states such as personal feelings, mental processes, or psychological activities. For example, we coded the consequent in (14) as a physical fact, and the consequent in (15) as a judgment. (16) was considered to have a speech-act consequent, and (17) a consequent representing a mental fact.

- (15) *Cong xueshu yanjiu de jiaodu kan, zhe xie guandian meiyou shenme xueshu jiazhi ke yan, yinwei tamen cong genben shang weibei le lishi shishi.*

‘From the academic point of view, these opinions do not have any academic value at all, **because** they fundamentally went against the historical facts.’

- (16) *Jiran zhexie shi luan shoufei, weishenme shi ji nian lai jin'erbuju ne?*

‘**Since** these are illegal charges, why didn’t they disappear after over ten years’ prohibition?’

- (17) *Qiaoqiao dui na ge furen yinxiang shenke, yinwei na ge furen zhang de hen xiang ta de mama.*

‘Qiaoqiao had a deep impression of that woman, **because** the woman resembled her mother a lot.’

Physical facts are considered to be less subjective than mental facts, judgments and speech acts because physical facts do not involve mental processes. Mental facts are less subjective than judgments and speech acts since more reasoning processes are involved to form an opinion, make a conclusion, or ask a question.

The presence of SoC

An SoC is an animate subject – a person – whose intentionality is conceptualized as the ultimate source of the causal event, be it an act of reasoning or some real-world activity (Pander Maat & Sanders 2001: 251). We made a three-way distinction between utterances with no SoC, with an implicit SoC, and with an explicit SoC. There is no SoC in utterances where the causal relation exists independent of human intention. For example, in (17) the fact that the woman looked like Qiaoqiao’s mother automatically leads to the deep impression that Qiaoqiao had of that woman. Qiaoqiao did not intentionally decide to remember that woman. There is no SoC in it. When there is an SoC, we made a distinction between implicit and explicit SoCs. If the SoC is explicitly referred to by linguistic elements in the

consequent, the SoC is explicit. There is an explicit SoC in (18), *the villagers*, whose intention is involved in the consequent. If the SoC is not linguistically realized, the SoC is implicit. In (15), there is an SoC, who is drawing a conclusion (i.e., *these opinions do not have any academic value at all*) on the basis of an argument (i.e., *they fundamentally went against the historical facts*). The SoC is the implicit speaker.

- (18) *2001 nian qian yinwei gongdian wu baozhang, dianbingxiang, xiyiji he kongtiao quandou yong bu liao, jin sifenzhiyi de cunmin bu yuan zai cun li juzhu.*

'Because the supply of electricity was not guaranteed before the year 2001, and all the fridges, washing-machines and air-conditioning units could not be used, nearly one-fourth of the villagers were not willing to live in the village.'

Since the SoC is the ultimate source of the subjective construction of the causal relation, a relation with an SoC is more subjective than one with no SoC. In addition, following Langacker's (1990) view of subjectivity in terms of on-stage or off-stage conceptualization, utterances with an explicit reference to the SoC (i.e., on-stage conceptualization) should be considered less subjective than utterances with an implicit SoC (i.e., off-stage conceptualization). According to Langacker (1990), an utterance is objectified when the speaker (as a conceptualizer) is put on stage and becomes observable.

Apart from the implicit concluder SoC exemplified in (15), subject-drop languages like Mandarin have another type of implicit SoC – the implicit agent SoC. Unlike the implicit concluder SoC that occurs in the epistemic domain, the implicit agent SoC appears in the volitional content domain functioning as the active agent of an event. The occurrence of an implicit agent SoC is exemplified in (19), where the agent of the event expressed in the second clause, *his two neighbors*, is implicit via subject-drop (indicated with \emptyset).

- (19) *Ta de liang ge linju youyu zanshi mei you anzhuang kuandai, \emptyset jingchang pao dao ta jia lai tongguo shipin xiang zhuanjia zixun.*

'Because his two neighbors did not install ISDN broadband at that time, (they) often run to his house to consult experts through the internet.'

We think implicit agent SoCs and implicit concluder SoCs should be analyzed as separate categories, because they express different degrees of subjectivity. The agent SoC in (19), *his two neighbors*, can be recovered from

the linguistic context. That is, *his two neighbors* is actually present in the preceding context. The concluder SoC in (15), however, cannot be recovered through such mechanisms. In this sense, it is more implicit than the SoC in (19). Hence, implicit concluder SoCs are considered more subjective than implicit agent SoCs.

The identity of SoC

In utterances in which an SoC is present, the SoC can be a speaker or a character. There is a speaker SoC in (15), and a character SoC in (19). In (15) it is the speaker/author who draws the conclusion that those opinions have no academic value at all, whereas in (19) it is the character *his two neighbors* who volitionally run to *his house* to consult experts through the internet. Speaker SoCs are considered more subjective than character SoCs. The author/speaker can be considered the first voice in the discourse, who has constant access to her feelings and thoughts. She does not have access to the feelings and thoughts of a third person. As a result, *I think Utrecht is nice* can be a direct report of an inner feeling, whereas *He thinks Utrecht is nice* is a description of an evaluation. As a consequence, first person evaluations are more subjective than third person evaluations. What distinguishes the speaker/author from a character as SoC is that the speaker concerns a first voice, which is grounded in the Deictic Centre of Communication (Sanders, Sanders & Sweetser 2009). This reflects Traugott's (1989, 1995) view on subjectivity as closeness to the communicative "here and now": the speaker/author here and now asserts that a particular state of affairs holds. By contrast, the character type concerns a third person in the discourse, who is more distant from the Deictic Centre of Communication. The notion of SoC, and the distinction we make here between speaker/author and character SoC, is related to the notion of *perspective* in literary studies (Fludernik 1993) and *voice* in linguistics (Ducrot 1980): the speaker/author is the first voice in the discourse, and the character is another person whose inner thoughts and evaluations can be accessible in a narrative (Sanders 2010).

In natural discourse we may encounter clauses in which the author is quoting a causal relation constructed by another person. For example, the concluder SoC *I* within the quotation marks in (20) refers to a third person (i.e., Haobu) rather than the author. Nevertheless, Haobu functions as the actual speaker in the communicative here and now. In the analysis, we coded cases like this as containing a speaker SoC.⁵

- (20) *Haobu biaooshi: "wo renwei keyi changshi dui zhexie shipin de sheru jiayi jiezhi, er bushi wanquan bu chi, yinwei rang suoyou ren meitian zhi chi shucui shuiguo wanquan bu keneng."*

‘Haobu said: “I think (we) may attempt to control the intake of this type of food, but not to completely stop eating it, **because** it is absolutely impossible to ask everyone to eat only vegetables and fruits.”’

4.3 Results

We conducted a comprehensive analysis on the subjectivity profiles of the connectives in question with the four subjectivity indicators introduced in the previous section. For each subjectivity indicator, we conducted General Log-linear Analyses to find the model with the lowest number of parameters (such as connective and genre), and the best fit to the observed data. Statistical details about the models per subjectivity factor can be found in Appendices 1 to 4.

4.3.1 Domain

Figure 1 shows the distribution of the three connectives over domains. The distribution over domains varies per connective ($\chi^2 [6] = 378.14, p < 0.001$). Neither main effects of genre nor interaction effects involving genre were observed, indicating that the three connectives under discussion show stable domain profiles across genres. In the following paragraphs, we report the findings on domain profiles per connective.

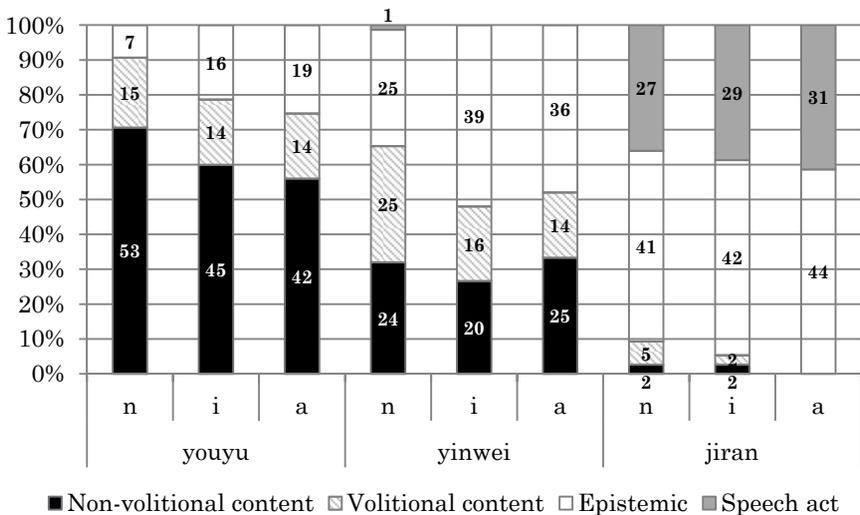


Figure 1. The distribution of connectives over domains per genre (a = argumentative opinion pieces; i = informative news reports; n = narrative novels)

Across genres, *youyu* is typically used in the non-volitional content domain: it is used less often in the volitional content domain ($z = -3.79$, $p < 0.001$), the epistemic domain ($z = -6.67$, $p < 0.001$), and the speech-act domain (cf. the main effect of domain: $z = -4.65$, $p < 0.001$). Fragment (21) illustrates the typical usage of *youyu*: it denotes that one fact (i.e., *water supply was lacking*) causes another (i.e., *farmers reaped nothing at harvest time*) in the natural world, and that the causal connection does not involve human intentions.

(21) *Youyu zheli quefa shuiyuan, di'er nian keliwushou.*

'Because water supply was lacking, in the next year not a single grain was reaped.'

Across genres, *jiran* is typically used to express epistemic and speech act causal relations (epistemic: $z = 5.90$, $p < 0.001$; speech act: $z = 7.11$, $p < 0.001$). As can be seen in Figure 1, *jiran* is rarely used in the content domain (only 11 cases in total). Fragments (22) and (23) demonstrate the epistemic and the speech-act use of *jiran*, respectively.

(22) *Jiran tamen quan dou xiangxin butong de shiwu, tamen bu keneng quan dou shi zhengque de.*

'Since they all believe in different things, it is not possible that all of them are right.'

(23) *Jiran zhenzhong zai Wenchuan, weihe Beichuan bi Wenchuan shou zai geng zhong?*

'Since the epicenter of the earthquake is Wenchuan, why is Beichuan more seriously affected than Wenchuan?'

In the epistemic relation (22), *jiran* links a claim (i.e., *it is not possible that all of them are right*) to its evidence (i.e., the fact that *they all believe in different things*). In the speech-act relation (23), *jiran* is used to connect a question (i.e., *why is Beichuan more seriously affected than Wenchuan?*) with an explanation for asking the question (i.e., the fact that *the epicenter of the earthquake is Wenchuan*).

As shown in Figure 1, language users can use *yinwei* to express all four domains of causality, although its speech-act use is rare (only one occurrence). And the statistics reveal that *yinwei* occurs relatively often in the epistemic domain (cf. the main effect of domain: $z = 2.36$, $p = 0.018$), and less often in the speech-act domain (cf. the main effect of domain: $z = -4.65$, $p < 0.001$). Fragment (24) exemplifies the typical usage of *yinwei*: its epistemic use. Fragments (25) and (26) exemplify the other two domains that *yinwei* often expresses: the non-volitional use (25) where one fact (i.e., *bats were not able to prey on insects under the protecting net*) leads to another fact (i.e.,

insects multiplied their descendants rapidly) in the objective world, and the volitional content use (26) where the fact that the speaker's legs felt numb leads to the speaker's intentional act of getting out of the car (to have a walk).

(24) *Yinwei zhengzhi tizhi gaige juyou jiaoda de fengxian xing, nong bu hao hui dai lai shehui de dongdangbu'an huo wuxu.*

'Because the risk of political reform is relatively high, perhaps it will bring about social instability or disorder.'

(25) *Yinwei zai wangfu baohu xia, bianfu wu fa bushi kunchong, kunchong xingwang fanyan.*

'Because bats were not able to prey on the insects under the protecting net, the insects multiplied their descendants rapidly.'

(26) *Yinwei tui ma le, wo jiu xia le che liuda liuda.*

'Because my legs felt numb, I got out of the car to have a walk.'

4.3.2 Propositional attitude

Figure 2 shows the distribution of the three connectives along the dimension of propositional attitude. Because the connective use in the mental fact category is rare, we collapsed the mental fact category and the other objective category, physical fact, for the sake of statistical power.

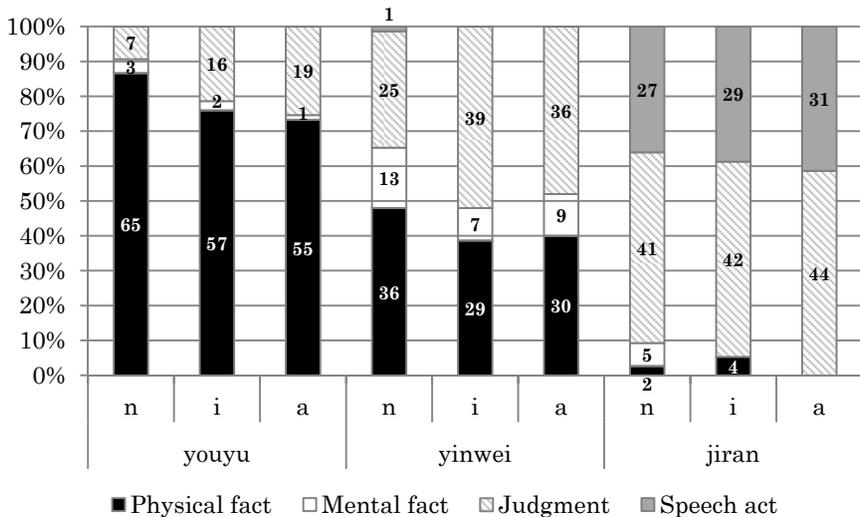


Figure 2. The distribution of connectives over propositional attitude per genre (a = argumentative opinion pieces; i = informative news reports; n = narrative novels)

As can be seen in Figure 2, the distribution over propositional attitude differs per connective ($\chi^2 [4] = 356.28, p < 0.001$). Across genres, *youyu* is typically followed by consequents containing facts (few judgments: $z = -5.76, p < 0.001$; few speech acts: cf. main effect of propositional attitude, $z = -5.38, p < 0.001$), whereas *jiran* typically precedes consequents containing judgments ($z = 7.80, p < 0.001$) and speech acts ($z = 7.33, p < 0.001$). Fragment (21) exemplifies the typical use of *youyu* with fact. Fragments (22) and (23) exemplify the co-occurrences of *jiran* with judgments and speech acts, respectively. *Yinwei* is rarely used in combination with speech acts (cf. main effect of propositional attitude: $z = -5.38, p < 0.001$). It often co-occurs with facts and judgments, and no significant difference was observed with respect to its distribution over these two types of propositional attitude. Fragment (24) is an example of its co-occurrence with a judgment (i.e., *perhaps it will bring about social instability or disorder*). Fragment (25) and (26) exemplify the co-occurrence of *yinwei* with facts (i.e. *insects multiplied their descendants rapidly; I got out of the car to have a walk*).

4.3.3 The presence of SoC

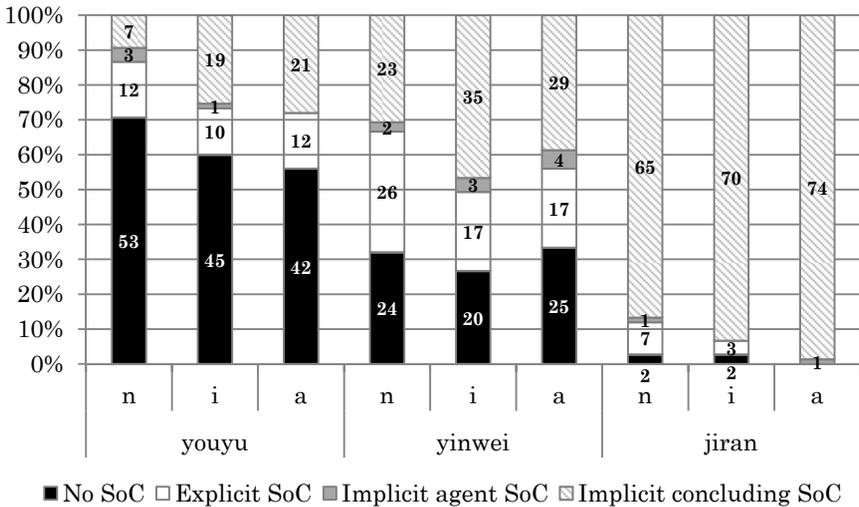


Figure 3. The distribution of connectives over presence of SoC per genre (a = argumentative opinion pieces; i = informative news reports; n = narrative novels)

Figure 3 shows the distribution of the three connectives over the kinds of SoC they can co-occur with. Because implicit agent SoCs are rare (only 15 cases), in the analysis we collapsed implicit agent SoCs and implicit concluder SoCs into a single category. The distribution profiles differ per connective ($\chi^2 [4] = 274.09, p < 0.001$). Furthermore, the distribution over the

presence of SoCs varies per genre ($\chi^2 [4] = 9.96, p = 0.041$): all connectives show an increase of explicit SoCs in novels ($z = 2.37, p = 0.018$).

Across genres, *youyu* is typically used to express causal relations in which no SoC is involved ($z = 5.90, p < 0.001$), and *jiran* is predominantly used to express causal relations in which SoCs remain implicit: it rarely occurs in causal relations without SoC ($z = -6.87, p < 0.001$), and it rarely co-occurs with an explicit SoC ($z = -7.10, p < 0.001$). The typical usages of *youyu* and *jiran* are clearly demonstrated in fragments (27) and (28). In (27), no SoC is involved in the construction of the causal relation. In (28), the causal relation is subjectively constructed by the implicit speaker SoC, who is not explicitly referred to by any linguistic element.

- (27) *Youyu shengyi hao, Gaozhuang ren hen kuai fu qilai le.*
 ‘Because the business was good, the people of Gao Village rapidly became rich.’
- (28) *Jiran wang shang jiu neng kan dao boke, weishenme haiyao mai shu ne?*
 ‘Since we can see the blog on the internet, why should we buy the book?’

Yinwei co-occurs with each category: no SoC, explicit SoC, and implicit SoC (see Figure 3). These uses are illustrated in (29) – (31), respectively.

- (29) *Yinwei tian leng, ta fan le xiaochuan.*
 ‘Because it was cold, he had an asthma attack again.’
- (30) *Tamen renwei Lao'er yiding hui bangmang de, yinwei Ruixuan shi ta de qin gege.*
 ‘They think Lao’er will surely help Ruixuan, because Ruixuan is his brother.’
- (31) *Yinwei Ø yijing sheji fubai, ta benren shi gongzhong renwu, yinci, dui ta de jiandu mei you renhe wenti.*
 ‘Because (it) has involved corruption and he himself is a public figure, there is nothing wrong with keeping him under supervision.’

The statistics show that *yinwei* co-occurs more often with implicit SoCs than with the other two categories (cf. main effects of the presence of SoC: fewer instances of no SoC, $z = -2.29, p = 0.022$; fewer instances of explicit SoC, $z = -2.90, p = 0.004$). Fragment (31) demonstrates this typical usage. The SoC who makes the judgment *there is nothing wrong with keeping him under supervision* is inferable from the context, but it is implicit in the consequent clause. Fragment (31) also exemplifies the combinatory use of a reason connective and a result connective in the same sentence. The ‘because P, so Q’

pattern is not acceptable in English anymore, but in Mandarin it is grammatical to double mark a causal relation in the pattern of ‘reason connective P, result connective Q’. This language-specific property of double-marking is to be dealt with in more detail in Section 4.4.

4.3.4 The identity of the SoC

Figure 4 shows the distribution of the three connectives over the types of SoC they can co-occur with. The distribution patterns vary per connective ($\chi^2 [2] = 87.54, p < 0.001$), and are not significantly affected by genre. We will interpret these findings in more detail.

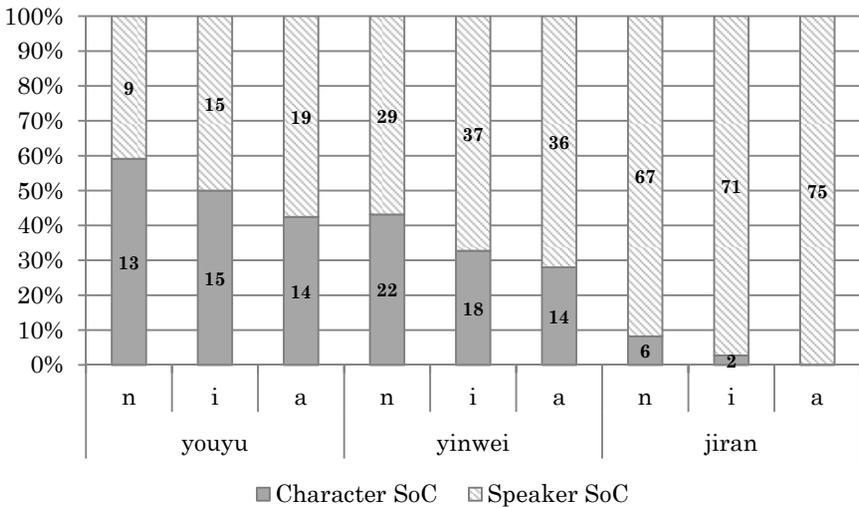


Figure 4. The distribution of connectives over identity of SoC per genre (a = argumentative opinion pieces; i = informative news reports; n = narrative novels)

Across genres, *jiran* has a strong preference for speaker SoCs (few instances of character SoC: $z = -6.66, p < .001$). Fragment (28) serves to exemplify such a usage. In (28) the speaker is the ultimate source of the speech act: *why should we buy the book?*

Yinwei adheres to the main effect of the identity of SoC. Across genres, it co-occurs less often with character SoCs ($z = -3.78, p < .001$) than with speaker SoCs. Fragment(32) exemplifies this usage. The SoC *I* in (32), the person who draws the conclusion, *I guess they had just drunk up the alcohol*, is the actual speaker.

- (32) *Yisheng shuo*: “**yinwei** wo wen dao tamen manshen jiuqi, wo cai tamen ganggang ba jiu he guang.”

‘The doctor said: “**because** I smelt that there were fumes of alcohol all around them, I guess they had just drunk up the alcohol.’”

Across genres, *youyu* is prototypically used in causal relations with character SoCs ($z = 2.23, p = 0.026$), exemplified in fragment (33). The SoC in (33) is the character *he*, who chose to give drawing classes in that beautiful area.

- (33) *Youyu zhe de hua er kai de tebie hao, ta ceng dai xuesheng lai shang xiesheng ke.*

‘**Because** the flowers are particularly good in this area, he once brought his students here to have drawing classes.’

4.3.5 Sentence-initial vs. inter-sentential *yinwei*

To address the third research question, we performed a separate subjectivity analysis on sentence-initial versus inter-sentential *yinwei* using the four subjectivity factors. Among the 225 *yinwei* fragments, 89 cases of *yinwei* take the sentence-initial position and 136 cases of *yinwei* take the inter-sentential position. Table 4 presents an overview of the distribution of two types of *yinwei* over domain, propositional attitude, the presence and identity of the SoC.

Table 4. The distribution of *yinwei* along four dimensions

Subjectivity factor		Sentence-initial <i>yinwei</i> (N=89)		Inter-sentential <i>yinwei</i> (N=136)	
Domain	Non-vol. content	39	(43.8%)	30	(22.1%)
	Volitional content	35	(39.3%)	20	(14.7%)
	Epistemic	15	(16.9%)	85	(62.5%)
	Speech act	0	(0%)	1	(0.7%)
Propositional attitude	Physical fact	69	(77.5%)	26	(19.1%)
	Mental fact	5	(5.6%)	24	(17.6%)
	Judgment	15	(16.9%)	85	(62.5%)
	Speech act	0	(0%)	1	(0.7%)
Presence of SoC	No SoC	39	(43.8%)	30	(22.1%)
	Explicit SoC	24	(27.0%)	36	(26.5%)
	Implicit SoC	26	(29.2%)	70	(51.5%)
Identity of SoC	Character	30	(33.7%)	24	(17.6%)
	Speaker	20	(22.5%)	82	(60.3%)
	d.n.a.*	39	(43.8%)	30	(22.1%)

*This concerns the fragments without SoC.

For each subjectivity factor, we conducted a General Log-linear Analysis to find the best model. Because there was only one instance of *yinwei* expressing the speech-act relation, we collapsed the speech-act relation with another subjective category, the epistemic relation. For the same reason, we collapsed the speech-act category with the judgment category in the analysis of propositional attitude. Hence, we are not able to discuss the speech-act use of *yinwei* in this study. Statistical details about the models per subjectivity factor can be found in Appendices 5 to 8.

It was found that the distribution of sentence-initial and inter-sentential *yinwei* varies along all four dimensions (domain: χ^2 [2] = 47.43, $p < 0.001$; propositional attitude: χ^2 [1] = 46.78, $p < 0.001$; presence of SoC: χ^2 [2] = 14.56, $p = 0.001$; identity of SoC: χ^2 [1] = 20.95, $p < 0.001$). No effects of genre were observed. Across genres, language users prefer to use sentence-initial *yinwei* objectively, to express content causal relations (fewer instances of epistemic relations: $z = -5.40$, $p < 0.001$) holding between facts ($z = 6.38$, $p < 0.001$), in which typically no SoC is involved ($z = 3.73$, $p < 0.001$). When an SoC is present, the identity of the SoC is most often a character rather than a speaker ($z = 4.39$, $p < 0.001$). By contrast, inter-sentential *yinwei* is often used to express more subjective causal relations. It appears more often in epistemic relations (cf. main effect of domain: $z = 4.95$, $p < 0.001$), with judgments (cf. main effect of propositional attitude: fewer facts, $z = -3.04$, $p = 0.002$) and implicit speaker SoCs (cf. main effects of the presence and identity of SoC: fewer instances of no SoC, $z = -3.87$, $p < 0.001$; fewer instances of explicit SoC, $z = -3.23$, $p = 0.001$; fewer character SoCs, $z = -5.28$, $p < 0.001$).

Fragment (34) exemplifies non-volitional content causal relations holding between objective facts where no SoC is involved, the typical objective context in which sentence-initial *yinwei* occurs. The volitional content use is exemplified in fragment (35). When sentence-initial *yinwei* expresses such a causal relation, the SoC is usually a character, such as the *peasants* in (35). Fragment (36) exemplifies the typical use of inter-sentential *yinwei*. The usage is very subjective: in (36) the inter-sentential *yinwei* expresses an epistemic causal relation involving judgment and an implicit speaker SoC.

(34) *Yinwei shangkou fayan, zhengge lian dou zhong qilai.*

'Because the wound is inflamed, the entire face is swollen up.'

(35) *Yinwei shouru keguan, bushao nongmin kaishi yangzhi zhe zhong yang.*

'Because the income is handsome, quite a few peasants began to breed this type of sheep.'

- (36) *Zhe zhong huaiyi shi you yiju de, yinwei tongyang de changjing yijing chongfu le 18 ci.*

‘This kind of suspicion has its basis, **because** the same situation has already repeated itself eighteen times.’

4.4 Discussion and conclusion

In this study, we examined the subjectivity profiles of Mandarin reason connectives *jiran*, *yinwei*, and *youyu* (all can be translated as ‘because’) on the basis of the CCL and PPD corpora. It was found that *jiran*, *yinwei*, and *youyu* are associated with different degrees of subjectivity. *Jiran* expresses very subjective causal relations (i.e., epistemic and speech-act relations) involving judgments, speech acts, and implicit speaker SoCs. In contrast, *youyu* expresses very objective causal relations (i.e., non-volitional content relations) holding between facts, where no SoC is discerned. *Yinwei* seems to be an in-between category. On the one hand, language users prefer to use it to express subjective epistemic relations with implicit SoCs. On the other hand, *yinwei* also expresses a considerable number of content relations (124 out of 225 cases), which are considered to be relatively objective. This is different from *jiran*, which is rarely used in the objective content domain (only 11 cases). Therefore, we conclude that *yinwei* is comparatively general in use.

Because the *yinwei* fragments are a heterogeneous mixture of forward-linking and backward-linking uses, we analyzed the subjectivity profile per sequence order. The results show that the sentence-initial and inter-sentential *yinwei* express different degrees of subjectivity, with the latter being more subjective than the former. Sentence-initial *yinwei* is very similar to *youyu* in that, across genres, it is used to express content causal relations holding between facts, which involve no SoCs. Inter-sentential *yinwei* typically expresses epistemic causal relations involving judgments and implicit speaker SoCs.

The subjectivity profiles of all three reason connectives under investigation (including the subjectivity profile of sentence-initial *yinwei* and that of inter-sentential *yinwei*) are robust across genres. Accordingly, we can suppose that the degrees of subjectivity involved are inherent semantic characteristics of these Chinese reason connectives, and the pragmatics of genre plays little role in determining the values of their subjectivity. This is different from what we have found for the result connectives *suoyi* and *yinci* (both translated as ‘so, therefore’). The subjectivity orientation of *suoyi* and *yinci* changes across genres, which is reflected by the values of all four subjectivity indicators. *Suoyi* and *yinci* were found to be subjectively oriented in news reports and opinion pieces (i.e., more epistemic relations, judgments, and implicit author SoCs), but become objectively oriented in novels (i.e.,

more volitional content relations, facts, and explicit character SoCs). This observation reinforces our previous suggestion that *suoyi* and *yinci* are truly underspecified in terms of subjectivity (see Chapter 3). While all other causal connectives, including the three reason connectives, have the subjectivity feature specified in the lexicon, *suoyi* and *yinci* depend on the pragmatics of the context for the exact subjectivity value.

Apart from the depiction of subjectivity profiles for Mandarin reason connectives, another contribution of the current study concerns previous findings (in Song & Tao 2009) about the function of sentence-initial *yinwei* as an information-management device for narration and exposition, and the function of inter-sentential *yinwei* in mitigating and diverting potential sources of trouble as well as achieving affiliation. We believe that our findings about the subjectivity meanings of the two types of *yinwei* help to provide a valid explanation for the observed differences between them in terms of function. In our view, it is because inter-sentential *yinwei* has a highly subjective meaning that it is capable of performing the so-called interactional function. Namely, if the speaker wants to react to potential objections and defend her own point of view, her utterances tend to be subjective. It follows that she needs a highly subjective form to express herself in the utterance. Alternatively, if the speaker merely wants to share information with readers, the objectively-oriented sentence-initial *yinwei* will be a more appropriate option for performing this information-sharing function. From this we can see that meaning is an intermediate link between form-function pairs, and gives the reason for certain forms being able to perform the corresponding functions.

Studies on English, German, and Dutch have all reported interactions between the linearization of adverbial clauses and discourse functions (see Evers-Vermeul 2005: 35-37 for an overview). These studies are based on causal clauses as well as other adverbial clauses such as purpose clauses and conditional clauses. The basic idea is that initial and final adverbial clauses are not merely structural variants, but specialize in different contexts of usage. Our present study corroborates this idea.

Our final point concerns the combinatory use of reason and result connectives. See (31), repeated here as (37).

(37) *Yinwei* Ø *yijing sheji fubai, ta benren shi gongzhong renwu, yinci, dui ta de jiandu mei you renhe wenti.*

'Because (it) has involved corruption and he himself is a public figure, there is nothing wrong to keep him under supervision.'

The combinatory use, or *double marking*, is a language-specific property of Mandarin complex clauses, frequently referred to in the literature (Biq 1995; Xing 2001; Zhu 2009). There are suggestions for possible combinations, such

as the combination of *yinwei* and *suoyi*, *youyu* and *yin'er*, or *jiran* and *yinci* (Xing 2001). An interesting question to ask is whether there are restrictions that govern the ways in which reason and result connectives can be combined. One possible restriction might be related to the degree of subjectivity expressed by those causal connectives. Presumably, an objectively-oriented reason connective such as *youyu* might impose a restriction upon the degree of subjectivity of the upcoming result connective, requiring it to be underspecified or to be objective as well. This seems to be the case when we looked at the 45 double-marked cases in the current study. We have to admit that the combinatory use forms only 6.67 percent of the total number of fragments. Nevertheless, these combinations show a tendency for speakers to prefer to use the underspecified result connectives *suoyi* and *yinci* together with reason connectives. Out of the 45 double-marked cases, 26 are connected with *suoyi* and 12 are connected with *yinci*. The uses of other result connectives are rare (*yin'er*: 3, *yushi*: 4, *kejian*: 0). This finding seems to suggest that the match between degrees of subjectivity is a restriction on the way in which a reason connective can be combined with certain result connectives. The underspecified result connectives *suoyi* and *yinci* are preferred by speakers because they are general in use. They can be matched with both subjectively- and objectively-oriented reason connectives. Speakers do not need to spend extra time and energy on calculating the match between them and a reason connective in terms of subjectivity. A topic for future studies is to examine the total amount of subjectivity of the double-marked coherence relations: What does double-marking add to the subjectivity of the relation as a whole?

With the current study, we have provided a corpus-based analysis of the differences between Mandarin reason connectives in terms of subjectivity. The fact that Mandarin reason connectives express different subjectivity meanings seems to hint at the universality of subjectivity as a cross-linguistic cognitive principle in categorizing causal connectives. Meanwhile, when linguists discuss such lexical differences, they frequently raise the point that every language divides up the world, or reality, in its own way (see Lyons 1981, and see Stukker & Sanders 2012 for the way causal relations are categorized in various languages). Hopefully, the current study has managed to show how Mandarin reason connectives divide up the world of causality and subjectivity in its own way.

Notes

1. The frequencies for the other two reason connectives, *jiran* and *youyu*, are 0.72 and 2.38 per 10,000 words, respectively. The frequencies are calculated on the basis of Lancaster Corpus of Mandarin Chinese (McEnery & Xiao 2004), a balanced corpus representing modern Mandarin Chinese as written on the mainland of China. It

consists of five hundred 2,000-word samples taken from a wide range of text-types, including news reports, novels, and opinion pieces.

2. *People's Daily Online* is the most authoritative, comprehensive and influential website, claiming the largest daily amount of news releases in China. It has a database of 20 billion characters. It is an excellent data source for analyzing newspaper texts.

(http://search.people.com.cn/rmw/GB/rmwsearch/gj_search_pd.jsp)

3. *The CCL Corpus* was created by the Center for Chinese Linguistics of Peking University. It is as large as 1.06 GB, consisting of 477,000,000 characters. It is a good database of Chinese literary works, including *Novels*, *Prose*, *Biographies*, etc.

4. In the literature, the distinction between subjectivity and objectivity is considered to be gradual rather than absolute (Lyons 1982b: 105). Therefore, we did not calculate an exact value of subjectivity for the connectives.

5. In Chapter 3, a distinction was made within the category of speaker SoCs: the author vs. the current/quoted speaker. Cases such as (20) were considered to contain a current speaker SoC. In that study, the value of current speaker SoC did not truly reflect the genre-sensitivity of connectives, because it is genre-sensitive itself. The results showed that current speaker SoCs were distributed more frequently in novels than in news reports and opinion pieces. For this reason, in the present study we did not code current speaker as a separate category.

Appendix 1. The analysis of *domain*

A. Results of General Log-linear analysis of *domain*

Model	χ^2 (model)	<i>df</i>	<i>p</i> (model)	χ^2 (factor)	<i>df</i>	<i>p</i> (factor)
1. constant+conn.	476.34	9	<.001			
+ 2. domain	378.14	6	<.001	98.20	3	<.001
+ 3. domain*conn.	.00	0	1.000	378.14	6	<.001

B. Parameter estimates *domain* for Model 3

Parameter	Estimate	Std. Error	<i>z</i>	<i>p</i>
constant	4.241	.120	35.359	<.001
[conn. = youyu]	.704	.147	4.800	<.001
[conn. = jiran]	-2.737	.486	-5.627	<.001
[domain = volitional content]	-.225	.180	-1.250	.211
[domain = epistemic]	.369	.156	2.364	.018
[domain = speech act]	-3.836	.825	-4.648	<.001
[conn. = youyu] * [domain = vol. content]	-.948	.250	-3.790	<.001
[conn. = youyu] * [domain = epistemic]	-1.565	.234	-6.672	<.001
[conn. = youyu] * [domain = speech act]	-1.802	1.639	-1.100	.271
[conn. = jiran] * [domain = vol. content]	.736	.623	1.181	.237
[conn. = jiran] * [domain = epistemic]	2.975	.504	5.899	<.001
[conn. = jiran] * [domain = speech act]	6.803	.956	7.114	<.001

Appendix 2. The analysis of *propositional attitude*

A. Results of General Log-linear analysis of *propositional attitude* (of the consequent, *Q*)

Model	χ^2 (model)	<i>df</i>	<i>p</i> (model)	χ^2 (factor)	<i>df</i>	<i>p</i> (factor)
1. constant+conn.	450.99	6	<.001			
+ 2. Q	356.28	4	<.001	94.71	2	<.001
+ 3. Q*conn.	.00	0	1.000	356.28	4	<.001

B. Parameter estimates *propositional attitude* for Model 3

Parameter	Estimate	Std. Error	<i>z</i>	<i>p</i>
constant	4.824	.090	53.829	<.001
[conn. = youyu]	.388	.116	3.341	.001
[conn. = jiran]	-2.382	.308	-7.729	<.001
[Q = judgment]	-.214	.134	-1.597	.110
[Q = speech act]	-4.419	.821	-5.380	<.001
[conn. = youyu] * [Q = judgment]	-1.249	.217	-5.762	<.001
[conn. = youyu] * [Q = speech act]	-1.487	1.637	-.908	.364
[conn. = jiran] * [Q = judgment]	2.620	.336	7.801	<.001
[conn. = jiran] * [Q = speech act]	6.448	.879	7.334	<.001

Appendix 3. The analysis of *presence of SoC*A. Results of General Log-linear analysis of *presence of SoC*

Model	χ^2 (model)	df	p	χ^2 (factor)	df	p (factor)
1. constant+conn.	480.48	24	< .001			
+ 2. genre	480.48	22	< .001	0	2	1.000
+ 3. presence_SoC	296.04	20	<.001	184.44	2	< .001
+ 4. presence_SoC*conn.	21.95	16	.145	274.09	4	< .001
+ 5. presence_SoC*genre	11.99	12	.446	9.96	4	.041

B. Parameter estimates *presence of SoC* for Model 5

Parameter	Estimate	Std. Error	z	p
constant	3.536	.124	28.460	< .001
[conn. = youyu]	-.633	.173	-3.650	< .001
[conn. = jiran]	.788	.123	6.397	< .001
[genre = narra.]	-.237	.133	-1.780	.075
[genre = argu.]	.008	.125	.062	.950
[presence = no SoC]	-.458	.200	-2.287	.022
[presence = explicit Soc]	-.685	.236	-2.898	.004
[conn. = youyu] * [presence = no SoC]	1.340	.227	5.896	< .001
[conn. = youyu] * [presence = explicit]	.065	.276	.234	.815
[conn. = jiran] * [presence = no SoC]	-3.635	.529	-6.874	< .001
[conn. = jiran] * [presence = explicit]	-2.579	.363	-7.104	< .001
[genre = narra.] * [presence = no SoC]	.402	.213	1.887	.059
[genre = narra.] * [presence = explicit]	.642	.271	2.373	.018
[genre = argu.] * [presence = no SoC]	-.008	.213	-.037	.971
[genre = argu.] * [presence = explicit]	-.042	.289	-.144	.885

Appendix 4. The analysis of *identity of SoC*A. Results of General Log-linear analysis of *identity of SoC*

Model	χ^2 (model)	df	p (model)	χ^2 (factor)	df	p (factor)
1. constant+conn.	212.73	15	< .001			
+ 2. identity_SoC	101.56	14	< .001	111.17	1	< .001
+ 3. identity_SoC*conn.	14.02	12	.299	87.54	2	< .001

B. Parameter estimates *identity of SoC* for Model 3

Parameter	Estimate	Std. Error	z	p
constant	3.526	.099	35.614	<.001
[conn. = youyu]	-.864	.182	-4.751	<.001
[conn. = jiran]	.736	.120	6.115	<.001
[identity_SoC = character]	-.636	.168	-3.779	<.001
[conn. = youyu] * [identity = character]	.612	.275	2.231	.026
[conn. = jiran] * [identity = character]	-2.646	.398	-6.656	<.001

Appendix 5. The analysis of *Domain*

A. Results of General Log-linear analysis of *Domain*

Model	χ^2 (model)	df	p (model)	χ^2 (factor)	df	p (factor)
1. constant+position	66.97	4	<.001			
+ 2. domain	47.43	2	<.001	19.54	2	<.001
+ 3. domain*position	.00	0	1.000	47.43	2	<.001

B. Parameter estimates *domain* for Model 3

Parameter	Estimate	Std. Error	z	p
constant	3.418	.181	18.875	<.001
[position = sentence-initial]	.259	.241	1.073	.283
[domain = volitional content]	-.397	.286	-1.391	.164
[domain = epistemic/speech act]	1.042	.211	4.950	<.001
[position = initial] * [domain = vol.]	.291	.367	.791	.429
[position = inter-sentential] * [domain = epistemic/speech act]	-1.978	.366	-5.400	<.001

Appendix 6. The analysis of *propositional attitude*

A. Results of General Log-linear analysis of *propositional attitude* (of the consequent, Q)

Model	χ^2 (model)	df	p (model)	χ^2 (factor)	df	p (factor)
1. constant+position	48.64	2	<.001			
+ 2. Q	46.78	1	<.001	1.86	1	.173
+ 3. position*Q	.00	0	1.000	46.78	1	<.001

B. Parameter estimates *propositional attitude* for Model 3

Parameter	Estimate	Std. Error	z	p
constant	4.460	.108	41.482	<.001
[position = sentence-initial]	-1.719	.276	-6.234	<.001
[Q = fact]	-.538	.177	-3.039	.002
[position = sentence-initial] * [Q = fact]	2.108	.331	6.377	<.001

Appendix 7. The analysis of *presence of SoC*A. Results of General Log-linear analysis of *presence of SoC*

Model	χ^2 (model)	df	p (model)	χ^2 (factor)	df	p (factor)
1. constant+position	25.00	4	<.001			
+ 2. presence	14.56	2	.001	10.44	2	.005
+ 3. position*presence	.00	0	1.000	14.56	2	.001

B. Parameter estimates *presence of SoC* for Model 3

Parameter	Estimate	Std. Error	z	p
constant	4.256	.119	35.738	<.001
[position = sentence-initial]	-.978	.228	-4.294	<.001
[presence = no SoC]	-.838	.217	-3.866	<.001
[presence = explicit SoC]	-.658	.204	-3.228	.001
[position = initial] * [presence = no SoC]	1.237	.332	3.730	<.001
[position = initial] * [presence = exp.]	.580	.347	1.673	.094

Appendix 8. The analysis of *identity of SoC*A. Results of General Log-linear analysis of *identity of SoC*

Model	χ^2 (model)	df	p (model)	χ^2 (factor)	df	p (factor)
1. constant+position	33.74	2	< .001			
+ 2. identity_SoC	20.95	1	<.001	12.79	1	< .001
+ 3. position*identity_SoC	.00	0	1.000	20.95	1	< .001

B. Parameter estimates *identity of SoC* for Model 3

Parameter	Estimate	Std. Error	z	p
constant	4.413	.110	40.081	<.001
[position = sentence-initial]	-1.392	.247	-5.642	<.001
[identity_SoC = character]	-1.214	.230	-5.277	<.001
[position = initial] * [identity= character]	1.611	.367	4.394	<.001

CHAPTER 5

On the online effects of subjectivity information encoded in causal connectives

5.1 Introduction

5.1.1 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 and Thompson 1986, 1988; Sanders, Spooren & Noordman 1993). Coherence relations (e.g., cause-consequence, additive, and concessive relations) are conceptual relations that hold between the propositional content of two discourse segments (see Sanders et al. 1993), which can (but need not) be explicitly marked by connectives such as *because*, *and*, and *but*. A causal coherence relation is a typical example of coherence relations. Consider sentences (1) – (4).

- (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. They probably left.
- (3) Many kingfishers died last year. Temperatures were below minus ten degrees for more than a month.
- (4) The neighbors probably left. Their car is not in the driveway.

In order to understand these sentences, readers need to interpret the coherence relations contained in them. The coherence relations expressed in (1) through (4) have something in common: causality that holds between an antecedent segment P and a consequent segment Q. Still, several aspects can be distinguished. One apparent distinction concerns the causal order; that is, the sequencing of P and Q. In both (1) and (2), the antecedent is expressed first and the consequent second. In both (3) and (4), the order is reversed: the consequent is placed in front of the antecedent. In the literature, the former causal order is referred to as ‘the basic order’ and the latter causal order is termed ‘the non-basic order’ (Sanders, Spooren & Noordman 1992).¹ The

¹ There might be two interpretations for the causal relation involved in (2) and (4). The first interpretation is as follows. The fact that “the neighbors’ car is not in the driveway” can be taken as a premise which leads to the conclusion that “they probably left”. In this case, there is a basic order in (2) and a non-basic order in (4). The second

basic order, at least for cause-consequence relations such as in (1), iconically reflects the order of the events in the real world.

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. Sentences (1) and (3) exemplify causal relations in the content domain; that is, the “real-world causality” holding between observable facts/events (p.77) – the observable fact that the temperatures were below minus ten degrees for more than a month led to the other observable fact that many kingfishers died. Sentences (2) and (4) illustrate causal relations in the epistemic domain, in which the speaker’s knowledge (e.g., *the neighbors’ car is not in the driveway*) is involved as the basis for a logical conclusion (e.g., *they probably left*). The conceptual differences between (1) versus (2) and between (3) versus (4) are also referred to as external versus internal (Halliday & Hasan 1976; Martin 1992), subject matter versus presentational (Mann & Thompson 1988), and semantic versus pragmatic (van Dijk 1979; Moeschler 1989; Sanders, Spooren & Noordman 1992; Sanders 1997).

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 & Verhagen 2008; Stukker & Sanders 2010b, 2012). 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, as in (1) and (3), are assumed to be objective, because they do not involve the speaker as a conscious subject. The speaker merely reports causality that has an origin in the natural world. By contrast, casual relations in the epistemic domain, as in (2) and (4), 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). In the current study, we distinguish between subjective and objective causal relations. In particular, epistemic relations and content

interpretation is: the judgment that “the neighbors probably left” can be taken as a premise which gives rise to a further judgment that “their car is not in the driveway”. In this case, there is a non-basic order in (2) and a basic order in (4). We will discuss later that this kind of ambiguity can be easily resolved by the insertion of a causal connective. For example, the second interpretation is the only one if we use *because* between the clauses in (2), but this interpretation is ruled out if we use *so* instead.

relations are selected as the targets of study because they typically represent the subjective causal relation and the objective causal relation, respectively.²

5.1.2 Causal connectives and subjectivity in discourse processing

Previous experimental studies on reading comprehension have shown that causal coherence relations are constructed online during discourse processing (Gernsbacher 1990; Graesser, Singer & Trabasso 1994; Singer, Revlin & Halldorson 1990). This paper focuses on the way in which readers construct different types of causal coherence relations online; in particular, how this process is guided by different causal connectives, which are explicit linguistic markers of causal relations between adjacent segments (Murray 1997; Stukker & Sanders 2012).

Causal connectives help readers to identify that the coherence relation in question is causal rather than additive or temporal. For example,

- (5) The snow began to melt gradually **when** the temperature rose.
- (6) The snow began to melt gradually **because** the temperature rose.
- (7) The snow began to melt gradually, **so** the temperature rose.

The temporal connective *when* in (5) only indicates that the two pieces of text are temporally related to each other, so the implicit causality has to be figured out on the basis of the propositional content of the clauses. The causal connective *because*, however, directly makes this causal relation explicit (see (6)). Moreover, causal connectives assist readers in ascertaining the way in which the antecedent and the consequent are ordered. *Because*, as a reason connective, explicitly indicates that the clause containing it is the antecedent segment of the causal relation, which is the second clause in (6). Readers will establish the causal relation in the non-basic order, accordingly, from the second clause “the temperature rose”, backward to the first clause “the snow began to melt gradually” (i.e., high temperature caused snow to melt). *So* specifies a linear order in which the antecedent precedes the consequent, as most result connectives do. If *so* is used (see (7)), readers will accordingly build up a causal relation in the basic order. That is, in (7), readers will establish the causal relation from the first clause forward to the second clause (i.e., the fact that *the snow began to melt gradually* leads to the

² According to Sanders and Spooren (2009), the speech-act domain is also associated with a high degree of subjectivity, but it rarely occurs in written discourse (Spooren, Sanders, Huiskes & Degand 2010; Li, Evers-Vermeul & Sanders 2013). Since the present experiment uses written materials, we do not include the speech-act domain as a target of study.

conclusion that *the temperature rose*). Without the use of a connective, the two causal interpretations are both acceptable. In other words, the use of a connective can to some extent disambiguate between candidate interpretations of a coherence relation.

From the perspective of psycholinguistics, connectives can be 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). This idea is in fact quite in line with one of the major functions of connectives identified by Noordman and Vonk (1997): the function of connectives as an integration device. Integration refers to the construction of the two connected text segments into a single representation, and is generally assumed to occur online during processing (Kintsch 1988; Millis & Just 1994). According to Noordman and Vonk (1997), the connective specifies to the reader that a specific coherence relation holds between the clauses, and instructs the reader to integrate the clauses by means of that relation. On the basis of these ideas, it can be predicted that connectives facilitate the processing of subsequent information when they are properly used. For example, when a connective encodes causality, it provides readers with a processing instruction that a causal relation, rather than an additive or other coherence relation, needs to be established between the two pieces of text. By providing such a concrete, disambiguating instruction, a causal connective is likely to facilitate the integration of the two conjoint segments into a coherent mental representation, compared to the same relation without a connective. Indeed, several online reading studies have shown that the presence of an appropriately-used causal connective speeds up the processing of the immediately following discourse segment: participants spent less time reading a clause when it was introduced by a causal connective than when the connective was absent (see Cozijn 2000; Haberlandt 1982; Millis & Just 1994).

As will become clear next, causal connectives can 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 suggest that some causal connectives carry a higher degree of subjectivity than others, whereas there are also causal connectives that are not encoded with any information regarding the degree of subjectivity (see Chapter 3 and Dancygier 2009; Evers-Vermeul, Degand, Fagard, & Mortier 2011; Keller 1995; Pander Maat & Sanders 2001; Pit 2003; Zufferey 2012). That is, some causal connectives specify the degree of subjectivity the relation is involved with, whereas other causal connectives have a more general usage. In this study, we refer to the former as ‘specified causal connectives’, and the latter as ‘underspecified causal connectives’. The

Dutch reason connectives *want* 'because' and *omdat* 'because' exemplify specified causal connectives. *Want* is associated with a high degree of subjectivity whereas *omdat* is associated with a high degree of objectivity (Degand & Pander Maat 2003; Pit 2003; Sanders & Spooren submitted; Sanders, Sanders & Sweetser 2012; Verhagen 2005). Specified causal connectives restrict the possible interpretations of the causal relation by specifying their degree of subjectivity. Underspecified causal connectives, however, do not impose any restrictions on the degree of subjectivity of the causal relation. The English *because* is an instance of underspecified causal connectives (see Couper-Kuhlen 1996; Ford 1993; Knott & Sanders 1998; Sweetser 1990). According to Knott & Sanders (1998), *because* is undefined for the parameter of 'source of coherence'; i.e., with respect to the distinction between semantic and pragmatic causality.

Subjective causal relations have been argued to be more complex than objective causal relations at the cognitive level – the so-called 'subjective complexity hypothesis' (Sanders 2005). This 'subjective complexity' has been demonstrated from different perspectives in the literature (see Noordman & De Blijzer 2000; Sanders 2005). Diachronically, subjective causal relations are derived from objective causal relations (Sweetser 1990; Traugott 1995). Furthermore, epistemic reasoning is often based on real-world content relations (Noordman & De Blijzer 2000; Pander Maat & Degand 2001; Sanders 2005; Spooren & Sanders 2005). In addition, subjective causal relations are assumed to contain more information than objective relations, such as a meta-representation of others' beliefs and conclusions (Canestrelli et al. 2013).

Empirical evidence in support of this subjective complexity hypothesis can be found in acquisition studies on the development of causal relations/connectives. A consistent finding 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.

Reading-time experiments have shown that the degree of subjectivity affects the online processing of causal relations. One line of study focuses on the processing of content versus epistemic causal relations, both marked with the underspecified causal connective *because* (Traxler, Bybee & Pickering 1997a; Traxler, Sanford, Aked & Moxey 1997b). The other line of study uses Dutch materials to explore the ways in which subjective and objective causal connectives respectively guide the online comprehension of the corresponding epistemic and content causal relations (Canestrelli 2013;

Canestrelli, Mak & Sanders 2013; Klein, Mak & Sanders 2011). We introduce the two lines of study in detail below.

The experiment by Traxler et al. (1997a) represents the former line of study. It revealed the way in which content causal relations (in their terms ‘simple-causal relations’) and epistemic causal relations (in their terms ‘diagnostic relations’) are processed when the causal connective provides no information with respect to the degree of subjectivity. English sentence-pairs such as (8) and (9) are used in their study.

- (8) The goal keeper won the game **because** the other team didn’t score any goals.
- (9) The goal keeper knew how to play the game **because** the other team didn’t score any goals.

Within each pair, one sentence (8) expresses a causal relation in the content domain and the other sentence (9) expresses a causal relation in the epistemic domain. For both cases, the relations are expressed in the non-basic order, and *because* is used to introduce the antecedent clause backward to the preceding consequent clause. As an underspecified connective, *because* does not specify the degree of subjectivity of the causal relation. 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 (9) expresses a subjective epistemic relation. The results have shown that, precisely at the predicate position of the second clause, a processing delay occurs for epistemic compared to content relations. It can be concluded that the longer reading times at the predicate of epistemic relations reflect, in essence, that the relation’s degree of subjectivity directly affects online processing. Following the subjective complexity hypothesis, comprehending subjective causal relations probably requires more cognitive effort than comprehending objective causal relations, which is why processing the former takes longer than processing the latter.

Traxler et al.’s (1997a) experiment demonstrates that when underspecified causal connectives are used, the degree of subjectivity encoded in the unfolding content of the relation has an immediate effect on processing. What if the causal connectives already specify the degree of subjectivity of the causal relation they mark? Canestrelli et al. (2013) (which represents the second line of study mentioned earlier) have compared reading times of Dutch content and epistemic causal relations, respectively marked by *omdat* and *want* in the non-basic order, basing their experimental materials on those in Traxler et al. (1997a). As specified causal connectives, objective *omdat* and subjective *want* impose restrictions on the degree of subjectivity of the relation. Accordingly, they provide information as to whether a content or an epistemic causal relation is likely to follow. Again,

the eye-tracking results indicated a processing delay for epistemic relations: the *want* condition generated longer processing times than the *omdat* condition. However, the 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, immediately after the connective, and even at the connective itself, but not at the predicate region (Canestrelli 2013; Canestrelli et al. 2013). Canestrelli et al. (2013) explained this by claiming that the epistemic relation was not constructed on the basis of the content of the clauses, but on the basis of the prior information provided by the specified causal connective. *Want* directly triggers the representation of a subjective relation whereas *omdat* immediately triggers the representation of an objective relation. Hence, the processing difference between the two conditions was observed as early as at the connective region. In the meantime, this explanation also fits another result observed in their study: longer reading times were observed at the end of the clause in the *omdat*-condition compared to in the *want*-condition, when both connectives were placed in a subjective relation that does not allow for an objective interpretation. That implies: *omdat* triggered the representation of an objective relation before the propositional content of the second clause unfolded; and this led to a processing difficulty at the end of the second clause, where it turned out that the relation was actually subjective.

In another Dutch experiment by Kleijn, Mak, and Sanders (2011), the same immediate processing difference (as the one observed in Canestrelli et al. 2013 for *want* and *omdat*) is observed between the objective connective *daardoor* ‘as a result’ and the subjective connective *dus* ‘so’. The materials consisted of Dutch sentences expressing content and epistemic relations in the basic order, in the form of “P, as a result/so Q”. Given that at or immediately after the connectives there was no information on the type of causal relation on the basis of the semantic content of the sentence, the observed processing asymmetry between the subjective *dus* condition and the objective *daardoor* condition must be due to the properties of the connectives themselves, to be precise, due to the high/low degree of subjectivity encoded in these connectives. In light of the findings from the above two Dutch studies, it is apparent that the degree of subjectivity encoded in causal connectives has an immediate effect on processing, no matter in which order P and Q are presented.

Taken together, the English and the Dutch processing studies indicate that the processing asymmetry between the objective and subjective conditions occurs at different positions in the sentence, depending on the specificity of the connective with respect to subjectivity. The processing difference occurred at or immediately after the connective when the specified causal connectives *omdat* and *want* were used, whereas it occurred at the predicate position (at which the bulk of the two propositions P and Q were

available) when the underspecified causal connective *because* was used. These results suggest that, in causally-related discourse, two sources of subjectivity information – the propositional content of the utterance itself (i.e., the contents of P and Q) and the specified causal connectives that connect P and Q – affect online processing, albeit at different stages of the reading process. More interestingly, as reported earlier, when the processing delay associated with the epistemic relation occurred earlier (i.e., at or immediately after the specified connective *want*), it did not occur at the predicate region as it did in conditions where the underspecified connective *because* was used.

However, the above claims about the processing consequences of the specificity of connectives with respect to subjectivity are only tentative, because they 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 only looked at languages with either an underspecified causal connective (i.e., *because*) or only specified causal connectives (i.e., *omdat* and *want*). Hence, the main question of the present research is the following. What does the degree of subjectivity encoded in specified causal connectives add to the online interpretation of causal relations compared to those marked with underspecified causal connectives? For example, is there a processing difference between an underspecified causal connective versus an objective causal connective when they are used to express causal relations in the content domain? Or do readers treat these types of connectives in the same way, and only process the information differently if they encounter a subjective connective? In other words, is the objective relation the default interpretation, as Traxler et al. (1997b) have suggested, or does the objective connective also facilitate the processing of objective relations? In order to address these issues, we need to set up a new eye-tracking experiment, which simultaneously takes both specified and underspecified causal connectives within the same language into consideration in the experimental design. A language with both specified and underspecified causal connectives could lead to clarifying results. As shown in Table 1, the Chinese language offers such a possibility whereas the English and the Dutch languages do not.

Table 1. A summary of the specified and underspecified causal connectives in English, Dutch, and Mandarin Chinese

Causal connective	English		Dutch		Mandarin Chinese	
	Reason	Result	Reason	Result	Reason	Result
Objective			<i>doordat</i> <i>omdat</i>	<i>daardoor</i> <i>daarom</i>	<i>youyu</i> <i>yinwei</i> (basic order)	<i>yin'er</i> <i>yushi</i>
Subjective	<i>since</i>	<i>so</i>	<i>want</i>	<i>dus</i>	<i>jiran</i> <i>yinwei</i> (non-basic order)	<i>kejian</i>
Underspecified	<i>because</i>					<i>suoyi</i> <i>yinci</i>

For English, no objective causal connective has been identified. Existing analyses suggest that demarcation between subjective and objective categories is realized by cue phrases rather than connectives (Knott & Dale 1994; Knott & Sanders 1998): “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. We choose to use *kejian*, *suoyi*, and *yin'er* (all translated as *so/therefore*), because they represent typical instances of subjective, underspecified, and objective causal connectives, respectively (see Chapter 3). In the next section, we introduce the usage of these connectives in more detail and, on the basis of that, we present our predictions regarding how the differences between causal connectives with respect to their specificity for subjectivity affect the way in which subjective and objective causal relations are processed online.

5.2 The causal connectives to be studied and predictions

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

Causal connectives can differ with respect to the degree of subjectivity they encode (Sanders & Sweetser 2009). In Mandarin Chinese, *kejian* is highly subjective whereas *yin'er* is highly objective, and *suoyi* is underspecified with respect to subjectivity (see Chapter 3). The corpus fragments (10) – (13) exemplify typical uses of the three connectives. The usage of *kejian* is restricted to subjective causal relations, such as in (12).³ *Kejian* is never used in content causal relations, such as in (10) and (11). *Yin'er* is typically used to

³ 221 out of 225 cases of *kejian* are used in the epistemic domain. The other four cases were used to mark the speech-act relation, which is also considered a highly subjective causal category.

mark content causal relations (62%) between facts in observable reality, such as in (10). When the two specified connectives *yin'er* and *kejian* are used, it is very likely that readers can predict whether the unfolding relation is going to be objective or subjective on the basis of their knowledge about the usage patterns of the two connectives. In this sense, we assume that specified causal connectives indicate the degree of subjectivity of the relation. The underspecified causal connective *suoyi* has a very flexible usage, and occurs in content relations like (11) as frequently as in epistemic relations like (13). When readers encounter *suoyi*, it is not possible for them to make usage-based predictions regarding the degree of subjectivity involved. In other words, *suoyi* does not provide information with respect to the degree of subjectivity of the relation.

- (10) *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.'

- (11) *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.'

- (12) *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.'

- (13) *Ta shou shang daile hao ji ge jiezh, suoyi wo yiwei ta jiehun le.*

'He wore quite a few rings on the hand, **so** I thought he was married.'

For the subjective connective *kejian*, it can be clearly seen even from its word formation that the word itself is encoded with the speaker/writer's perspective. *Kejian* consists of two morphemes. The morpheme *ke* 'can' concerns epistemic modality, and hence involves the speaker/writer's or another person's perspective on and assessment of the plausibility or truthfulness of a certain state of affairs. The other morpheme *jian* 'see' has been demonstrated as a subjectivity indicator, which encodes perspective, hence the involvement of the speaker/writer (see Tao 2007 for evidence that *jian* indicates subjectivity in the context of existential/presentative constructions). A point to note here is that *kejian* encodes only the speaker/writer's perspective, and not another person's perspective. Corpus-based analyses have revealed that *kejian* is exclusively used to introduce the speaker/writer's beliefs or conclusions (see Chapter 3). For both the objective

connective *yin'er* and the underspecified connective *suoyi*, however, the words themselves never refer to the speaker's or another person's perspective.

Apart from the differences in terms 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 only used in the basic order (i.e., P precedes Q). In terms of parts of speech, 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.

5.2.2 Predictions

As we have put forward in the introduction, the present study is intended to increase our knowledge on the way in which the degree of subjectivity encoded in causal connectives affects online discourse processing. In particular, we want to explore exactly how specified causal connectives differ from underspecified causal connectives in guiding the processing of causally-related texts. Will a subjective/objective causal connective facilitate the online interpretation of epistemic/content causal relations compared to an underspecified causal connective? The English version of one set of our test items, (14) through (17), will suffice here for the purpose of illustrating our predictions. A full explanation of the materials and the design will be given in the next section.

- (14) My bike ran over a nail when I rode on it, *yin'er* the tire leaked.
- (15) My bike ran over a nail when I rode on it, *suoyi* the tire leaked.
- (16) The tire became flat immediately after I pumped it up, *kejian* the tire leaked.
- (17) The tire became flat immediately after I pumped it up, *suoyi* the tire leaked.

Sentences (14) and (15) 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 (16) and (17) 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). The connective use is appropriate for each sentence. With such a design, we seek any processing differences between (14) and (15), and between (16) and (17). For the latter comparison between the epistemic conditions, we predict that *kejian* will

facilitate the reading of the second clause, because we assume that, as a specialized subjective causal connective, *kejian* immediately guides readers to construct the causal relation in the subjective direction. In contrast, the underspecified connective *suoyi* provides no guidance or instruction in this respect. Whether the relation is to be represented as an objective connection existing in the outside world or as a subjective construal in one's mind is determined by the propositional content of the clauses. In that case, the reader probably starts to construct the epistemic relation at the predicate of the second clause (i.e., leaked), which will probably lead to a slowdown in reading compared to the sentence with *kejian*. Hence, we predict that at the predicate of the second clause, epistemic relations with *kejian* (e.g., (16)) are read faster than epistemic relations with *suoyi* (e.g., (17)). However, the high degree of subjectivity encoded in *kejian* itself may well induce an immediate slowdown in reading, as it does for Dutch *want*, since a high degree of subjectivity has been shown to be cognitively complex. Given that different lexical items are involved and the frequency of the connectives is different, we will not focus on the effects only at the connective. Instead, the effects at the connective might be meaningful when we combine them with the effects at the predicate of the second clause. We will pick up this point again in section 5.5.

For the comparison between the content conditions (14) and (15), a possible prediction is that *yin'er* speeds up the processing of the predicate of the second clause (i.e., leaked), the position at which the reader generally starts to construct the causal relation when the causal connective is underspecified (i.e., *suoyi*). As an objective causal connective, *yin'er* is likely to immediately instruct and prepare the reader to construct an objective causal relation, and in this way facilitate the processing of the content causal relation. Alternatively, perhaps no processing differences will arise between (14) and (15) throughout the sentences. We can assume that people might generally expect to construct an objective interpretation anyway, on the basis of the claim proposed by Traxler et al. (1997b) that readers might prefer to build the simplest possible discourse representation during interpretation. In this case, *yin'er* will not facilitate the processing of the content relation compared to *suoyi*.

In addition, we expect to replicate the result in Traxler et al. (1997a). We predict that, when both the content relation and the epistemic relation are marked with the underspecified connective *suoyi* ((15) vs. (17)), a processing asymmetry will occur between these two conditions at the predicate of the second clause.

5.3 Method

An eye-tracking experiment was conducted. Participants' eye movements, such as saccades and fixations, were measured and recorded, on the basis of which reading times were computed for regions of interest throughout the materials. Subsequently, we assessed the reading times of these regions across experimental conditions.

5.3.1 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.

5.3.2 Materials and design

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

Each set consisted of four versions, and each version of the story was embedded in a different experimental condition. There were altogether four experimental conditions: content causal relations marked by the objective causal connective *yin'er*, content causal relations marked by the underspecified causal connective *suoyi*, epistemic causal relations marked by the subjective causal connective *kejian*, and epistemic causal relations marked by the underspecified causal connective *suoyi*.

There were two critical manipulations. Firstly, within each set of stories, the first clauses were manipulated. As a result of the manipulation, the relation between the first and second clause alternated between content and epistemic causal relations. The second manipulation concerned the choice of connective. Either a specified or an underspecified causal connective was used to explicitly mark the inter-clausal relation. 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*. The selection of connectives was based on results from the previous corpus-based study (Chapter 3) introduced earlier.

To make sure that the constructed causal relations were without ambiguity (i.e., they don't simultaneously 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". As is apparent from its name, 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. The participants' response data were analyzed with SPSS. We first took the

means of those scores by items, and then calculated the item's mean for each condition. Table 3 summarizes the means and the standard deviations for the four conditions.

Table 3. Mean scores and standard deviations of the rating for the paraphrases of the constructed causal relations

Causal relation	Paraphrase	60 sets	40 sets
Objective	1. objective	4.07 (0.49)	4.13 (0.40)
	2. subjective	2.33 (0.57)	2.23 (0.46)
Subjective	3. objective	2.42 (0.62)	2.30 (0.55)
	4. subjective	4.09 (0.61)	4.30 (0.37)

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 stimuli for the experiment. A repeated measures ANOVA was then conducted on the item's means. We found a significant interaction between the type of the causal relation and the type of the paraphrase ($F_2(1, 39) = 980.66, p < 0.001$). In order to figure out exactly how the conditions differ from each other in terms of mean score, a series of pairwise comparisons were conducted across the four conditions, using a Bonferroni adjustment. Pairwise comparisons indicated that neither the correct conditions 1&4 differed ($p = 0.356$), nor the incorrect conditions 2&3 ($p = 1.000$); meanwhile, the correct conditions scored significantly higher than the incorrect conditions (1 vs. 2: $p < 0.001$; 4 vs. 3: $p < 0.001$). 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.

The selected stimuli, 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.

5.3.3 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 practice items, the real test started. Each participant read 105 texts in the experiment, which took about 25 minutes on average.

5.3.4 Analysis

The second clause of each text was the target of our analysis, which was specified to contain a connective region, a subject region, and a predicate region. In Table 4, we illustrate the way in which a target clause was split into these three regions.

Table 4. The division between three target regions

	Connective region	Subject region	Predicate region
Chinese item:	<i>suoyi/yin'er/kejian</i>	<i>ta bi yiqian</i>	<i>shou le bushao.</i>
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. Every target clause was divided as such, because we predicted separate effects for the connective, the words immediately following the connective on the one hand, and the predicate on the other hand. It should be noted that the effect at the connective region is perhaps difficult to interpret because different lexical items are involved, and the frequency of the connectives varies. Yet, the effects at the connective might be meaningful when we combine them with the effects at the predicate. For this reason, we coded the region containing the connective as one of the critical regions. The three critical regions were displayed near the center of the screen during the experiment, in order to prevent ‘edge effects’ (Koornneef & van Berkum 2006; Rayner 1998).⁴

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. Further clean-up was done as follows. Any observations 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.

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.

Following Traxler et al. (1997a), we focused on four reading-time measures. Three of them concerned first pass reading. *First pass reading time* (FP) is the total reading time (including fixation 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

⁴ 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’.

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), is related not only to 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.

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.⁵ We used both *subject* and *item* as random effects (see 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.

5.4 Results

In Table 5, we present the mean reading times and standard deviations for each reading-time measure and for each region under investigation. In the subsections following this table, we first report significant effects for the first pass measures, then significant effects observed from total fixation duration. The latter measure captures the time required for reanalysis when a piece of information is not completely processed during the first reading (Rayner & Stereno 1994). It contains rereading times of a certain region after the region is left in a progressive manner. Therefore, effects observed for total fixation duration at a certain region relate to both the words within and the words after the region.

⁵ 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.

Table 5. Mean reading times and standard deviations in terms of FP, TG, RP, and TF (in ms)

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

5.4.1 First pass reading

The connective region contains only a specified/underspecified causal connective, and the content of the relation is not yet available there. Significant effects (if any) cannot be due to the relation, but must be due to the properties of the connective by itself. At this region, the *first pass reading time* presented an interaction effect between relation and connective ($\beta = -0.13$, $SE = 0.04$, $t = -3.03$, $p = 0.003$). In order to pinpoint the direction of the interaction, firstly a pairwise comparison was performed between the content condition with *suoyi* and the epistemic condition with *suoyi*. No differences were observed between these two conditions ($\beta = 0.04$, $SE = 0.03$, $t = 1.65$, $p = 0.099$). The pairwise comparison between the *yin'er* and *kejian* cases revealed that the former took longer to read than the latter ($\beta = -0.10$, $SE = 0.04$, $t = -2.84$, $p = 0.009$). The pairwise comparison between the two epistemic conditions showed no difference between *kejian* and *suoyi* ($\beta = 0.01$, $SE = 0.03$, $t = 0.51$, $p = 0.602$), whereas the pairwise comparison between the two content conditions showed that *yin'er* resulted in longer reading times than *suoyi* ($\beta = 0.15$, $SE = 0.33$, $t = 4.40$, $p < 0.001$).

Two other reading-time measures, *first pass total gaze duration* and *regression path duration*, presented a main effect of connective: the specified connectives *yin'er* and *kejian* were read more slowly than the underspecified

connective *suoyi* (TG: $\beta = 0.08$, $SE = 0.02$, $t = 3.58$, $p < 0.001$; RP: $\beta = 0.07$, $SE = 0.03$, $t = 2.90$, $p = 0.005$). No other effects were observed at this region for TG and RP.

The subject region contains the subject of the second clause, and sometimes also includes an adverbial. This is a region at which we might observe a processing difficulty associated with *kejian*. In the Dutch experiments, it was found that sometimes the processing difficulty induced by the subjective connective *want* spilled over to this region. In the current experiment, however, no significant effects were observed at this region for the first pass reading-time measures under investigation.

At the predicate region, which is also the final region of the second clause, the actual content of the clause is fully available. Here we may expect that specified causal connectives in general speed up the processing of the corresponding epistemic/content causal relation. Or else, only the subjective connective *kejian* plays such a facilitating effect, whereas the objective connective *yin'er* does not; if people might generally expect to build up a content causal relation anyway.

An interaction effect between relation and connective was observed (FP: $\beta = -0.13$, $SE = 0.05$, $t = -2.39$, $p = 0.016$; TG: $\beta = -0.13$, $SE = 0.05$, $t = -2.63$, $p = 0.010$; RP: $\beta = -0.17$, $SE = 0.06$, $t = -2.92$, $p = 0.004$). In order to figure out the direction of the interaction effect, a series of pairwise comparisons was performed. In the subjective sentences, it was found that participants spent shorter times at this region when *kejian* was used than they did when the connective was underspecified (i.e., *suoyi*) (FP: $\beta = -0.10$, $SE = 0.04$, $t = -2.65$, $p = 0.008$; TG: $\beta = -0.10$, $SE = 0.04$, $t = -2.57$, $p = 0.012$; RP: $\beta = -0.14$, $SE = 0.04$, $t = -3.12$, $p = 0.002$). Meanwhile, no significant difference was observed between the reading times of content causal relations with *yin'er* and those with *suoyi* (FP: $\beta = 0.03$, $SE = 0.04$, $t = 0.78$, $p = 0.448$; TG: $\beta = 0.04$, $SE = 0.04$, $t = 1.10$, $p = 0.280$; RP: $\beta = 0.03$, $SE = 0.04$, $t = 0.86$, $p = 0.381$).

Additional pairwise comparisons were performed between epistemic causal relations with *suoyi* and content causal relations with *suoyi*. No significant difference was observed (FP: $\beta = -0.02$, $SE = 0.04$, $t = -0.46$, $p = 0.638$; TG: $\beta = 0.05$, $SE = 0.04$, $t = 1.23$, $p = 0.245$; RP: $\beta = 0.08$, $SE = 0.04$, $t = 1.86$, $p = 0.067$), whereas the pairwise comparison between the conditions with specified causal connectives showed that content causal relations marked by *yin'er* resulted in longer reading times than epistemic causal relations marked by *kejian* (FP: $\beta = -0.14$, $SE = 0.04$, $t = -3.76$, $p < 0.001$; TG: $\beta = -0.10$, $SE = 0.04$, $t = -2.59$, $p = 0.011$; RP: $\beta = -0.10$, $SE = 0.04$, $t = -2.50$, $p = 0.011$).

5.4.2 Total fixation duration

Total fixation duration captures effects that occur after the end of the second clause (including the predicate) has been read. There was an interaction effect at all three regions (the connective region: $\beta = -0.28$, $SE = 0.06$, $t = -4.92$, $p < 0.001$; the subject region: $\beta = -0.18$, $SE = 0.06$, $t = -3.09$, $p = 0.001$; the predicate region: $\beta = -0.19$, $SE = 0.05$, $t = -3.51$, $p < 0.001$). In order to figure out the direction of the interaction effect, firstly a pairwise comparison was performed between the two conditions with *suoyi*. It was found that the epistemic relations with *suoyi* led to longer reading times than the content relations with *suoyi* (the connective region: $\beta = 0.16$, $SE = 0.04$, $t = 4.24$, $p < 0.001$; the subject region: $\beta = 0.18$, $SE = 0.04$, $t = 4.25$, $p < 0.001$; the predicate region: $\beta = 0.10$, $SE = 0.04$, $t = 2.65$, $p = 0.012$). Moreover, the 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 all three regions (the connective region: $\beta = -0.09$, $SE = 0.04$, $t = -2.23$, $p = 0.030$; the subject region: $\beta = -0.12$, $SE = 0.04$, $t = -2.86$, $p = 0.005$; the predicate region: $\beta = -0.16$, $SE = 0.04$, $t = -4.04$, $p < 0.001$). For the pairwise comparison between the content causal relations, the only effect found was at the connective region: *yin'er* resulted in longer reading times than *suoyi* at that position ($\beta = 0.20$, $SE = 0.04$, $t = 4.89$, $p < 0.001$). As reported earlier, this processing delay at the connective *yin'er* was also observed in FP, TG, and RP. In addition, as already reported for first pass reading-time measures, at both the connective and the predicate the content condition with *yin'er* showed longer reading times than the epistemic condition with *kejian* (the connective region: $\beta = -0.13$, $SE = 0.04$, $t = -2.87$, $p = 0.004$; the predicate region: $\beta = -0.10$, $SE = 0.04$, $t = -2.57$, $p = 0.009$).

5.5 Discussion

With the current experiment, we aimed to increase our understanding as to whether and how the subjectivity information (objectivity versus subjectivity) encoded in specified causal connectives affects the online processing of causally-related discourse. In particular, we focused on the research question: what does the degree of subjectivity encoded in specified causal connectives add to the online interpretation of causal relations compared to those marked with underspecified causal connectives?

Unlike previous processing studies, the current study simultaneously included specified causal connectives and underspecified causal connectives within the same language in the experimental design, and hence enabled us to directly compare the roles of these two types of causal connectives during online processing. For the epistemic causal relations, a comparison was made

between the subjective causal connective *kejian* and the underspecified causal connective *suoyi*. For the content causal relations, we compared the objective causal connective *yin'er* and the underspecified connective *suoyi*.

As predicted, the results from the comparison between the two epistemic conditions showed that, at the predicate region, the epistemic causal relations marked with *kejian* resulted in shorter reading times than the epistemic causal relations marked with *suoyi*. The difference between the two epistemic conditions solely concerns the connective, to be precise, the specificity of connectives with respect to subjectivity. Accordingly, we can conclude that the high degree of subjectivity encoded in *kejian* has facilitated the reading of the subjective epistemic relation. However, 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*. We observed no processing differences between the *yin'er* condition and the *suoyi* condition, except at the connective region. This finding suggests 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*. Probably, people generally expect objective content causal relations more than subjective epistemic causal relations. This result, in fact, conforms to the proposal in previous studies that readers might prefer to build the simplest possible discourse representation during interpretation (see Traxler et al. 1997b), in this case the objective relation.

As another important finding, the current experiment replicated the results reported in Traxler et al. (1997a). In their experiment, it was found that there were significant differences in total time between epistemic and content causal relations marked with *because* from the connective region through the predicate region. In the current experiment, we observed similar results: the epistemic causal relations with *suoyi* showed longer total reading times than the content causal relations with *suoyi* at all three critical regions. In both *suoyi* conditions, the connective's degree of subjectivity did not play a role, since the underspecified connective *suoyi* does not provide any information with respect to subjectivity. Therefore, this result probably reflects the role of the relation's degree of subjectivity itself. It should be noted that, even if the effect was observed at the connective region or at the subject region, the effect would still reflect the role of the relation's degree of subjectivity. This is because, as discussed earlier, the measure of total fixation duration captures effects that occur after readers have reached the end of the sentence. This is an important finding because it reconfirms the subjective complexity hypothesis. Namely, a subjective relation is cognitively more complex than an objective relation, hence the former requires more processing time (more re-reading times in the current study) than the latter.

Crucially, we want to stress that the effects observed at the connective and at the predicate, taken together, reveal the way in which the high degree of subjectivity encoded in the connective and the high degree of subjectivity associated with the relation affect each other during online processing. As discussed above, at the predicate region, when the connectives were underspecified with respect to subjectivity, subjective epistemic causal relations resulted in longer total reading times than objective content causal relations. This finding has been related to the inherent cognitive complexity of epistemic causal relations in general. 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 time than the predicate region of the *suoyi* clauses. It might be the case that the processing difficulty induced by the high degree of subjectivity arose at the specified causal connective, which contributed to the main effect of connective at the connective region for first pass total gaze duration and regression path duration: overall, specified causal connectives *kejian* and *yin'er* took longer to read than the underspecified causal connective *suoyi*. Admittedly, the frequency of use of the connective is a confounding factor to subjectivity at the connective region. As introduced earlier, *suoyi* is more frequently used than both *kejian* and *yin'er*. However, frequency is not very likely to affect later reading-time measures such as TG and RP. These two measures are associated with regressive eye movements, and thus they are generally assumed to be sensitive to cognitive complexity and comprehension difficulty. Further research is needed to disentangle the role of the frequency of the connective from the role of subjectivity in the connective. At the moment, we can tentatively assume that a high degree of subjectivity plays the role of activating specific cognitive processes during reading, whether it is encoded in the connective or in the relation, and, moreover, the role of the former (i.e., a high degree of subjectivity of the connective) can cancel out the role of the latter (i.e., a high degree of subjectivity associated with the relation) when they co-occur in the same sentence.

The effect that the *yin'er* condition took longer processing times than the *kejian* condition at the connective region is unexpected. What makes the objective connective *yin'er* difficult to read? Perhaps this is because *yin'er* is perceived as formal, and therefore does not fit in with informal occasions. 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. Probably, the formal style of *yin'er* makes it

look inappropriate in the informal materials. This could be an explanation for the unexpected finding that the objective causal connective *yin'er* leads to longer reading times than the subjective causal connective *kejian*. This could also explain the findings in all four measures that *yin'er* takes longer to read than the underspecified causal connective *suoyi* at the connective region. This assumption can be tested in future experiments using the same experimental design, but with materials in a more formal style.

Finally, we want to discuss what possibly makes texts encoded with subjectivity difficult to process during online reading. There are different theoretical explanations. On the grounds of the indications from previous studies that the word *kejian* itself encodes the speaker/writer's perspective whereas the objective causal connective *yin'er* and the underspecified causal connective *suoyi* do not (see Section 5.2.1), we think the most plausible explanation is as follows. Subjective information requires more cognitive efforts 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)' (Canestrelli et al. 2013; De Smet & Verstraete 2006; Ross 1970; Rutherford 1970).⁶ *Kejian* directly encodes the speaker's perspective, 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 at the word itself. 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 in subjectivity account' can explain the major results of the current experiment in a very sensible way. Moreover, this explanation is plausible also because it is in line with native speakers' intuitions about the conceptual difference between *kejian* and *suoyi*. Compare (18), (19), and (20).

- (18) *Meng Na na tiao kuzi xianzai xiande hen fei, suoyi wo renwei ta bi yiqian shou le bu shao.*

'That old pair of trousers now look very baggy on Meng Na, **so I think** she has become much thinner now than before.'

⁶ In the case of *kejian*, the previous corpus-based study has shown that it only encodes the speaker's perspective (see Chapter 3), so the extra proposition should purely be 'I think'.

- (19) **Meng Na na tiao kuzi xianzai xiande hen fei, **kejian wo renwei** ta bi yiqian shou le bu shao.*

‘That old pair of trousers now look very baggy on Meng Na, **so I think** she has become much thinner now than before.’

- (20) *Meng Na na tiao kuzi xianzai xiande hen fei, **kejian** ta bi yiqian shou le bu shao.*

‘That old pair of trousers now look very baggy on Meng Na, **so** she has become much thinner now than before.’

As shown in (18), the underspecified causal connective *suoyi* allows the addition of the words *wo renwei* ‘I think’ right after it, whereas (19) illustrates that adding *wo renwei* ‘I think’ after the subjective connective *kejian* makes the originally appropriate sentence (20) 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: the epistemic causal relation marked with *kejian* (20) and the epistemic causal relation marked with “*suoyi* I think” (18). 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 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 et al. 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. 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 have to establish

whether the subjective complexity observed in the current and previous processing studies can be related to these cognitive processes .

CHAPTER 6

Conclusion and discussion

The messages in discourse are not restricted to objective propositions about events or facts that can be observed in the outside world. Sometimes, the speaker “imprints a specific seal” upon the message, even without being aware of it (Kristeva 1989: 11). Messages involving the speaker’s imprint are assumed to be encoded with a high degree of subjectivity. Previous research has shown that the notion of subjectivity is useful in explaining the usage patterns of causal connectives in several European languages. So far, it has remained unclear whether the usages of Chinese causal connectives can also be described in terms of subjectivity. This dissertation provided a systematic analysis of Chinese causal connectives in their actual use with the subjectivity approach. We made a thorough corpus-based investigation of a wide range of Chinese causal connectives in terms of subjectivity across different genres of written text. Furthermore, with eye-tracking techniques, we explored the roles of causal connectives in the online processing of causally related texts. By combining corpus-linguistic and psycholinguistic methods, the present dissertation contributes to the understanding of the ways in which subjectivity, as a cognitive principle, constrains human production and comprehension of causally related messages.

This chapter first presents summaries of the results that have been achieved in the present dissertation, and provides answers to the research questions raised in Chapter 1. Following the summaries, topics and directions for future research in the study of Chinese causal connectives are suggested.

6.1 Assessing the relevance of subjectivity for the characterization of causal connectives in Chinese

Chapter 2 provided a literature review exploring the potential for subjectivity to explain the usage patterns of Chinese causal connectives. This literature review aimed to address the following research questions.

Research question 1

How do Chinese linguists define subjectivity? Can their approaches be related to western approaches of subjectivity?

Research question 2

Which analytical categories have been used in previous studies on Chinese causal connectives? Can these categories be related to analyses in terms of subjectivity, and if so, how?

Research question 3

Do Chinese and European studies address the issue of subjectivity in causality with similar methods?

In answer to Research question 1, this literature review has shown that the definitions of subjectivity in previous Chinese studies can already be linked to the theories of the western linguists Lyons (1977), Langacker (1990), and Traugott (1995, 2010). Based on these three theories, Chinese linguists have defined subjectivity in terms of the expression of *self*, speaker-relatedness, and on/off-stage conceptualization. In line with these definitions, a linguistic expression is considered to be subjective when it involves the speaker's perspective (that is, it concerns the aspect from which the speaker views the circumstances described in the utterance), the speaker's affect (dealt with mainly in terms of empathy) and epistemic modality (concerning the speaker's judgment on the likelihood that the propositional content of the utterance is true).

An increasing number of studies have applied the subjectivity approach to the analysis of various linguistic phenomena in Chinese discourse, and have demonstrated the usefulness of the notion of subjectivity in describing and explaining language use in Mandarin Chinese. So far, only a few studies have analyzed the uses of causal connectives in Chinese such as *jiran*, *yinwei* and *youyu* (all translated as 'because') in terms of subjectivity. In these studies, the distinction between the content, epistemic, and speech-act domains (Li 2011; Li & Liu 2004) and the distinction between descriptive and inferential causality (Gao 2013; Guo 2008; Deng 2007) were assumed to be relevant to the notion of subjectivity. Moreover, results of this literature review demonstrated that many other analytical categories for describing differences in the usage of Chinese causal connectives can also be related to subjectivity, although previous Chinese studies have not established such links explicitly. These categories include the modality of the consequent clause (Guo 2006, 2008; X. Li 2009; Zhao 2003), the distinction between ideational and interpersonal function (Biq 1995; Song & Tao 2009), causal sequencing (i.e., forward order vs. backward order) (Gao 2013; Guo 2008; Li 2011; Li & Liu 2004; Qu 2002; Zhang & Zhang 2011), and the presence/absence of the so-called speaker subject (Li & Liu 2004). In answer to Research question 2, these findings allowed us to hypothesize that the subjectivity account for connective use is generalizable to the entire family of Mandarin causal connectives. In other words, we hypothesized that the subjectivity approach should be able to provide a unified account for the distributional patterns of the full range of causal connectives that commonly occur in Chinese discourse.

In answer to Research question 3, we evaluated the methods with which previous Chinese studies have addressed the issue of subjectivity. Based on

this evaluation, we concluded that the current qualitatively-oriented methods need to be extended with quantitative and statistical approaches in order to enhance the generalizability and replicability of the results. Ways to improve the methodology include broadening the range of causal connectives as objects of examination for an overall view of their distributional patterns; analyzing language in actual use (i.e., corpus-based analysis) rather than constructed examples; combining qualitative analysis with quantitative approaches and related statistical techniques to guarantee the statistical validity of the results; involving more subjectivity indicators in analyses (apart from Sweetser's domains of use), in order to specify more fine-grained distinctions between different degrees of subjectivity; and developing a detailed analytical model with precise specifications and decision rules for assessing degrees of subjectivity in order to ensure high inter-coder reliability and replication.

Chinese studies have not yet paid attention to the relationship between genre and degree of subjectivity that causal connectives encode, a topic that has recently received increasing academic interest in western literature (see Sanders & Spooren 2009; Spooren, Sanders, Huiskes & Degand 2010; Stukker & Sanders 2012). Genre is conventionally defined as a recognizable communicative event characterized by a set of communicative purposes identifiable and mutually understood by the members of the discourse community (Swales 1990; Trosborg 1997). News reports, novels, opinion pieces are three genres frequently mentioned in discourse studies, as they are characterized by salient but distinctive communicative purposes, which cohere with the three functions of language given in Bühler's *Organon* model – information, expression, and persuasion (see the model in Renkema 2004). Similarly, Chinese discourse analysts often divide written texts into three genres: the narrative genre, the informative/descriptive genre, and the argumentative genre (Guo 2006). It is generally assumed that opinion pieces display an overall higher degree of subjectivity than the other two genres, given that opinion pieces often involves arguments to convince readers that the claim or conclusion is true (Sanders 1997). Is it likely that causal connectives encoded with higher degrees of subjectivity tend to occur more often in opinion pieces than news reports and novels? If so, the robustness of a connective's usage pattern with respect to subjectivity can be called into question. If a causal connective occurs in its less typical genre of use, does it assimilate to the inherent features of the genre in such a way that it obtains a higher or lower degree of subjectivity than usual? Keeping these questions in mind, I intended to take the impact of genre into consideration in the subjectivity analysis.

Based on results from corpus-based analyses of connective use, a few pioneering studies have started investigating connective use from the perspectives of discourse processing and language acquisition (Canestrelli,

Mak & Sanders 2013; Evers-Vermeul & Sanders 2009, 2011; Spooren & Sanders 2008; Traxler, Bybee & Pickering 1997a; Traxler, Sanford, Aked & Moxey 1997b; van Veen 2011). Insights have been obtained from these areas, revealing that causal connectives varying in degrees of subjectivity show different patterns. These findings suggest that subjectivity is cognitively relevant. As Sanders and Spooren (2009) suggest, subjectivity is likely to be one of the cognitive principles that organize our knowledge of linguistic use of causal connectives. If subjectivity is really cognitively relevant, we should expect that it is systematically encoded in Chinese discourse, and that linguistic expressions of subjectivity manifest themselves in some way in cognitive activities, such as those involved in online discourse processing. In previous Chinese studies, differences in subjectivity (or domain) have been used as descriptive entities, but to the best of our knowledge no attempts have been made to explore whether these differences also play a role during language processing.

6.2 Subjectivity profiles identified, genre-sensitivity, and language-specific properties

Two corpus-based studies, one on Chinese result connectives and the other on Chinese reason connectives, were conducted in order to address the following research questions.

Research question 4

Do Chinese causal connectives show systematic variation in terms of the degree of subjectivity they encode? If so, how? Are there language-specific properties in this respect?

Research question 5

Do subjectivity profiles of causal connectives vary across genre?

A set of eight Chinese causal connectives were investigated, which nearly exhausts the full list of those generally acknowledged in the literature: five result connectives translated as ‘so/therefore’ (i.e., *kejian*, *suoyi*, *yinci*, *yin’er*, and *yushi*) and three reason connectives translated as ‘because’ (i.e., *jiran*, *yinwei*, and *youyu*). To study the impact of genre on connectives’ subjectivity meanings, Research question 5, we collected samples in a balanced way from three genres: news reports, opinion pieces, and novels. Fragments of the former two genres were taken from the newspaper corpus *People’s Daily Online*. Fragments of novels were taken from the CCL Corpus that was created by the Center for Chinese Linguistics at Peking University. For each genre, seventy-five occurrences of each connective were randomly selected (225 fragments per connective).

An analytical model was created on the basis of Sanders and Spooren (2009), which consists of four subjectivity indicators: domain, propositional attitude, the presence and identity of Subject of Consciousness (i.e., the person whose intentionality is conceptualized as the ultimate source of the causal event). Table 1 illustrates the values associated with each of these four variables.

Table 1. Model for subjectivity analysis with variables and their values

Variable	- Subjectivity value +
Domain	Non-vol content < Vol content < Speech act / Epistemic
Propositional attitude	Physical fact < Mental fact < Speech act / Judgment
The presence of SoC	No SoC < Explicit SoC < Implicit SoC
The identity of SoC	Character SoC < Current speaker SoC < Author SoC

There are minor differences between the two corpus-based studies with respect to the exact values we specify for some variables. For example, the results of the first corpus-based study (i.e., the study on result marking) suggest that the observed genre-sensitivity associated with ‘current speaker SoC’ does not truly reflect the genre-sensitivity of the connective. The value of current speaker SoC is genre-sensitive itself: overall current speaker SoCs were found more frequently in novels than in news reports or opinion pieces (Chapter 3). Consequently, this contributes to the observed variation across genres – that connectives show an increase in current speaker SoCs in novels compared to the other two genres. In order to remove the confounding effect of the genre-sensitivity of current speaker SoC, and to truly capture the connective’s sensitivity to genre in the second corpus-based study (i.e., the one on reason marking), the two subjective categories – current speaker SoC and author SoC – were combined under the label of speaker SoC.

6.2.1 Six stable subjectivity profiles

In answer to Research question 4, it was found that, overall, Chinese causal connectives show systematic variation in terms of subjectivity. Six causal connectives – three reason connectives, *jiran*, *yinwei*, *youyu* (all translated as ‘because’), and three result connectives, *kejian*, *yin’er*, *yushi* (all translated as ‘so/therefore’) – display subjectivity profiles that differ from each other, and are robust across genres. Tables 2 and 3 summarize the subjectivity profiles of these result connectives and reason connectives, respectively.

Table 2. Three result connectives' subjectivity profiles that are robust across genres

Result connective (translated as 'so/therefore')	Domain	Propositional attitude	Presence of SoC	Identity of SoC
<i>kejian</i>	+epistemic	+judgment	+implicit	+speaker
<i>yin'er</i>	+non-vol. content	+physical fact	+no SoC	d.n.a
<i>yushi</i>	+vol. content	+physical fact	+explicit	+character

Table 3. Three reason connectives' subjectivity profiles that are robust across genres

Reason connective (translated as 'because')	Domain	Propositional attitude	Presence of SoC	Identity of SoC
<i>jiran</i>	+epistemic +speech act	+judgment +speech act	+implicit	+speaker
<i>yinwei1</i>	+epistemic	+judgment	+implicit	+speaker
<i>yinwei2</i>	+content	+facts	+no SoC	d.n.a
<i>youyu</i>	+non-vol. content	+physical fact	+no SoC	d.n.a

Note: *Yinwei1* refers to inter-sentential *yinwei*, and *yinwei2* refers to sentence-initial *yinwei*.

Kejian 'so' is restricted to subjective uses. It specializes in expressing causal relations in the epistemic domain (see (1)), and occasionally occurs in the speech-act domain, but is not used in the content domains. It always occurs in highly subjective contexts involving judgments and implicit speaker SoCs (see (1)).

- (1) *Cong zhe xiang weiqi shi nian de yanjiu laikan, zhengmian qingxu youzhu yufang guanxinbing, **kejian** baohu xinzang xuyao zengjia zhengmian qingxu.*

'According to the ten-year long study, positive moods help to prevent the occurrence of coronary heart disease, **kejian** one needs to increase his positive moods to protect the heart.'

The other five connectives can each be used to express different degrees of subjectivity, e.g., causal relations in the content domain as well as those in the epistemic domain. However, these five causal connectives differ in terms of their distributional frequency. *Jiran* 'because' frequently expresses epistemic and speech-act relations involving judgments, speech acts, and implicit speaker SoCs (e.g., the type of contexts that *kejian* is tied to, such as (1)). The speech-act use of *jiran* is exemplified in (2). By contrast, *youyu* 'because' frequently expresses non-volitional content relations holding between facts, where no SoC is discerned. Very similar to *youyu*, *yin'er* 'as a result, so' frequently occurs in the non-volitional content domain, where no SoC is explicitly or implicitly present (contexts such as (3)).

- (2) *Jiran* zhenzhong zai Wenchuan, weihe Beichuan bi Wenchuan shou zai geng zhong?
 ‘*Jiran* the epicenter of the earthquake is Wenchuan, why is Beichuan more seriously affected than Wenchuan?’
- (3) *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, *yin’er* the petals are blue.’

Yushi ‘so’ frequently expresses volitional content causal relations, and frequently co-occurs with an explicit character SoC (see (4)).

- (4) *Ta tingdao men wai you dongjing, yushi ta tui kai men qu kan ge jiuqing.*
 ‘He heard some sounds outside the door, *yushi* he pushed open the door to see what actually happened.’

Yinwei ‘because’ shows different usage patterns, depending on its position in the sentence. The inter-sentential *yinwei* (in the sequence of ‘Q (consequent clause), *yinwei* P (antecedent clause)’) frequently expresses causal relations in the epistemic domain that involve the off-stage/implicit speaker’s judgments (e.g., *This kind of suspicion has its basis, yinwei the same situation has already repeated itself eighteen times.*), whereas the sentence-initial *yinwei* (in the sequence of ‘*yinwei* P, Q’) is frequently used to mark content causal relations holding between observable states or events (e.g., *Yinwei the wound is inflamed, the entire face is swollen up.*)

To sum up, what do the above-discussed findings add to our previous knowledge about the uses and functions of causal connectives in Chinese discourse? First, our finding that *jiran* ‘because’ is more subjectively-oriented than *youyu* ‘because’ corroborates the corpus-based finding reported by Li and Liu (2004). As a point to note, what the present outcome contributes is not simply a replication, because three other variables (i.e., propositional attitude, the presence and identity of the SoC) were taken into consideration in addition to the one used in Li and Liu (2004) (i.e., domain). Consequently, a more detailed description of the two connectives’ profiles was produced in terms of subjectivity.

Second, some of the introspective observations in Xing (2001) were borne out. Empirical evidence was obtained which confirms Xing’s (2001) claim that *kejian* is (in his terms) inferential, whereas *yin’er* is descriptive. In the terms of the present approach, *kejian* is associated with a higher degree of subjectivity than *yin’er*.

Third, findings about the two forms of *yinwei* – sentence-initial and inter-sentential – are compatible with the functional account proposed by Song and Tao (2009). According to them, inter-sentential *yinwei* performs an interactional function in mitigating and diverting potential sources of trouble, whereas sentence-initial *yinwei* performs an information-sharing function. In our view, the difference in degree of subjectivity can explain why the two forms are used to perform different communicative functions. The inter-sentential *yinwei* functions as an interactional device because it has a highly subjective meaning. Following the idea that meaning is an intermediate link between form-function pairs (Evers-Vermeul 2005), the speaker will generally choose the form with a highly subjective meaning (in this case the inter-sentential *yinwei*) for the function of defending her own point of view. Alternatively, the speaker will use the form with a highly objective meaning (in this case the sentence-initial *yinwei*) for the function of sharing information with the addressee. There are exceptions, such as the use of objective causal connectives to express subjective epistemic causal relations. Some such exceptional uses of connectives are explained in terms of the speaker's rhetorical purposes in studies such as Sanders and Spooren (2013), Stukker, Sanders, and Verhagen (2008, 2009) and van den Hoven (1997). For example, on some occasions, the speaker might want to background her role as a concluder in the causal relation in order to be more convincing (see Sanders & Spooren 2013; Stukker, Sanders, and Verhagen 2008). The rhetorical use of causal connectives should be an interesting topic for future research.

Fourth, our corpus-based study on result connectives revealed that *yushi* typically expresses causal relations in the volitional content domain, but that it can also be used to express causal relations in the epistemic domain involving inferences or conclusions. This finding runs counter to Lu (2000), who argues on the basis of intuition and generalization that *yushi* cannot be used to express inferential causality. This new finding about the epistemic use of *yushi* adds to our knowledge on language use, and demonstrates the importance of combining linguists' intuitions with genuine distributional data.

6.2.2 The issue of genre-sensitivity

An important finding of the present dissertation concerns the genre-sensitivity of connectives' subjectivity profiles. In answer to Research question 5, our results show that the subjectivity profiles of two frequent causal connectives – *yinci* and *suoyi* (both translated as 'so/therefore') – are genre-sensitive. *Yinci* and *suoyi* are not as purely descriptive as Xing (2001) suggests. They show a preference for descriptive uses (e.g., volitional content causal relations between physical facts, character SoCs) in novels, but in

news reports and opinion pieces they are frequently used in the subjective epistemic domain involving judgments and implicit speaker SoCs.

Two implications can be drawn from the observed genre-sensitivity. First, Xing (2001) focused only on novels, and, for that matter, his study only uncovered part of the picture, i.e., the part of the descriptive usage. Narrative texts focus on narrating real-world events/states of affairs in sequential order, and hardly involve claims or argumentation. Probably due to this type of bias associated with the narrative texts, the inferential use did not come into notice in Xing (2001). To reduce the risk of being biased by the genre during analyses of linguistic expressions in discourse, we suggested taking different genre types into account. What is more, we proposed analyzing genre as a separate factor in our statistical models. There is a danger that a corpus containing a mixture of genres will neutralize genre-specific patterns, if we do not analyze the linguistic element in question per genre.

The second implication concerns the following question. Are the subjectivity profiles really part of the inherent semantic characteristics of the connectives themselves, or are they (partially) determined by the pragmatics of the context? Given that results show genre-sensitive profiles as well as robust ones, it is likely that subjectivity meanings are specified in some causal connectives (in the lexicon) but not in all. For the five causal connectives that maintain stable subjectivity profiles across genres, the feature of subjectivity is probably firmly specified in the lexicon in either the objective or the subjective direction. That is, they have a meaning in and of themselves with respect to the degree of subjectivity they typically express, which enables them to influence the reader's interpretation of the causal relation in a way that is not directly affected by the reader's 'genre knowledge'. We call these causal connectives 'specified causal connectives'.

Figure 1 tentatively illustrates the role of the so-called specified causal connectives, abbreviated as 'CON', in the process of relational interpretation.

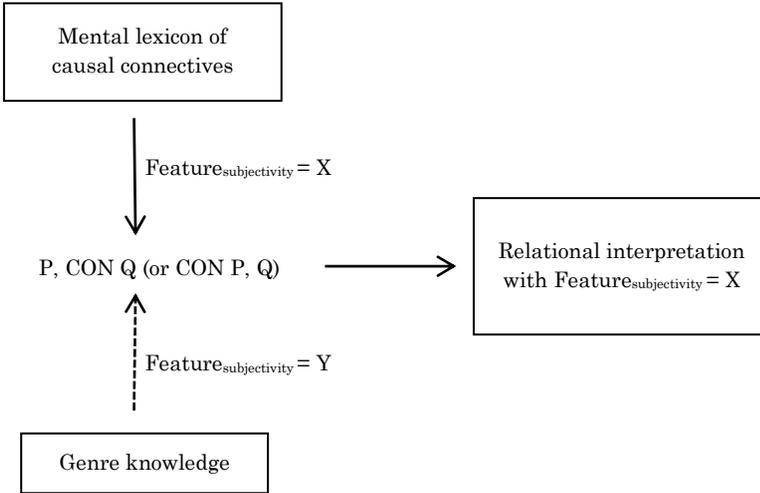


Figure 1. The process of interpreting causal relations with specified causal connectives¹

We assume that both genre knowledge and the specification of connectives play a role when readers interpret causal relations with respect to subjectivity. In Figure 1, X stands for the value of subjectivity specified in the mental lexicon of causal connectives, and Y represents reader knowledge about the degree of subjectivity that is typically associated with the genre in question. Furthermore, we suppose that these two factors differ in the strength of their roles: specified causal connectives impose a strong constraint on relational interpretation whereas the constraint imposed by genre knowledge is rather weak, which is symbolized by using the dotted arrow in the latter case and by using the solid arrow in the former in Figure 1.

For the two genre-sensitive causal connectives, *yinci* and *suoyi*, their subjectivity meanings are probably underspecified in the lexicon, and thus are likely to be derived from the pragmatics of the context, or in other words, genre knowledge. We call this type of causal connectives ‘underspecified causal connectives’ (see also Spooren 1997). Figure 2 tentatively schematizes the process of interpreting causally-related clauses containing underspecified causal connectives, abbreviated as ‘con’.

¹ Figure 1 does not represent the role of causal connectives that are used for rhetorical purposes.

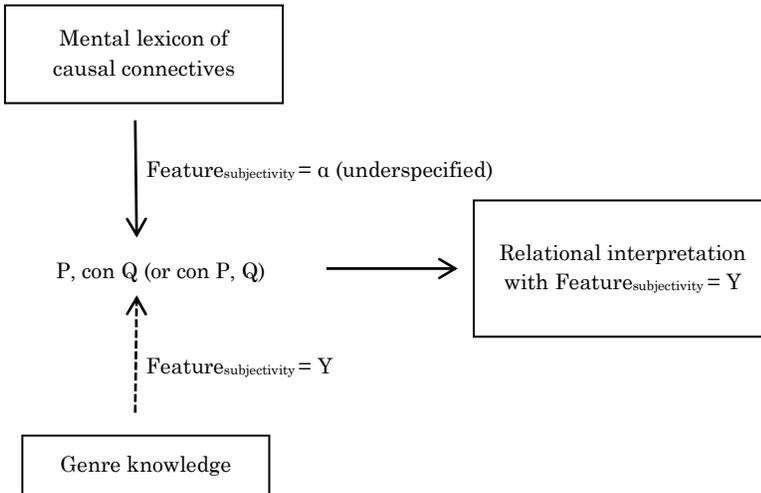


Figure 2. The process of interpreting causal relations with underspecified causal connectives

α means that the feature of subjectivity is underspecified in the lexicon of causal connectives, and Y stands for the value of subjectivity that is suggested by a reader's genre knowledge. It can be seen that, in this case, genre knowledge (a weak constraint as it is) plays a crucial role during interpretation, given that the causal connective does not provide information on the degree of subjectivity.

Our assumption about the role of pragmatics in affecting the interpretation of an utterance as a whole has its theoretical basis in the literature. It is generally assumed that elements that are contextually implied contribute to the content conveyed by an utterance without being part of what the utterance literally means, by virtue of the processes referred to as 'pragmatic enrichment', or 'conversational implicature' (Grice 1989; Levinson 2000; Recanati 1993, 2004). Furthermore, Traugott proposes that elements from pragmatic contexts in which an expression typically occurs can become part of a word's conventional meaning, referred to as "the semanticization of pragmatics" (see Barlow & Kemmer 2000; Traugott 2011). Her proposal is based on qualitative analyses of a few European languages, which reveal that pragmatic inferences may eventually become conventionalized in connectives such as English *while* and German *weil* (see Keller 1995; Traugott 1995). In German, a causal inference originates from the temporal meaning of *weil* 'during', and this causal inference has gradually been semanticised in *weil* 'because'. As for the English *while*, the inference became conventionalized as "surprise concerning the overlap in

time or the relations between event and ground”, which led to the adversative, concessive meaning (Traugott 1995: 41).

6.2.3 Language-specific properties

European languages such as Dutch, French, and German have shown clear differences in the way they structure the lexicon of causal connectives in terms of subjectivity. For example, Dutch speakers have a dedicated marker *daardoor* ‘as a result’ for non-volitional causality (Pander Maat & Degand 2001; Pander Maat & Sanders 2001), whereas French and German do not (Degand & Pander Maat 2003; Stukker & Sanders 2012). In Polish, causality is most typically expressed via epistemic or intersubjective construals marked with connectives *bo* ‘because’ and *to* ‘then’, and seldom focuses on objective relations in the real world (Dancygier 2009).

Not surprisingly, the Chinese language also displays some specific properties with respect to the organization of its causal connectives. Chinese speakers have a connective *kejian* at their disposal, which is restricted to expressing epistemic causality (and only occasionally expresses speech-act causality). More specifically, our results show that *kejian* exclusively occurs in extremely subjective epistemic relations – those involving implicit speaker SoCs. That is, Chinese speakers have a specified connective (i.e., *kejian*) that is confined to expressing the perspective of the speaker, who is obligatorily off-stage. This is different from European languages such as Dutch, French, and German, in which a causal connective specializing in expressing epistemic causal relations (e.g., *dus* ‘so’ in Dutch) is also used (albeit less frequently) in the volitional content domain.

Another property associated with connective use is not specific to the Chinese language itself. It is a shared property between Chinese and Dutch. In both languages, the concept of volitionality is highlighted in the lexicon of causal connectives. The Dutch *daarom* ‘that is why’ and the Chinese *yushi* ‘so’ are both typically used to express relations in the volitional content domain. These findings confirm the necessity of supplementing Sweetser’s (1990) domain theory with Stukker et al.’s (2008) notion of volitionality in characterizing the use of causal connectives in various languages.

6.3 The impacts of subjective, objective, and underspecified causal connectives on discourse processing

Previous experimental research on Dutch and English has revealed that causal relations or connectives that are associated with different degrees of subjectivity show different patterns in online processing. To be concrete, objective causal relations and connectives are processed faster than subjective causal relations and connectives during online reading

(Canestrelli 2013; Canestrelli, Mak & Sanders 2013; Noordman & De Blijzer 2000; Traxler, Bybee & Pickering 1997a; Traxler, Sanford, Aked & Moxey 1997b). These findings provide support for the idea that subjectivity is cognitively relevant. More precisely, these findings are in line with the so-called ‘subjective complexity hypothesis’ (Noordman & De Blijzer 2000; Sanders 2005). That is, due to subjective complexity, comprehending subjective information requires more cognitive effort than comprehending objective information, which is why processing the former takes longer than processing the latter. Based on these findings, we raised the sixth research question.

Research question 6

How are differences in degree of subjectivity between Chinese causal connectives reflected in online discourse processing?

The experiment by Traxler et al. (1997a) used English sentence pairs such as (5) and (6).

- (5) The goal keeper won the game **because** the other team didn’t score any goals.
- (6) The goal keeper knew how to play the game **because** the other team didn’t score any goals.

Within each pair, one sentence (5) expresses a content (or in their terms *simple-causal*) relation and the other sentence (6) expresses an epistemic (or in their terms *diagnostic*) relation. For both cases, *because* is used to introduce the antecedent clause backward to the preceding consequent clause. As an underspecified connective, *because* does not provide any information with respect to the degree of subjectivity or the type of causal relation that is likely to follow. In this case, the predicate (*didn’t score*) is the first possible position where the necessary information becomes available for readers to determine that the sentence (6) expresses a relation in the epistemic domain. The results showed that precisely at the predicate position of the second clause, a processing delay occurred in the epistemic compared to the content relations.²

Canestrelli et al. (2013) focused on other experimental conditions, i.e., cases in which specified causal connectives (i.e., objective vs. subjective ones) are used. Specified causal connectives are assumed to provide explicit

² For *first pass time*, Traxler et al. (1997a) also observed a processing delay at the region immediately after the predicate, i.e., at the final region of the *because*-clause in the epistemic compared to the content conditions. This effect can be taken as a spill-over effect from the predicate region.

information on the degree of subjectivity of the unfolding relation, which matches with the degree of subjectivity of the causal connective. Canestrelli et al. compared reading times of Dutch content and epistemic relations respectively marked by an objective causal connective *omdat* and a subjective causal connective *want* (both translated as ‘because’), their experimental materials inspired by 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. However, the processing delay associated with the epistemic relation was observed earlier in the Dutch materials than in the English materials: in the *want* condition, there was a delay immediately after the connective, and even at the connective itself, but not at the predicate region. Canestrelli et al. (2013) explained this by claiming that in their experiments the epistemic relation was not constructed on the basis of the content of the clauses, but on the basis of the information provided by the specified causal connective. They conclude that *omdat* immediately triggers the representation of an objective relation whereas *want* directly triggers the representation of a subjective relation.

So far, psycholinguistic studies have either looked at a language (i.e., English) with an underspecified causal connective (i.e., *because*) or a language with only specified causal connectives (e.g., *omdat* and *want*). Thus, it is not clear exactly how specified causal connectives differ from underspecified causal connectives in steering the processing of causally related texts. For example, is there a processing difference between an underspecified causal connective versus an objective causal connective when they are used to express causal relations in the content domain? Or do readers treat these types of connectives in the same way, and only process the information differently if they encounter a subjective causal connective? In order to address these unsettled issues, we conducted an eye-tracking experiment using Chinese materials, which allowed us to directly compare the impacts of specified and underspecified causal connectives within a single experimental design.

The causal connectives *kejian*, *suoyi*, and *yin'er* (all translated as *so/therefore*) were used in the materials, because they represent typical instances of subjective, underspecified, and objective causal connectives, respectively (see Chapter 3). Four experimental conditions were involved: epistemic relations with the subjective causal connective *kejian*, epistemic relations with the underspecified causal connective *suoyi*, content relations with the objective causal connective *yin'er*, and content relations with *suoyi*, which are exemplified in (7) to (10), respectively. The connective usage is appropriate in each of the four conditions.

- (7) *Meng Na na tiao kuzi xianzai xian de hen fei, **kejian** ta bi yiqian shou le bushao. 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.’
- (8) *Meng Na na tiao kuzi xianzai xian de hen fei, **suoyi** ta bi yiqian shou le bushao. 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.’
- (9) *Meng Na yi nian lai baoshou weibing de zhemo, **yin’er** ta bi yiqian shou le bushao. 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.’
- (10) *Meng Na yi nian lai baoshou weibing de zhemo, **suoyi** ta bi yiqian shou le bushao. 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.’

One characteristic of this design is that the difference between the two epistemic conditions solely concerns whether the connective is encoded with a high degree of subjectivity (in the case of *kejian*) or the connective is underspecified with respect to subjectivity (in the case of *suoyi*). Therefore, processing differences (if any) between these two conditions can directly reflect what the high degree of subjectivity encoded in *kejian* adds to the online processing of epistemic causal relations. In the same way, the results also shed light on what a highly objective causal connective *yin’er* contributes to the online interpretation of content causal relations, compared to a causal connective underspecified for subjectivity (i.e., *suoyi*).

The second clause of each text was the target of our analysis, which was specified to contain a connective region, a subject region, and a predicate region. Table 4 illustrates the way in which a target clause was split into these three regions.

Table 4. The division between three target regions

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

The current experiment (partially) 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) showed a higher processing cost, evidenced by longer total reading times at the connective region, the region immediately following the connective, and the predicate region, compared to content relations with the same causal connective. This finding reconfirms the subjective complexity hypothesis: a subjective relation is cognitively more complex than an objective relation, and hence the former requires more processing time (more re-reading in the current study) than the latter.

Results also showed that epistemic causal relations marked with *kejian* such as (7) have a lower processing cost (shorter reading times with respect to both first-pass reading and re-reading) than epistemic causal relations marked with *suoyi* such as (8) at the predicate region of the consequent clause. This finding enabled us to conclude that the high degree of subjectivity encoded in *kejian* facilitated the reading of the subjective epistemic relation. Why does *kejian* have this facilitating effect at the predicate region? We proposed that subjective information, being cognitively complex as it is, immediately activates specific cognitive processes (e.g., setting up and instantiating specific mental spaces) during reading, whether it is encoded in the connective or in the relation. Moreover, we proposed that the activating role of the former (i.e., the high degree of subjectivity encoded in the connective) can cancel out the activating role of the latter (i.e., the high degree of subjectivity encoded in the relation) when they co-occur in a sentence. Accordingly, when epistemic relations are marked with the subjective connective *kejian*, the cognitive processes required for the representation of subjective information are already activated at the causal connective, and thus are not activated at the predicate region. When epistemic relations are marked with the causal connective *suoyi*, which is not encoded with subjective information, the cognitive processes required for the representation of subjective information are probably activated no earlier than at the predicate region, where the propositional content of the segments is adequate for readers to determine that a representation of a subjective relation is to be built. The above arguments can explain the observed processing differences between the two epistemic conditions: the predicate region of the *kejian* clauses showed shorter reading times than the predicate region of the *suoyi* clauses.

Meanwhile, the above arguments suggest that extra processing time should arise at the connective *kejian* or at the subject region. Indeed, between the two epistemic conditions, we observed longer processing times at *kejian* than at *suoyi*, indicated by the measure of *first pass total gaze duration* (i.e., the sum of durations of fixations that fall within a region before the region is left progressively) and the measure of *regression path duration* (i.e., 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).³ These two measures are generally assumed to be indicators of cognitive complexity and comprehension difficulty, because they are associated with regressive eye movements. The frequency of the connective is not likely to be a crucial factor affecting these two measures. That is, this finding regarding the processing disadvantage associated with *kejian* is in support of the subjective complexity hypothesis.

An unexpected finding concerns the two conditions with specified causal connectives: at the connective region, the objective causal connective *yin'er* resulted in longer reading times than the subjective causal connective *kejian* (in terms of first-pass reading and re-reading). What makes this objective causal connective difficult to read compared to a subjective causal connective that is associated with extra cognitive complexity? We cannot explain this in terms of frequency in its general sense, given that the objective causal connective *yin'er* (1.50 per 10,000 words) is more frequent than the subjective causal connective *kejian* (0.63 per 10,000 words) in written discourse. However, in spoken discourse (as reflected in Taiwan Mandarin Spoken Wordlist), we found that *yin'er* never appeared in conversations whereas *suoyi* and *kejian* did. This finding corresponds with the linguistic intuition that *yin'er* is perceived as formal. For the sake of the materials' easy readability, we refrained from using formal style. It is likely that the formal style associated with *yin'er* did not fit in with the relatively informal style of our experimental materials, and it's probable that this mismatch in style gave rise to a processing disadvantage for the *yin'er* condition. This assumption can explain the unexpected finding reported at the beginning of this paragraph, but needs to be tested in future, perhaps using the same experimental design with materials in a more formal style.

At the predicate region, no processing differences were observed between the content causal relations marked with *yin'er* and those marked with *suoyi*, indicating that the objective causal connective *yin'er* did not facilitate the processing of the content causal relations more than the underspecified causal connective *suoyi* did. It is likely that people generally expect objective content causal relations more than subjective epistemic causal relations, when there are no explicit indicators in the text that suggest the other way around. This account, in fact, is in line with the proposal raised in previous studies that readers might prefer to build the simplest possible discourse representation during interpretation (see Traxler et al. 1997b), in this case the objective relation. In the next section, I propose an investigation of another possibility, which concerns the genre-sensitivity of the underspecified causal connective *suoyi*.

³ At the subject region, no significant effects were observed.

At the connective region, processing differences were observed between the two content conditions discussed above. The objective causal connective *yin'er* took longer to read than the underspecified causal connective *suoyi*, with respect to two late processing measures (*first pass total gaze duration* and *regression path duration*). We cannot attribute the longer reading times observed for *yin'er* to processes of activating an objective representation of the relation. We cannot argue that *yin'er* activated the objective representation of the content causal relation at the connective region whereas *suoyi* did not. If that were the case, we should have observed shorter reading times at the predicate region in the *yin'er* condition compared to the *suoyi* condition. Neither can we attribute the longer reading times observed at *yin'er* to the fact that *yin'er* (1.50 per 10,000 words) is less frequent than *suoyi* (3.55 per 10,000 words). As discussed earlier, the frequency of the connective is not likely to be a crucial factor affecting late processing measures such as *first pass total gaze duration* and *regression path duration*. The longer reading times observed at *yin'er* in comparison with *suoyi* were also likely induced by the mismatch between the formal style of *yin'er* and the relatively informal style of our experimental materials.

6.4 Discussion and topics for future research

6.4.1 Variation in connective use: scalar view or categorical perspective?

The two corpus-based studies presented in the current dissertation presuppose that the distributive patterns of causal connectives over different contexts relate to their intrinsic meanings with respect to subjectivity. A tricky issue, however, concerns the indication from our corpus-based studies as well as from European studies (Pander Maat & Sanders 2000; Pander Maat & Degand 2001; Degand & Pander Maat 2003; Stukker, Sanders & Verhagen 2008, 2009) that connective use cannot be described with reference to clearly delineated categories of causality, e.g., domains of use. Usually, a causal connective can be used across different domains of causality that are associated with different degrees of subjectivity. In other words, connective use is not strictly categorical.

Two different approaches have been proposed to describe connective use, searching for a balance point between the relative stability and the restricted flexibility of connective use. One is the scalar approach, which proposes the characterization of connective meanings as positions on a scale of subjectivity (Pander Maat & Degand 2001; Pit 2006, 2007; Degand & Pander Maat 2003). According to the scalar approach, “the distribution of each causal connective occupies a continuous area on the scale of speaker involvement” (Pander Maat & Degand 2001:230). This stipulated continuous area, to some extent, accounts for the flexibility of a connective’s meaning and the overlap between

the usages of different causal connectives. The scalar approach “allows for a certain amount of variation in degrees of subjectivity ‘tolerated’ by the respective causal connectives” (Stukker, Sanders & Verhagen 2009: 156).

The other approach, the prototype-based categorical approach (see Stukker & Sanders 2012; Stukker, Sanders & Verhagen 2008, 2009), seeks to explain the patterns of variation and stability with reference to more general cognitive capacities. Starting from the cognitive phenomenon that conceptual categories display prototype structure, Stukker et al. (2008, 2009) propose an analysis of connective meaning and use in terms of ‘prototypicality’ (Rosch 1973) on the basis of the connective’s distributional frequency. A more frequent use of a type is assumed to result in a more entrenched representation, which, in turn, is the more prototypical member of a category; the less frequent type of use is taken to be the more peripheral (non-prototypical) member of the same category, related to the prototype by way of conceptual affinity (Stukker & Sanders 2012; Stukker et al. 2009). This prototype perspective on connective use provides an explanation for differences between causal connectives with reference to their respective prototypical meanings. At the same time, this approach excels in accounting for the flexibility of connective use with reference to the property of the prototype structure, the structure consisting of a dominant area surrounded by a less salient periphery.

Results of the present dissertation conform to the scalar view, in the sense that we did not find any single Chinese causal connective that covers a discontinuous area on the scale of subjectivity. For example, we found that a connective expressing causal relations in the non-volitional content and the epistemic domains was also able to express causal relations in the volitional content domain. However, several Chinese causal connectives entirely overlap with each other on the scale of subjectivity. Only *kejian* has a restricted usage, i.e., subjective uses. All other causal connectives occupy an identical stretch of area on the scale, e.g., ranging from the objective content domains to the subjective epistemic domain.⁴ Clearly, the scalar approach does not allow us to explain the fundamental issue of why several causal connectives coexist at the same level in a given language. According to Pander Maat and Degand (2001: 231), the primary reason for a language to have more than one causal connective is to be able to express several levels of speaker involvement.

The prototype-based approach provides a solution to the above problem, as this approach highlights the differences between connectives in distributional frequency. The more frequent type of use reflects the prototypical features of meaning, whereas the less frequent type of use

⁴ The speech-act domain of use was rarely found in our data, so we are not able to discuss the speech-act use in depth.

reflects non-prototypical features of meaning. With this approach, we can easily distinguish Chinese causal connectives, such as *yin'er* and *yushi* (both translated as 'so/therefore'), from each other: non-volitionality is the prototypical meaning of *yin'er*, and volitionality is the prototypical meaning of *yushi*. Meanwhile, representing connectives' semantics in the prototype structure (see the discussion on 'radial set models' or 'schematic networks' in Geeraerts & Cuyckens 2007) can faithfully reveal that connectives' meanings are not restricted to their respective prototypical cores. That is, with the help of the prototype structure, the flexibility of connective use is also well accounted for. Hence, the results reported in this dissertation seem to be in favor of the prototype-based categorical approach, because this approach produces a higher level of descriptive accuracy than the scalar approach. Furthermore, the prototype-based categorical approach is preferable because it aims to represent the lexical semantics of causal connectives at a more general level, on a par with human categorization of other concepts such as "color" and "bird". With this approach, we may expect to relate usage patterns of language to the cognitive structures associated with human categorization (see Stukker & Sanders 2012).

In order to ascertain the extent to which a Chinese causal connective displays prototype characteristics, future research is needed to explore the conceptual links between non-prototypical meanings/uses and their prototypical core. First, following the methodology proposed in studies on a few European languages (Sanders & Spooren 2013; Stukker & Sanders 2012; Stukker, Sanders & Verhagen 2009), we can compare between natural occurrences of epistemic causal relations marked with a causal connective that is prototypically subjective (e.g., *kejian*) and those marked with a causal connective that is prototypically objective (e.g., *yin'er*), in relation to the subjective elements involved in the segments, such as epistemic markers (e.g., *I think*, *perhaps*) and evaluative markers. We can hypothesize, for example, that *yin'er* co-occurs more often with epistemic markers such as 'I think' than *kejian*, given that the uses of epistemic markers function to objectify the causal relation by putting the SoC on stage. Other comparisons and hypotheses can be made along the same line of reasoning, in order to test the extent to which the non-prototypical uses are restricted by the respective prototypes.

A second way to investigate the conceptual relation between prototypical and non-prototypical use of a connective concerns the defeasibility of the consequent clauses introduced by connectives in either their prototypical or their non-prototypical context. The Chinese literature (Li & Liu 2004) mentions that, for an epistemic causal relation such as (11), it is possible to cancel out the consequent with a subsequent clause, e.g., *but you didn't put him into an important position*, whereas the cancelation is not possible if an objective causal connective *youyu* is used in the place of the

subjective causal connective *jiran* (see (12)). The symbol # in (12) indicates that the sentence is not interpretable.

- (11) ***Jiran*** *ta you nengli, name jiu yinggai zhong yong ta, keshi ni que bu zhong yong ta.* (Li & Liu 2004: 127)
 ‘***Jiran*** he is a capable man, we should put him into an important position, but you didn’t put him into any important position.’
- (12) # ***Youyu*** *ta you nengli, name jiu yinggai zhong yong ta, keshi ni que bu zhong yong ta.* (Li & Liu 2004: 128)
 ‘***Youyu*** he is a capable man, we should put him into an important position, but you didn’t put him into any important position.’

From (11) and (12), we can infer that the connective *youyu* might have brought some objectiveness (i.e., its prototypical core) to the epistemic causal relation, rendering it somewhat factual and hence infeasible. This is the so-called ‘prototype effect’ (Geeraerts & Cuyckens 2007): even in a non-prototypical context, a lexical item (e.g., a connective) still reflects features of its prototypical use. Sanders and Spooren (2013) and Stukker, Sanders and Verhagen (2008, 2009) have also discussed cases of connectives importing elements of their prototypical use into their non-prototypical context. The pattern “Reason connective P, Q, but not Q” can be developed into a substitution test in future research, with which we can assess an objective connective’s contribution to the interpretation of subjective causal relations. We may hypothesize that due to prototype effects, objective causal connectives systematically objectify epistemic causal relations to the extent that the epistemic relation in question becomes infeasible. We may also develop a way to alter the pattern, so as to make it applicable to result connectives as well, e.g., “P, result connective Q, but not Q”.

6.4.2 Subjectivity and correlatively used causal connectives

Another interesting topic concerns a special characteristic of the Chinese language in its use of connectives. Connectives can be used in a correlative way in Chinese discourse, for example, in the form of “*Yinwei* P, *suoyi* Q”. This characteristic is not unfamiliar to European linguists, since double marking (*Because* P, *therefore* Q) used to be the old English norm, which was later taken over by the Modern English norm of single marking (Traugott 1989). The fact that the Chinese language keeps both ways of causality marking at the same time provides us with an opportunity to investigate whether subjectivity functions as one of the constraints regulating the way in which a reason connective can be combined with a result connective. Will an

objective reason connective such as *youyu* impose a restriction upon the degree of subjectivity of the upcoming result connective, requiring it to be underspecified or to be objective as well? It is not possible to draw any statistically valid answer yet on the basis of the present research, because in the present sample of linguistic data the combinatory usages were rare. Future research is expected to pool and zoom in on more correlative uses to address this question.

However, the present study allows us to describe a tentative tendency. Out of forty-five double-marked cases, twenty-six were connected with *suoyi* and twelve were connected with *yinci*. The uses of other result connectives were rare (*yin'er*: 3, *yushi*: 4, *kejian*: 0). It seems that speakers prefer to use reason connectives together with underspecified result connectives *suoyi* and *yinci*. It is worthwhile to investigate whether this tendency can be ascertained by future research on the basis of a larger quantity of data, because it might have to do with speaker versus listener economy (Knott & Sanders 1998). For the sake of speaker economy, underspecified result connectives are easy to use, as they fit in with reason connectives of different degrees of subjectivity. For the sake of listener economy, specified result connectives are preferred because they add to the reader's knowledge about the overall degree of the unfolding causal relation that is intended by the speaker, on top of the information provided by the reason connective. Do Chinese speakers indeed prefer to choose a result connective that is general in use (i.e., an underspecified causal connective) to combine with a preceding reason connective, when a correlative usage is necessary under the communicative situation? If so, readers are required to draw implicatures to judge between candidate relational interpretations, which are not disambiguated on the basis of the information provided by the preceding reason connective alone. If this tendency is indeed borne out, then it is a case of speaker economy outdoing listener economy: speakers choose the most general forms to express a relation, in which case readers are required to draw implicatures to disambiguate (see Knott & Sanders 1998).

Still, it is unclear what double-marking adds to the subjectivity of a relation as a whole, compared to a relation that is marked by a single causal connective. This question is raised on the basis of a fundamental presupposition: there must be a reason for the co-existence of both single-marking such as “*yinwei* P, Q” and double-marking such as “*yinwei* P, *suoyi* Q” in the same language. Future studies might examine the total amount of subjectivity of the double-marked coherence relations, taking the properties of both segments into account.

6.4.3 Cognitive processes associated with subjectivity

What cognitive processes make subjective connectives, or subjective information in general, relatively complex to process? A reasonable suggestion in the literature is that subjective information involves an additional proposition, a meta-representation of the speaker's or another person's beliefs or conclusions, which could be in the form of 'I/someone think(s)' (Canestrelli et al. 2013). It takes extra processing time to represent this additional proposition in mind, which is why it takes longer to read subjective information. Canestrelli et al. (2013) have tested this hypothesis on the basis of materials in Dutch, using only specified causal connectives. They compared three experimental conditions: objective causal relations marked with the objective causal connective *omdat* 'because' in the pattern of "Q, *omdat* P", epistemic causal relations marked with the subjective causal connective *want* 'because' in the pattern of "Q, *want* P", and epistemic causal relations marked with *want* and an extra epistemic marker *according to Peter* in the pattern of "*Volgens Peter* Q, *want* P" (see examples (13) – (15)).

- (13) *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.'
- (14) *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.'
- (15) *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.'

At the region directly following the connective (e.g., *ze was vier trappen*), it was found that *first fixation duration* (i.e., the duration of the first fixation on a region) was longer in epistemic causal relations without the epistemic marker *according to Peter* (e.g., (14)) compared to both content causal relations (e.g., (13)) and epistemic causal relations containing the epistemic marker (e.g., (15)). According to the authors, this result reveals that *want* triggers the same cognitive processes – probably a meta-representation of other's beliefs – as epistemic markers such as *according to Peter* do, which is why the observed processing disadvantage associated with *want* was canceled out by the presence of the preceding epistemic marker *according to Peter*.

In Mandarin Chinese, the situation is a bit different: it is not possible to combine an epistemic marker such as *according to Peter* or *I think* with the subjective causal connective *kejian* ‘so’ in the consequent clauses. Compare (16), (17), and (18).

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

As shown in (16), the underspecified causal connective *suoyi* allows the addition of the words *wo renwei* ‘I think’ right after it, whereas (17) illustrates that adding *wo renwei* ‘I think’ after the subjective connective *kejian* makes the originally appropriate sentence (18) unacceptable. How can we explain these linguistic intuitions? Previous studies suggest that *kejian* directly encodes the speaker’s perspective (see Chapter 3; see also Tao 2007 for evidence of *jian* indicating perspective in the context of existential/presentative constructions). Accordingly, it can be argued that the connective *kejian* is encoded with 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 such information, so adding ‘I think’ to the *suoyi* clause will not create any problems.

If this introspective explanation can be borne out empirically, we can provide further evidence for the “meta-representation account” proposed by Canestrelli et al. (2013). One way to test this idea is to set up a design with two experimental conditions: the epistemic relation marked with *kejian* and the epistemic relation marked with *suoyi wo renwei* ‘so I think’, which can be exemplified by examples (18) and (16), respectively. If it is the construction of the extra proposition ‘I think’ that increases the processing complexity 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. This experimental design can avoid problems associated with the scope of epistemic markers, which can be an

issue in the Dutch materials. It is not apparent whether “according to Peter” covers both P and Q or whether it covers only Q in the Dutch materials. In the Chinese materials, the scope of ‘I think’ is indisputable: it covers only the consequent segment. Moreover, in the Dutch materials, the condition containing the epistemic marker was constructed in free indirect style, in which the narrator introduces the perspective of other persons (see Sanders, Sanders & Sweetser 2009 for a discussion on the free indirect style). Possibly, it concerns blending of spaces (by which the speaker identifies with another person and sees through another person’s eyes), especially if the scope of “according to Peter” covers only Q. Using the Chinese materials as presented on the previous page, we can reduce the potential complexities associated with the free indirect speech.

In light of suggestions in the literature that language users build, and constantly update, a network of mental spaces when they communicate (Fauconnier 1998; Sanders, Sanders & Sweetser 2009), to represent the proposition ‘I/someone think(s)’ would involve setting up and instantiating a mental space to present the speaker’s or another person’s thoughts. Accordingly, we may presume that subjective causal connectives such as *kejian* or epistemic markers such as *I think* have the function of activating the relevant mental spaces, whereas objective causal connectives or underspecified causal connectives do not. Moreover, in all these cases, the reader’s knowledge of genre might also play a role. As we discussed earlier, it is reasonable to assume that readers have knowledge as to the degree of subjectivity that a certain genre is typically associated with. For example, readers are expected to know that opinion pieces typically involve one’s claims or conclusions, and novels typically involve objective descriptions of events in temporal sequence. In the present experiment, narrative mini-stories were used as materials, which probably have a bias towards the objective interpretation. Accordingly, we cannot exclude an alternative interpretation for the result that the objective causal connective *yin’er* did not facilitate the processing of causal relations in the content domain more than the underspecified connective *suoyi*. In section 6.3, we provided an explanation for this result: people generally expect objective content causal relations more than subjective epistemic causal relations, when there are no explicit indicators in the text that suggest otherwise. An alternative explanation that we can add here is the following: in both experimental conditions, the genre of the materials imposed a constraint – that the unfolding relation should be interpreted in the objective direction, and hence there is no need to set up or instantiate a mental space. This accounts for the fact that no processing differences were observed.

Is this account generalizable to the processing of content causal relations embedded in other genres? For example, what if we embed the set of materials in a text category characterized by argumentation and

persuasion, such as an opinion piece? We can hypothesize that in content causal relations containing either *yin'er* or *suoyi*, the argumentative flavor of a genre (e.g., opinion pieces) imposes a constraint: the relation should be interpreted in the subjective direction. In the condition with the objective causal connective *yin'er*, there should be another constraint. The connective itself imposes a constraint that points to the opposite direction: the unfolding relation is likely to be objective. Under the circumstances, which constraint is more decisive? Does genre knowledge play a minor role in the way we have speculated in section 6.2.2 (see Figure 1)? In the condition with the underspecified causal connective *suoyi*, there is only one constraint on the relational interpretation, i.e., the one imposed by the genre. In the case of opinion pieces, will readers first attempt to establish an epistemic causal relation based on their genre knowledge, e.g., setting up or instantiating related mental spaces? Or will they prefer to build the simplest possible representation (in this case a causal relation in the content domain), as proposed by Traxler et al. (1997a)? More experimental work is needed to address these questions.

6.5 Conclusion

The present dissertation reveals that subjectivity constrains the use of causal connectives in Chinese as well as in European languages, in spite of the differences in the exact ways these languages cut up the domain of causality in terms of subjectivity. The current subjectivity approach not only accounts for and corroborates the observations made in prior studies, but has also contributed to new findings on the use of causal connectives in Chinese discourse. With the present dissertation, we hope to have shown how a theory-driven and corpus-based approach to the categorization of causal connectives produces interesting results for further cross-linguistic comparison.

Moreover, we consider the organization of the lexicon of causal connectives as a window into conceptual categorization. The sometimes subtle objective-subjective distinctions proposed by discourse analysts on the one hand, and the processing effects discovered by psycholinguistic research on the other hand, correspond with each other and jointly reveal in what way subjectivity, as a general cognitive principle, affects the way in which we use and interpret causal connectives in our daily communication. Still, many questions remain unanswered. A closer integration of theoretical and corpus-linguistic work with psycholinguistic studies is expected to lead to significant progress in the field of studies on connective use and, on a higher level of abstraction, will provide new insights into the link between linguistic structure and conceptual structure.

REFERENCES

- Baayen, R. H. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- Baayen, R. H. 2011. Language R: Data sets and functions with “Analyzing Linguistic Data: A practical introduction to statistics”. R package version 1.4.
- Baayen, R. H., D. J. Davidson & D. M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59: 390-412.
- Barlow, M. & S. Kemmer. 2000. *Usage-Based Models of Language*. Stanford, CA: CSLI Publications.
- Bates, D., M. Maechler & B. Bolker. 2012. Lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0.
- Bhatia, V. K. 1993. *Analysing Genre. Language Use in Professional Settings*. London: Longman.
- Biber, D. 1989. A typology of English texts. *Linguistics* 27: 3-43.
- Biber, D. 1993. Representativeness in corpus design. *Literature and Linguistic Computing* 8(4): 243-257.
- Biq, Yung-O. 1995. Chinese causal sequencing and *yinwei* in conversation and press reportage. *Proceedings of the twenty-first annual meeting of the Berkeley Linguistics Society: Special Session on Discourse in Southeast Asian Languages*, 47-60.
- Bloom, L., M. Lahey, L. Hood, K. Lifter & K. Fiess. 1980. Complex sentences: acquisition of syntactic connectives and the semantic relations they encode. *Journal of Child Language* 7: 235-261.
- Britton, B. K. 1994. Understanding expository text. Building mental structures to induce insights. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics*, 641-674. San Diego, CA: Academic Press.
- Canavan, A. & G. Zipperlen. 1996. *CALLFRIEND Mandarin Chinese-Mainland dialect*. Philadelphia, PA: Linguistic Data Consortium.
- Canestrelli, A. R. 2013. *Small words, big effects? Subjective versus objective causal connectives in discourse processing*. Ph.D. dissertation, Utrecht University.
- Canestrelli, A. R., W. M. Mak & T. J. M. Sanders. 2013. Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements. *Language and Cognitive Processes* 28(9): 1394-1413.
- Cantos, P. 2009. *Statistical methods in language and linguistic research*. Sheffield: Equinox.

- Chao, Y. R. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Chinese Academy of Social Sciences (Eds.). 2002. *Contemporary Chinese dictionary (Chinese-English edition)*. Beijing: Foreign Language Teaching and Research Press.
- 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, Tiburg University.
- Dancygier, B. 2009. Causes and consequences: Evidence from Polish, English, and Dutch. In Sanders & Sweetser (2009), 91-118.
- Degand, L. & H. Pander Maat. 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*, 175-199. Utrecht: LOT.
- Deng, Yuhui. 2007. Guobiao 'yinci' he 'yin'er' de yongfa bianxi [Analysis of the usage of *yinci* 'therefore' and *yin'er* 'consequently']. *Journal of Guangzhou University (Social Science Edition)* 6(8): 79-82.
- De Smet, H. & J. C. Verstraete. 2006. Coming to terms with subjectivity. *Cognitive Linguistics* 17(3): 365-392.
- Dik, S. C. 1989. *The theory of functional grammar*. Dordrecht: Foris.
- Ding, Shengshu & Shuxiang Lü (Eds.). 1999. *Xiandai hanyu yufa jianghua [A speech on the grammar of Modern Chinese]*. Beijing: The commercial press.
- Dong, Weiguang. 1999. Jushi zai fuju fenlei zhong de diwei [The role of sentence potential in the classification of compound sentences]. *Linguistics study* 2: 37-40.
- Ducrot, O. 1980. Essai d'application: MAIS – les allusions à l'énonciation – délocutifs, performatifs, discours indirect. In P. Herman (Ed.), *Le langage en context: Etudes philosophiques et linguistiques de pragmatique*, 487-575. Amsterdam: John Benjamins.
- Evers-Vermeul, J. 2005. *The Development of Dutch Connectives: Change and acquisition as windows on form-function relations*. Ph.D. dissertation, Utrecht University.
- Evers-Vermeul, J., L. Degand, B. Fagard & L. Mortier. 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. & T. J. M. Sanders. 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. & T. J. M. Sanders. 2011. Discovering domains – On the acquisition of causal connectives. *Journal of Pragmatics* 43(6): 1645-1662.

- Fagard, B. & L. Degand. 2010. Cause and subjectivity, a comparative study of French and Italian. *Linguisticae Investigationes* 33(2): 179-193.
- Fan, Xiao. 1998. *Hanyu juzi de leixing [The types of Chinese sentences]*. Taiyuan: Shuhai.
- 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*, 251-279. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Feng, Guangwu. 2006. Yuyan de zhuguanxing jiqi xiangguan yanjiu [A critical survey of the studies on linguistic subjectivity]. *Shandong Foreign Language Teaching Journal* 5: 26-33.
- Field, A. 2009. *Discovering statistics using SPSS (third edition)*. London: Sage.
- Finegan, E. 1995. Subjectivity and subjectivisation: an introduction. In D. Stein & S. Wright (Eds.), *Subjectivity and subjectivisation: Linguistic perspectives*, 1-24. Cambridge: Cambridge University Press.
- Fleischman, S. 1990. *Tense and narrativity. From medieval performance to modern fiction*. Austin: University of Texas Press.
- Fludernik, M. 1993. *The fictions of language and the languages of fiction: The linguistic representation of speech and consciousness*. London: Routledge.
- Ford, C. E. 1993. *Grammar in interaction: Adverbial clauses in American English conversations*. Cambridge: Cambridge University Press.
- Frohning, D. 2007. *Kausalmarker zwischen pragmatik und kognition: Korpusbasierte analysen zur variation im Deutschen*. Tubingen: Niemeyer.
- Gao, Li & Xu Wen. 2012. Yuyan 'zhuguanxing' yanjiu de duowei shijiao [Studying linguistic subjectivity from multiple perspectives]. *Foreign Language Education* 33(6): 11-14.
- Gao, Mingkai. 1948. *Hanyu yufa lun [On Chinese grammar]*. Shanghai: Kaiming.
- Gao, Zailan. 2013. Qian, houzhi 'yinwei' de yinxian ji gongneng chayi [On the omission of the pre-posed and post-posed 'yinwei' and the functional differences between them]. *Chinese Linguistics* 2: 57-65.
- Geeraerts, D. & H. A. C. Cuyckens (Eds.). 2007. *The Oxford handbook of cognitive linguistics*. Oxford: Oxford University Press.
- Gernsbacher, M. A. 1990. *Language comprehension as structure building*. Hillsdale, NJ: Erlbaum.
- Graesser, A. C., M. Singer & T. Trabasso. 1994. Constructing inferences during narrative text comprehension. *Psychological Review* 101: 371-395.
- Grice, H. P. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

- Guo, Jimao. 2006. 'Yushi' he 'suoyi' de yitong [A contrastive analysis between *yushi* and *suoyi*]. *Chinese Linguistics* 4: 27-34.
- Guo, Jimao. 2008. 'Yinwei suoyi' ju he 'jiran name' ju de chayi [A contrastive analysis between sentences with 'yinwei' and 'jiran']. *Chinese Language Learning* 3: 22-29.
- Haberlandt, K. F. 1982. Reader expectations in text comprehension. In J. Le Ny & W. Kintsch (Eds.), *Language and comprehension*, 239-249. Amsterdam: North-Holland.
- Haegeman, L. 1994. *Introduction to government and binding theory*. Oxford, etc.: Blackwell Publishers.
- Halliday, M. A. K. 1994. *An introduction to functional grammar*. London: Arnold.
- Halliday, M. A. K. & R. Hasan. 1976. *Cohesion in English*. London: Longman.
- Hobbs, J. R. 1983. Why is discourse coherent? In F. Neubauer (Ed.), *Coherence in natural language texts*, 29-70. Hamburg: Buske.
- Hobbs, J. R. 1985. *On the coherence and structure of discourse*. Stanford, CA: Stanford University, Center for the Study of Language and Information.
- Hsieh, Chia-Ling. 2008. Evidentiality in Chinese newspaper reports: subjectivity/objectivity as a factor. *Discourse Studies* 10: 205-229.
- Hu, Yushu. 1995. *Xiandai Hanyu [Modern Chinese] (6th edition)*. Shanghai: Shanghai Education Press.
- Huang, Chengwen. 1990. *Fuju [Compound sentences]*. Beijing: People's Education Press.
- Jiang, Tao. 2010. Jiran de yufahua guocheng [The grammaticalization of *jiran*]. *Journal of Language and Literature Studies* 13: 102-105.
- Keller, R. 1995. The epistemic 'Weil'. In D. Stein & S. Wright (Eds.), *Subjectivity and subjectivisation: Linguistic perspectives*, 16-30. Cambridge: Cambridge University Press.
- Kintsch, W. 1988. Role of knowledge in discourse comprehension. *Psychological Review* 95: 163-182.
- Kleijn, S., W. M. Mak & T. J. M. Sanders. 2011. Daardoor dus! Effecten van specificiteit en subjectiviteit op de verwerking van connectieven. *Toegepaste Taalwetenschap in Artikelen* 85(1): 103-112.
- Knott, A. & R. Dale. 1994. Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes* 18: 35-62.
- Knott, A. & T. J. M. Sanders. 1998. The classification of coherence relations and their linguistic markers: An exploration of two languages. *Journal of Pragmatics* 30: 135-175.
- Koornneef, A. W. & T. J. M. Sanders. 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. & J. J. A. van Berkum. 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.
- Kristeva, J. 1989. *Language the Unknown: An Initiation into Linguistics*, trans. A. M. Menke. New York: Columbia University Press.
- Kuno, S. 1987. *Functional syntax: Anaphora, discourse and empathy*. Chicago and London: University of Chicago Press.
- Lakoff, G. 1987. *Women, fire and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- Lakoff, G. & M. Johnson. 1999. *Philosophy in the Flesh. The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.
- Langacker, R. 1985. Observations and speculations on subjectivity. In J. Haiman (Ed.), *Iconicity in syntax*, 109-150. Amsterdam: Benjamins.
- Langacker, R. 1987. *Foundations of cognitive grammar. Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, R. 1990. Subjectification. *Cognitive Linguistics* 1: 5-38.
- Levinson, S. C. 2000. *Presumptive meanings: the theory of generalized conversational implicature*. Cambridge, MA: MIT Press.
- Li, Fang, J. Evers-Vermeul & T. J. M. Sanders. 2013. Subjectivity and result marking in Mandarin: A corpus-based investigation. *Chinese Language and Discourse* 4(1): 74-119.
- Li, Jinxi & Shiru Liu. 1957. Hanyu fujū xīn tixī de lilun [A new system of Chinese complex sentences]. *Zhongguo Yuwen* 8: 476-490.
- Li, Jinxia. 2005. 'Hao' de yufahua yu zhuguanxing [The grammaticalization and subjectivity of 'hao']. *Chinese Teaching in the World* 1: 44-49.
- Li, Jinxia. 2011. Lun 'youyu' yu 'yinwei' de chayi [On the differences between *youyu* and *yinwei*]. *Chinese Teaching in the World* 25(4): 490-496.
- Li, Jinxia & Yun Liu. 2004. 'Youyu' yu 'jiran' de zhuguanxing chayi [The differences of *youyu* and *jiran* in subjectivity]. *Chinese Language* 2: 123-128.
- Li, Jinxia & Zhongling Wang. 2013. Lun 'yinwei' 'suoyi' dan yong shi xuanze qingxiang yu shiyong chayi [On the choice tendency and usage differences of *yinwei* and *suoyi* used alone]. *Studies in Language and Linguistics* 33(1): 19-26.
- Li, Shaoqun. 2012. Kejian de biaoji gongneng he yufahua guocheng [The function of *kejian* in discourse marking and the process of its grammaticalization]. *Journal of Northwest University (Philosophy and Social Sciences Edition)* 42(3): 177-179.
- Li, Wenshan. 2009. Cong yufahua de jiaodu kan 'youyu' yu 'jiran' de jufa weizhi chayi [A diachronic study of the difference of syntactic

- distribution between *youyu* and *jiran*]. *Journal of Weinan Teachers University* 24(4): 59-63.
- Li, Xiaojun. 2009. 'Cong'er', 'yin'er' de gongneng chayi jiqi lishi jieshi [The functional differences between 'cong'er' and 'yin'er']. *Chinese Language Learning* 1: 50-56.
- Li, Xiaojun & Xiaowei Tang. 2007. Yinguo lianci 'yin'er' 'cong'er' de cihuihua [The lexicalization of causal conjunctions 'yin'er' and 'cong'er']. *Journal of Huaibei Coal Industry Teachers College (Philosophy and Social Sciences)* 28(2): 38-42.
- Lim, Ni-Eng. 2011. From subjectivity to intersubjectivity: epistemic marker *wo juede* in Chinese. In Yun Xiao, Liang Tao & Hooi Ling Soh (Eds.), *Current issues in Chinese linguistics*, 265-300. New Castle: Cambridge Scholars Press.
- Liu, Jin. 2009. Yuyan zhuguanxing gainian tanxi [An explorative analysis of linguistic subjectivity]. *Journal of Xi'an International Studies University* 17(3): 39-42.
- Liu, Jin. 2011. Jin shinian guonei yuyan zhuguanxing he zhuguanhua yanjiu de jinzhan [The progress made in studies on subjectivity and subjectification in the decade]. *Journal of Guizhou Normal University (Social Science Edition)* 2: 131-135.
- Liu, Yahui & Xiaopeng Yao. 2011. Kejian de qingtaihua yu guanlianhua – jian lun hanyu liang lei shijueci de yanhua chayi [On the modalitization and relevantization of 'kejian']. *Hanyu Xuebao [Chinese Linguistics]* 4: 27-33.
- Liu, Zhenfeng. 1986. *Xiandai Hanyu Fuju [Complex sentences of modern Chinese]*. Tianjin: Tianjin People's Publishing House.
- Longacre, R. E. 1983. *The grammar of discourse*. New York: Plenum.
- Lu, Qinghe. 2000. Yushi yu shili chengjie ['Yushi' and logic connection]. *Journal of Yangzhou University (The Humanities and Social Science Edition)* 6: 40-44.
- Luo, Xiaosuo. 1999. *Xiandai hanyu yinlun [Introduction to modern Chinese]*. Kunming: Yunnan University Press.
- Lü, Shuxiang. 1979. *Hanyu yufa fenxi wenti [An analysis of Chinese grammar]*. Beijing: Commercial Publishing House.
- Lü, Shuxiang. 1982. *Zhongguo wenfa yaolue [The key elements of Chinese grammar]*. Beijing: The commercial press.
- Lü, Shuxiang. 1999. *Xiandai hanyu babai ci [Eight hundred words of Chinese]*. Beijing: Commercial Publishing House.
- Lü, Shuxiang & Dexi Zhu. 1952. *Yufa xiuci jianghua [A speech on grammar and rhetoric]*. Beijing: China Youth Publishing Group.
- Lyons, J. 1977. *Semantics*. Cambridge: Cambridge University Press.
- Lyons, J. 1981. *Language, meaning and context*. Bungay: Chaucer Press.

- Lyons, J. 1982a. Subjecthood and subjectivity. In M. Yaguello (Ed.), *Subjecthood and subjectivity: The status of the subject in linguistic theory*, 9-17. Paris: Ophrys.
- Lyons, J. 1982b. Deixis and subjectivity: loquor ergo sum. In R. J. Jarvella & W. Klein (Eds.), *Speech, place and action: Studies in deixis and related topics*, 101-124. Chichester / New York: John Wiley & Sons.
- Lyons, J. 1995. *Linguistic semantics: An introduction*. Cambridge: Cambridge University Press.
- Mak, W. M. & T. J. M. Sanders. 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. & S. A. Thompson. 1986. Relational propositions in discourse. *Discourse Processes* 9: 57-90.
- Mann, W. C. & S. A. Thompson. 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. & A. Hardie. 2012. *Corpus Linguistics*. New York: Cambridge University Press.
- McEnery, T. & A. Wilson. 2001. *Corpus Linguistics (2nd edition)*. Edinburgh: Edinburgh University Press.
- McEnery, T. & Z. Xiao. 2004. The Lancaster Corpus of Mandarin Chinese: A corpus for monolingual and contrastive language study. *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC)*, 1175-1178. Lisbon, May 24-30, 2004.
- Millis, K. K. & M. A. Just 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. & F. De Blijzer. 2000. On the processing of causal relations. In E. Couper-Kuhlen & B. Kortmann (Eds.), *Cause, Condition, Concession and Contrast: Cognitive and Discourse Perspectives*, 35-56. Berlin: Mouton de Gruyter.
- Noordman, L. G. M. & W. Vonk. 1997. The different functions of a conjunction in constructing a representation of the discourse. In J. Costermans & M. Fayol (Eds.), *Processing interclausal relationships. Studies in the production and comprehension of text*, 75-93. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Nuyts, J. 2001. Subjectivity as an evidential dimension in epistemic modal expressions. *Journal of Pragmatics* 33(3): 383-400.

- Núñez, R. 2007. Inferential statistics in the context of empirical cognitive linguistics. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson & M. J. Spivey (Eds.), *Methods in cognitive linguistics*, 87-118. Amsterdam: John Benjamins.
- Pander Maat, H. & L. Degand. 2001. Scaling causal relations and connectives in terms of Speaker Involvement. *Cognitive Linguistics* 12: 211-245.
- Pander Maat, H. & T. J. M. Sanders. 2000. Domains of use or subjectivity? The distribution of three Dutch causal connectives explained. In E. Couper-Kuhlen & B. Kortmann (Eds.), *Cause, Condition, Concession and Contrast: Cognitive and Discourse Perspectives*, 57-81. Berlin: Mouton de Gruyter.
- Pander Maat, H. & T. J. M. Sanders. 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.
- Pit, M. 2006. Determining subjectivity in text: the case of backward causal connectives in Dutch. *Discourse Processes* 41: 151-174.
- Pit, M. 2007. Cross-linguistic analyses of backward causal connectives in Dutch, German and French. *Languages in Contrast* 7: 53-82.
- Qiu, Juanjuan. 2006. Shi lun yinguo lianci 'yinci' de chansheng shidai he yuanyin [On the appearance of the causal conjunction *yinci* and the cause of its appearance]. *Journal of China West Normal University (Philosophy and Social Sciences)* 1: 40-43.
- Qu, Shaobing. 2002. 'Youyu' ju de yuyi pianxia [On the semantic preference of *youyu* clause]. *Chinese Language* 1: 22-24.
- Quirk, R., S. Greenbaum, G. Leech & J. Svartvik. 1985. *A comprehensive grammar of the English language*. New York: Longman.
- Rayner, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124(3): 372-422.
- Rayner, K. & S. C. Sereno. 1994. Eye movements in reading: Psycholinguistic studies. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics*, 57-82. New York: Academic Press.
- Recanati, F. 1993. *Direct Reference: From Language to Thought*. Oxford: Blackwell.
- Recanati, F. 2004. *Literal Meaning*. Cambridge: Cambridge University Press.
- Renkema, J. 2004. *Introduction to discourse studies*. Amsterdam / Philadelphia: Benjamins.
- Rong, Huali. 2011. Hanyu yinguo fuju yanjiu zongshu [An overview of the study on Chinese cause-effect complex sentence]. *Journal of*

- Changchun Normal University (Humanities and Social Sciences)* 30(5): 47-51.
- Rosch, F. 1973. Natural categories. *Cognitive Psychology* 4: 328-350.
- Ross, J. R. 1970. On declarative sentences. In R. A. Jacobs & P. S. Rosenbaum (Eds.), *Readings in English transformational grammar*, 222-272. Washington: Georgetown University Press.
- Rutherford, W. 1970. Some observations concerning subordinate clauses in English. *Language* 46(1): 97-115.
- Sanders, J. 2010. Intertwined voices: Journalists' modes of representing source information in journalistic subgenres. *English Text Construction* 3(2): 226-249.
- Sanders, J. & G. Redeker. 1996. Perspective and the representation of speech and thought in narrative discourse. In G. Fauconnier & E. Sweetser (Eds.), *Spaces, Worlds and Grammars*, 290-317. Chicago: University of Chicago Press.
- Sanders, J., T. J. M. Sanders & E. Sweetser. 2012. Responsible subjects and discourse causality. How mental spaces and perspective help identifying subjectivity in Dutch backward causal connectives. *Journal of Pragmatics* 44(2): 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*, 31-46. Toulouse, France: Universite de Toulouse-le-Mirail.
- Sanders, T. J. M., J. Sanders & E. Sweetser. 2009. Causality, cognition and communication: A mental space analysis of subjectivity in causal connectives. In Sanders & Sweetser (2009), 19-59.
- Sanders, T. J. M. & W. Spooren. 2009. Causal categories in discourse – Converging evidence from language use. In Sanders & Sweetser (2009), 205-246.
- Sanders, T. J. M. & W. Spooren. 2013. Exceptions to rules: a qualitative analysis of backward causal connectives in Dutch naturalistic discourse. *Text and Talk* 33(3): 377-398.
- Sanders, T. J. M. & W. Spooren. Submitted. Causality and subectivity in discourse: The meaning and use of causal connectives in spontaneous conversation, chat interactions and written text.
- Sanders, T. J. M., W. Spooren & L Noordman.1992. Towards a taxonomy of coherence relations. *Discourse Processes* 15: 1-35.
- Sanders, T. J. M., W. Spooren & L. Noordman. 1993. Coherence relations in a cognitive theory of discourse representation. *Cognitive Linguistics*

- 4(2): 93-133.
- Sanders, T. J. M. & E. Sweetser (Eds.). 2009. *Causal categories in discourse and cognition*. Berlin: Mouton de Gruyter.
- Schmied, J. 1993. Qualitative and quantitative research approaches to English relative constructions. In C. Souter & E. Atwell (Eds.), *Corpus based computational linguistics*, 85-96. Amsterdam: Rodopi.
- Shao, Jingmin. 2001. *Xiandai Hanyu Tonglun [General theory on Modern Chinese]*. Shanghai: Shanghai Education Press.
- Shen, Jiaxuan. 2001. Yuyan de 'zhuguanxing' he 'zhuguanhua' [A survey of studies on subjectivity and subjectivisation]. *Foreign Language Teaching and Research* 33(4): 268-275.
- Shen, Jiaxuan. 2002. Can the disposal construction be disposed of? *Zhongguo Yuwen* 5: 387-399.
- Shen, Jiaxuan. 2003. Compound sentences in three conceptual domains: acting, knowing, and uttering. *Zhongguo Yuwen* 3: 195-204.
- Shen, Jiaxuan. 2008. San ge shijie [Physical world, mental world, and linguistic world]. *Foreign Language Teaching and Research* 40(6): 403-408.
- Shen, Jiaxuan. 2009. Hanyu de zhuguanxing he hanyu yufa jiaoxue [Subjectivity of Chinese and the teaching of Chinese grammar]. *Chinese Language Learning* 1: 3-12.
- Singer, M., R. Revlin & M. Halldorson. 1990. Bridging-inferences and enthymemes. In A. C. Graesser & G. H. Bower (Eds.), *Psychology of Learning and Motivation: Inferences and Text Comprehension*, 35-51. San Diego: Academic Press.
- Song, Zuoyan & Hongyin Tao. 2009. A unified account of causal clause sequences in Mandarin Chinese and its implications. *Studies in Language* 33(1): 69-102.
- Sperber, D., F. Clément, C. Heintz, O. Mascaro, H. Mercier, G. Origi & D. Wilson. 2010. Epistemic vigilance. *Mind and Language* 25(4): 359-393.
- Spooren, W. 1997. The processing of underspecified coherence relations. *Discourse Processes* 24: 149-168.
- Spooren, W. & T. J. M. Sanders. 2005. Chatting about causality. Presentation at the *International Workshop on Causality in Language and Cognition*. Utrecht institute of Linguistics OTS, Utrecht University, December 1-2, 2005.
- Spooren, W. & T. J. M. Sanders. 2008. The acquisition order of coherence relations: On cognitive complexity in discourse. *Journal of Pragmatics* 40: 2003-2026.
- Spooren, W., T. J. M. Sanders, M. Huiskes & L. Degand. 2010. Subjectivity and causality: A corpus study of spoken language. In S. Rice & J.

- Newman (Eds.), *Empirical and experimental methods in cognitive/functional research*, 241-255. Chicago: CSLI publications.
- Stein, D. & S. Wright. 1995. *Subjectivity and subjectivisation: Linguistic perspectives*. Cambridge: Cambridge University Press.
- Stukker, N. & T. J. M. Sanders. 2010a. Prototype effects in causal connectives across discourse contexts: deviant use as a rhetorical strategy. *Pre-publication for 34th International LAUD Symposium*, University of Koblenz-Landau.
- Stukker, N. & T. J. M. Sanders. 2010b. Diverging frequency effects in causal connectives across discourse contexts. A usage-based interpretation. *LAUD symposium Cognitive Sociolinguistics: Language variation in its structural, conceptual and cultural dimensions*. Essen: LAUD Linguistic Agency/Linse Linguistik-server.
- Stukker, N. & T. J. M. Sanders. 2012. Subjectivity and prototype structure in causal connectives: A cross-linguistic perspective. *Journal of Pragmatics* 44(2): 169-190.
- Stukker, N., T. J. M. Sanders & A. Verhagen. 2008. Causality in verbs and in discourse connectives: Converging evidence of cross-level parallels in Dutch linguistic categorization. *Journal of Pragmatics* 40: 1296-1322.
- Stukker, N., T. J. M. Sanders & A. Verhagen. 2009. Categories of subjectivity in Dutch causal connectives: a usage-based analysis. In Sanders & Sweetser (2009), 119-171.
- Su, Lily I-Wen. 2002. A cognitive exploration of Chinese connectives. In Lily I-Wen Su, Chinfa Lien & Kawai Chui (Eds.), *Form and function: Linguistic studies in honor of Shuanfan Huang*, 305-329. Taipei: Crane Publishing.
- Swales, J. 1990. *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Sweetser, E. 1990. *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.
- Tao, Hongyin. 2007. Subjectification and the development of special-verb existential/presentative constructions. *Language and Linguistics* 8(2): 575-602.
- Traugott, E. C. 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65: 31-55.
- Traugott, E. C. 1995. Subjectification in grammaticalization. In D. Stein & S. Wright (Eds.), *Subjectivity and subjectivisation: Linguistic perspectives*, 31-54. Cambridge: Cambridge University Press.
- Traugott, E. C. 2010. (Inter)subjectivity and (inter)subjectification: A reassessment. In H. Cuyckens, K. Davidse & L. Vandelanotte (Eds.), *Subjectification, intersubjectification and grammaticalization*, 29-71. Berlin / New York: Mouton de Gruyter.

- Traugott, E. C. 2011. Pragmatics and language change. In K. Allan and K. Jaszczolt (Eds.), *The Cambridge Handbook of Pragmatics*, 549-565. Cambridge: Cambridge University Press.
- Traxler, M. J., M. D. Bybee & M. J. Pickering. 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., A. J. Sanford, J. P. Aked & L. M. Moxey. 1997b. Processing causal and diagnostic statements in discourse. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23(1): 88-101.
- Trosborg, A. 1997. Text typology: Register, genre and text type. In A. Trosborg (Ed.), *Text typology and translation*, 3-24. Amsterdam, etc.: Benjamins.
- Tsai, Mei-chih. 1996. A discourse approach to causal sentences in Mandarin Chinese. *Proceedings of the 11th Pacific Asia conference on language, information, and computation (PA-CLIC 11)*, 93-98. Seoul: Korea.
- Van den Hoven, P. 1997. Niet wij doen het, maar het recht! *Tijdschrift voor Taalbeheersing* 9: 207-214.
- 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 University.
- Verhagen, A. 2005. *Constructions of intersubjectivity. Discourse, syntax, and cognition*. Oxford: Oxford University Press.
- Wang, Canlong. 2009. Yi ge binyu xiaowang de zhuguanxing biaojici – xiangshi [Xiangshi as a subjectivity marker on the brink of disappearance]. *Contemporary Linguistics* 11(1):35-46.
- Wang, Li. 1957. *Zhongguo xiandai yufa [Modern Grammar in China]*. Beijing: Zhonghua Book Company.
- Wang, Songmao. 1982. Lue tan yinguo guanxi de lianci [A brief discussion on conjunctions expressing the causal relation]. *Chinese Language Learning* 2: 22-30.
- Wang, Weihui. 2002. ‘Suoyi’ wanquan biancheng lianci de shidai [When did *suoyi* thoroughly change into a conjunction]. *Research in Ancient Chinese Language* 2: 30-33.
- Wang, Xiang. 1985. *Fuju, juqun, pianzhang [Complex sentences, sentence groups, and text]*. Xi’an: Shanxi People’s Publishing House.
- Wang, Yu-Fang. 2002. The preferred information sequences of adverbial linking in Mandarin Chinese discourse. *Text* 22(1): 141-172.
- Wang, Yu-Fang. 2006. The information structure of adverbial clauses in Chinese discourse. *Taiwan Journal of Linguistics* 4(1): 49-88.

- Wang, Yu-Fang, Pi-Hua Tsai & Ya-Ting Yang. 2010. Objectivity, subjectivity and intersubjectivity: Evidence from *qishi* ('actually') and *shishishang* ('in fact') in spoken Chinese. *Journal of Pragmatics* 42: 705-727.
- Wang, Zushu. 1999. Shi lun lianci 'yushi' de chengjie fangshi jiqi zuoyong [On the patterns of transition and functions of the conjunction *yushi*]. *Journal of Hubei University (Social Science)* 26(3): 69-71.
- Wen, Xue & Bei Huang. 2009. Jixing Chengdu fuci 'ji' de zhuguanhua [The subjectification of the degree adverb 'ji']. *Foreign Languages Research* 5: 9-13.
- Wong, May Lai-Yin. 2005. *Adverbial Clauses in Mandarin Chinese: A Corpus-based Study*. Ph.D. dissertation, Lancaster University.
- Xing, Fuyi. 1996. *Hanyu yufaxue [Chinese Grammar]*. Changchun: Northeast Normal University Press.
- Xing, Fuyi. 2001. *Hanyu fuju yanjiu [A study of Chinese complex sentences]*. Beijing: Commercial Publishing House.
- Xing, Fuyi. 2002. 'Youyu' ju de yuyi pianxiang bian [On the semantic preference of the pattern introduced by *youyu*]. *Chinese Language* 4: 337-342.
- Xu, Yizhong & Yiming Yang. 2005. Fuci 'dou' de zhuguanxing, keguanxing ji yuyong qiyi [The subjectivity, objectivity, and pragmatic ambiguity of Chinese adverb 'dou']. *Studies in Language and Linguistics* 25(3): 24-29.
- Xu, Zhichun, Zemin Du & Jingyi Peng. 1987. *Xin tixi yufa ji xiuci [A new system of grammar and rhetoric]*. Chengdu: Chengdu University of Science and Technology Press.
- Zeng, Liying. 2005. 'Wokan' yu 'nikan' de zhuguanhua [The subjectivisation of 'Wo-kan' and 'Ni-kan']. *Chinese Language Learning* 2: 15-22.
- Zhang, Baolin. 1996. Lianci de zaifenlei [Subdivision within conjunctions]. In M. Hu (Ed.), *Cilei wenti kaocha [An examination of syntactic categories]*, 431-445. Beijing: Beijing Language and Culture University Press.
- Zhang, Bin. 2008. *Xin bian xiandai hanyu [A new study of modern Chinese]*. Shanghai: Fudan University Press.
- Zhang, Hongmin. 1994. The grammaticalization of 'bei' in Chinese. In P. J. K. Li, C. R. Huang & C. C. J. Tang (Eds.), *Chinese languages and linguistics II: Historical linguistics*, 321-360. Taipei: Academia Sinica.
- Zhang, Qi. 2012. 'Kejian' de yufahua [The grammaticalization of 'kejian']. *Journal of Hubei University of Education* 29(5): 28-32.
- Zhang, Wangxi & Huimin Li. 2009. Duihua yujing yu fuci 'ke' de jiaohu zhuguanxing [Interlocution Contexts and Inter-subjectivity of the Adverb 'Ke']. *Chinese Teaching and Studies* 2: 1-8.

- Zhang, Xichang & Yaru Zhang. 2011. 'Yinwei' yinguo ju tanxi [An explorative analysis into the 'yinwei' clause]. *Modern Chinese* 2: 51-55.
- Zhang, Yan. 2012. Yinguo fuju guanlian biaoji jufa – yuyi yanjiu: jiyu 'jiaohu zhuguanxing' renzhiguan [A syntax-semantics interface study of causal connectives]. *Journal of Foreign Languages* 35(3): 42-50.
- Zhang, Ya-ru. 2008. Yushi ju de duo jiaodu fenxi [A comprehensive analysis of 'yushi' sentences]. *Journal of Yunnan Normal University (Teaching and Research on Chinese as a Foreign Language Edition)* 1: 51-57.
- Zhang, Zhigong. 1959. *Hanyu zhishi [Knowledge of Chinese]*. Beijing: People's Education Press.
- Zhao, Xin. 2003. 'Yinci, yushi, cong'er' de duo jiaodu fenxi [An analysis of 'yinci, yushi and cong'er' from multiple perspectives]. *Linguistic Research* 1: 26-34.
- Zhao, Xiufeng. 2010. Yuyan de zhuguanxing yanjiu gailan [A survey of studies on linguistic subjectivity]. *Foreign Language Education* 31(1): 21-26.
- Zhao, Zunli. 1985. *Xiandai hanyu fuju bianxi [Analysis of complex sentences of modern Chinese]*. Xi'an: Shanxi People's Publishing House.
- Zhu, Dexi. 2009. *Yufa jiangyi [Explanations of Chinese Grammar]*. 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.

SAMENVATTING IN HET NEDERLANDS

Subjectiviteit in Mandarijn Chinees: De betekenis en het gebruik van causale connectieven in geschreven tekst

1. Subjectiviteit

Taal wordt gebruikt om boodschappen te communiceren. Deze boodschappen kunnen feitelijke uitspraken over de wereld betreffen, zoals in zin (1), maar ook oordelen van de spreker, zoals in zin (2).

- (1) Beijing is de hoofdstad van China.
- (2) Beijing is een fantastische stad.

Zin (1) is een feitelijke uitspraak over een situatie in de wereld, namelijk dat Beijing de hoofdstad van China is. Zin (2) daarentegen bevat de mening van de spreker. Het verschil tussen (1) en (2) wordt getypeerd als een verschil in de mate van subjectiviteit: zin (2) is subjectiever dan zin (1).

De term *subjectiviteit* is in de jaren 1990 opgekomen. Lyons, Langacker en Traugott vertegenwoordigen drie belangrijke westerse visies op de notie subjectiviteit. Lyons (1977) beschouwt subjectiviteit als de 'zelf-expressie van de spreker' in de uiting, waarbij 'zelf' wordt gezien als de cognitie, perceptie, houding en intentie van de spreker ten opzichte van de propositionele inhoud van de uiting. Vanuit dit standpunt is zin (2) subjectief omdat de spreker zichzelf – zijn mening over Beijing – uitdrukt, terwijl dat bij (1) niet het geval is. De beschrijving van Traugott (1995, 2010) is vergelijkbaar met die van Lyons (1977). Traugott ziet subjectiviteit als een spreker-gerelateerde eigenschap; er is een relatie met de spreker en de overtuigingen en houdingen van de spreker. Langacker (1990) beschrijft subjectiviteit vanuit een ander standpunt. De spreker conceptualiseert objecten en gebeurtenissen vanuit een *on-stage* een *off-stage* perspectief. Een *off-stage* perspectief verwijst naar een situatie waarbij het perspectief van de spreker impliciet is in de uiting, zoals in (2). Een *on-stage* perspectief komt voor wanneer de spreker wel expliciet aanwezig is in de uiting (bijvoorbeeld in de zin *Ik vind Beijing een fantastische stad*). Een uiting met een *off-stage* perspectief is volgens Langacker subjectiever dan een uiting met een *on-stage* perspectief.

Langacker, Lyons en Traugott richtten zich op de mate van subjectiviteit in geïsoleerde zinnen. Subjectiviteit kan echter ook relevant zijn in relaties tussen zinnen (of zinsdelen), zogeheten coherenterelaties (Degand & Pander Maat 2003; Pander Maat & Sanders 2000, 2001; Sanders & Spooren 2009, 2013; Sweetser 1990; Zufferey 2012). Een coherenterelatie

is subjectief wanneer een spreker of een ander persoon verantwoordelijk is voor het leggen van het verband tussen de uitingen. De persoon die dit verband legt, wordt de ‘Subject of Consciousness’ (SoC) genoemd (Pander Maat & Sanders 2001). Als een relatie geen SoC bevat, dan is deze objectief. Het verschil tussen objectieve en subjectieve relaties kan geïllustreerd worden aan de hand van zin (3) en (4).

- (3) De temperatuur was onder min tien graden gedurende meer dan een maand. Als gevolg hiervan zijn veel ijsvogels doodgegaan.
- (4) De lichten in het huis zijn uit, dus niemand is thuis.

Zin (3) bevat geen SoC: de causale relatie is een oorzaak-gevolg relatie die zonder menselijke tussenkomst ontstaat en waarneembaar is in de wereld. De causale relatie in (3) is daarom objectief. De relatie in (4) daarentegen bevat wel een SoC: de causale relatie verbindt een argument en een conclusie in het hoofd van de spreker, die in deze uiting niet expliciet genoemd wordt. Zin (5) en (6) geven voorbeelden van zinnen waarin de SoC wel linguïstisch gerealiseerd wordt.

- (5) De lichten in het huis zijn uit, dus ik denk dat niemand thuis is.
- (6) De lichten in het huis zijn uit, dus niemand is thuis, dacht Jan.

In (5) is de spreker verantwoordelijk voor het leggen van de relatie tussen de twee zinsdelen en wordt hij ook expliciet genoemd (“*ik denk dat...*”). In zin (6) introduceert de spreker het perspectief van een ander. In dit voorbeeld is Jan de SoC die verantwoordelijk is voor de relatie en treedt de spreker slechts op als rapporteur. Toch is ook (6) subjectiever dan een zin als (3), aangezien de causale relatie tussen de twee zinsdelen kan worden toegeschreven aan een SoC en de relatie niet waarneembaar is in de werkelijkheid.

De domeintheorie van Sweetser (1990) maakt een soortgelijk onderscheid tussen objectieve en subjectieve relaties. Volgens deze theorie kunnen causale relaties uitgedrukt worden in verschillende domeinen: het content-domein, het epistemische domein en het speech-act-domein. Het content-domein verwijst naar objectieve relaties waarbij een gebeurtenis de oorzaak is van een andere gebeurtenis in de wereld, zoals in zin (3) en zin (7). Het epistemische domein verwijst naar relaties waarbij de redenering van de spreker wordt genoemd, zoals in zin (4). Het speech-act domein verwijst naar relaties die verbonden zijn op het niveau van illocuties: zin (8) geeft een voorbeeld waarin de spreker motiveert waarom hij een vraag stelt. Epistemische en speech-act-relaties zijn beide subjectief.

- (7) Jan kwam terug omdat hij van Marie houdt.
- (8) Komt Jan nog terug? Want Marie wacht op hem.

Binnen het content domein kan er een onderscheid worden gemaakt tussen volitionele en non-volitionele content-relaties (Pander Maat & Sanders 2001). In volitionele content-relaties zoals die in (7) spelen menselijke intenties een rol. In non-volitionele content-relaties zoals die in (3) spelen menselijke intenties geen rol.

2. Cognitieve relevantie van subjectiviteit voor de categorisering van causale connectieven

In veel talen bestaan meerdere connectieven om causale relaties uit te drukken. Zo kunnen causale relaties in het Nederlands onder andere uitgedrukt worden met de connectieven *want*, *omdat*, *dus* en *daardoor*. Studies naar causale connectieven in het Frans, Duits, Pools en Mandarijn hebben aangetoond dat ook deze talen meerdere connectieven hebben om causale relaties uit te drukken, maar dat deze connectieven niet altijd alle in dezelfde context gebruikt kunnen worden zonder dat de betekenis van de relatie verandert (Deng 2007; Guo 2006; X. Li 2009; Zhao 2003; voor Europese causale connectieven zie het overzicht in Stukker & Sanders 2012). Een voorbeeld hiervan is te zien in zin (9) en (10).

- (9) Het regende de hele dag hard, dus / daardoor was de picknick afgezegd.
- (10) Het was koud, dus / *daardoor deed Jan een trui aan.

Dus en *daardoor* kunnen beide de zinsdelen in (9) verbinden, maar in (10) kan alleen *dus* gebruikt worden. Hieruit kan worden opgemaakt dat de twee connectieven niet volledig identiek zijn. Een verklaring hiervoor is te vinden in de mate van subjectiviteit die de connectieven uitdrukken. Verschillende Europese studies hebben namelijk laten zien dat causale connectieven een verschillende mate van subjectiviteit kunnen hebben (Degand & Pander Maat 2003; Evers-Vermeul, Degand, Fagard & Mortier 2011; Pit 2003; Spooren, Sanders, Huiskes & Degand 2010; Zufferey 2012). Zin (10) maakt bijvoorbeeld duidelijk dat *daardoor* niet in volitionele contentrelaties kan voorkomen.

Subjectiviteit wordt verondersteld cognitief relevant te zijn, omdat het in meerdere Europese talen onze kennis van causaliteit en causale connectieven categoriseert en organiseert (Canestrelli 2013; Sanders & Spooren 2009; Stukker & Sanders 2012; Van Veen 2011). Psycholinguïstische experimenten en studies naar taalverwerving ondersteunen dit idee van cognitieve relevantie. Het is aangetoond dat causale relaties en connectieven

met verschillende gradaties van subjectiviteit anders worden verwerkt en geleerd. Zo laten psycholinguïstische experimenten zien dat objectieve causale relaties en connectieven sneller worden verwerkt dan subjectieve causale relaties en connectieven (Canestrelli 2013; Canestrelli, Mak & Sanders 2013; Noordman & De Blijzer 2000; Traxler, Bybee & Pickering 1997a; Traxler, Sanford, Aked & Moxey 1997b). En taalverwervingsonderzoek heeft aangetoond dat objectieve causale relaties eerder worden verworven dan subjectieve causale relaties (Evers-Vermeul 2005; Evers-Vermeul & Sanders 2011; Spooren & Sanders 2008; van Veen 2011). Deze bevindingen kunnen verklaard worden door de ‘subjectieve-complexiteitshypothese’ (Sanders 2005): subjectieve relaties zijn cognitief complexer dan objectieve relaties, waardoor subjectieve relaties later worden verworven en langzamer worden verwerkt.

Zoals andere talen vertoont ook het Chinees een uitgebreid scala aan causale connectieven. Gezien de cognitieve relevantie van de notie subjectiviteit, is in dit proefschrift onderzocht in hoeverre de subjectiviteitsbenadering gegeneraliseerd kan worden naar causale connectieven in Mandarijn Chinees.

3. Subjectiviteit in Chinese literatuur

In hoofdstuk 2 wordt Chinese literatuur over causale connectieven en causale coherentierelaties besproken. Voor deze literatuurbespreking zijn de volgende onderzoeksvragen geformuleerd:

- Vraag 1 Hoe definiëren Chinese taalwetenschappers subjectiviteit? Kunnen hun benaderingen gerelateerd worden aan westerse benaderingen van subjectiviteit?
- Vraag 2 Welke analytische categorieën zijn gebruikt in voorgaande studies naar Chinese causale connectieven? Kunnen deze categorieën gerelateerd worden aan analyses in termen van subjectiviteit, en zo ja, hoe?
- Vraag 3 Hanteren Chinese en Europese studies vergelijkbare methoden bij de studie van subjectiviteit in causaliteit?

In antwoord op onderzoeksvraag 1 laat de literatuurbespreking zien dat de definities van subjectiviteit in voorgaande studies gerelateerd kunnen worden aan de theorieën van Lyons (1977), Langacker (1990) en Traugott (1995, 2010). Chinese taalwetenschappers definiëren subjectiviteit, gebaseerd op de theorieën van deze drie westerse taalwetenschappers, in termen van ‘zelf-expressie’, spreker-gerelateerdheid en *on-/off-stage*-conceptualisatie. Een talige uiting wordt als subjectief beschouwd als er een link is met spreker-perspectief, spreker-affect (wat voornamelijk behandeld wordt in

termen van empathie) of epistemische modaliteit (het oordeel van de spreker over de waarschijnlijkheid dat de propositionele inhoud van de uitspraak waar is).

De literatuurbespreking laat tevens zien dat een groeiend aantal studies de subjectiviteitsbenadering toepast op de analyse van verschillende linguïstische fenomenen in het Mandarijn Chinees, waaruit blijkt dat de notie van subjectiviteit bruikbaar is om ook dit taalgebruik te beschrijven en te verklaren. Slechts een klein aantal studies heeft zich daadwerkelijk gericht op het analyseren van de mate van subjectiviteit van causale connectieven in het Chinees, zoals *jiran*, *yinwei* en *youyu* (alle te vertalen als *omdat*). In deze studies wordt vooral onderscheid gemaakt in termen van domeinen (Li 2011; Li & Liu 2004) en tussen beschrijvende en inferentiële causaliteit (Gao 2013; Guo 2008; Deng 2007). In antwoord op onderzoeksvraag 2 duiden deze bevindingen erop dat de subjectiviteitsbenadering voor connectiefgebruik gegeneraliseerd kan worden naar de categorisering van Chinese causale connectieven.

Om onderzoeksvraag 3 te beantwoorden is geïnventariseerd welke methoden gebruikt zijn in eerdere studies naar subjectiviteit in het Chinees taalgebruik. Hieruit kan worden geconcludeerd dat de huidige methoden veelal kwalitatief georiënteerd zijn, en dat uitbreiding met kwantitatieve en statistische benaderingen wenselijk is, zodat de generaliseerbaarheid en repliceerbaarheid van resultaten verbeterd worden. De methodologie kan verbeterd worden door meerdere causale connectieven tegelijkertijd in ogenschouw te nemen, door connectieven in natuurlijke data te analyseren in plaats van in geïsoleerde uitingen, door kwalitatieve benaderingen met kwantitatieve benaderingen te combineren, door meerdere indicatoren van subjectiviteit (naast de drie domeinen van Sweetser) in de analyses op te nemen, en een analytisch model met precieze definities en beslisregels voor de beoordeling van subjectiviteit te ontwikkelen.

Onderzoek naar Chinese causale connectieven heeft nog niet de relatie tussen genre en mate van subjectiviteit in ogenschouw genomen. In de westerse literatuur heeft dit onderwerp recentelijk meer aandacht gekregen (Sanders & Spooren 2009; Spooren, Sanders, Huiskes & Degand 2010; Stukker & Sanders 2012). Genres kunnen worden gedefinieerd in termen van de communicatieve functies die ze uitdrukken. Vaak wordt hierbij een onderscheid gemaakt tussen informatieve, narratieve, en argumentatieve teksten (Renkema 2004; Guo 2006). Hierbij wordt aangenomen dat argumentatieve teksten een hogere mate van subjectiviteit bevatten dan de andere twee typen, omdat argumentatieve teksten vaak argumenten bevat om lezers te overtuigen van bepaalde standpunten (Sanders 1997). Dit roept de vraag op of bepaalde connectieven voorkeur hebben voor specifieke genres, of dat het subjectiviteitsprofiel van connectieven verschilt per genre. In dit proefschrift is daarom de impact van genre meegenomen in de analyses.

4. Subjectiviteitsprofielen van Chinese causale connectieven

Hoofdstuk 3 en 4 beschrijven corpusonderzoek waarin bekeken is hoe causale connectieven in Mandarijn Chinees kunnen worden gecategoriseerd. Hierbij stonden de volgende onderzoeksvragen centraal:

Vraag 4 Tonen Chinese causale connectieven systematische variatie in hun mate van subjectiviteit? Zo ja, welke variatie? Zijn er op dit vlak taalspecifieke kenmerken?

Vraag 5 Variëren de subjectiviteitsprofielen van causale connectieven per genre?

Acht causale connectieven zijn onderzocht: vijf resultaat-connectieven die vertaald kunnen worden als *du*, namelijk *kejian*, *suoyi*, *yinci*, *yin'er* en *yushi*; en drie reden-connectieven die vertaald kunnen worden als *omdat/want*, namelijk *jiran*, *yinwei* en *youyu*.

Relaties in het corpus zijn geanalyseerd aan de hand van een analytisch model gebaseerd op Sanders en Spooren (2009) en bestaande uit vier indicatoren van subjectiviteit: domein, propositionele attitude, de aanwezigheid van een SoC en de identiteit van de SoC. De waarden van deze indicatoren staan in Tabel 1.

Tabel 1. Model voor de subjectiviteitsanalyse met bijbehorende variabelen en waarden.

Variabelen	-Subjectiviteitswaarden..... +
Domein	Non-vol. content < Vol. content < Speech act / Epistemisch
Propositionele attitude	Fysiek feit < Mentaal feit < Speech act / Oordeel
Aanwezigheid van SoC	Geen SoC < Expliciete SoC < Impliciete SoC
Identiteit van SoC	Karakter SoC < Huidige spreker SoC < Auteur SoC

In antwoord op onderzoeksvraag 4 blijkt uit het corpusonderzoek dat zes Chinese causale connectieven systematische variatie vertonen in hun mate van subjectiviteit. Drie reden-connectieven (namelijk *jiran*, *yinwei*, *youyu*) en drie resultaat-connectieven (namelijk *kejian*, *yin'er* en *yushi*) vertonen subjectiviteitsprofielen die van elkaar verschillen en consistent zijn over genres. Tabel 2 geeft de subjectiviteitsprofielen van deze connectieven weer.

Tabel 2. Subjectiviteitsprofielen van robuuste connectieven

Connectief	Domein	Propos. attitude	Aanwezigheid van SoC	Identiteit van SoC
<i>kejian</i>	+epistemisch	+oordeel	+impliciet	+spreker
<i>jin'er</i>	+non-vol. content	+fysiek feit	+geen SoC	n.v.t.
<i>yushi</i>	+vol. content	+fysiek feit	+expliciet	+karakter
<i>jiran</i>	+epistemisch +speech act	+oordeel +speech act	+impliciet	+spreker
<i>yinwei1</i>	+epistemisch	+oordeel	+impliciet	+spreker
<i>yinwei2</i>	+content	+feiten	+geen SoC	n.v.t.
<i>youyu</i>	+non-vol. content	+fysieke feiten	+geen SoC	n.v.t.

N.B. *Yinwei1* verwijst naar het gebruik van *yinwei* aan het hoofd van een achteropgeplaatste bijzin; *yinwei2* verwijst naar het gebruik van *yinwei* in een vooropgeplaatste bijzin.

Het gebruik van *kejian* 'dus' is beperkt tot subjectieve relaties in het epistemische domein. Het connectief wordt soms ook gebruikt in het speech-act-domein, maar komt nooit voor in het content-domein. *Kejian* wordt gebruikt in sterk subjectieve contexten met oordelen en impliciete-spreker-SoC's. Deze mate van specialisatie in het epistemische domein was nog niet bekend vanuit andere talen. Het Nederlands, Frans en Duits bevatten weliswaar connectieven die voornamelijk gebruikt worden in epistemische relaties (zoals het Nederlandse *dus*), maar die connectieven kunnen daarnaast in het volitionele contentdomein gebruikt worden, al komt dit minder vaak voor.

De andere vijf connectieven kunnen ieder gebruikt worden om verschillende gradaties van subjectiviteit uit te drukken, variërend van content- tot epistemische relaties. De vijf connectieven verschillen echter wel in hun distributionele frequentie. *Jiran* 'omdat/want' wordt gebruikt in epistemische en speech-actrelaties met oordelen, speech acts en impliciete-spreker-SoC's. *Youyu* 'omdat/want' wordt daarentegen gebruikt in non-volitionele content relaties (zonder SoC) tussen twee feiten. *Yin'er* 'dus' lijkt op *youyu*. Dit connectief wordt vooral gebruikt in het non-volitionele contentdomein, waar geen SoC aanwezig is. *Yushi* 'dus' drukt volitionele contentrelaties uit, en komt vaak voor met een expliciet karakter-SoC. Het gebruik van *yinwei* 'omdat' hangt af van de positie in de zin. *Yinwei* in een achteropgeplaatste bijzin (*Q yinwei P*) drukt vaak epistemische oordelen van

een impliciete spreker uit. *Yinwei* in een vooropgeplaatste bijzin (*yinwei P, Q*) drukt vaak contentrelaties tussen twee feiten uit.

Bovenstaande bevindingen dragen op verschillende manieren bij aan de huidige kennis over causale connectieven in Mandarijn Chinees taalgebruik. De bevinding dat *jiran* ‘omdat/want’ vaker gebruikt wordt om subjectieve relaties uit te drukken dan *youyu* ‘omdat/want’ komt overeen met de conclusie van het corpusonderzoek gerapporteerd in Li en Liu (2004). Het huidige corpusonderzoek breidt de bevindingen van Li en Liu (2004) uit, omdat naast de indicator domein nog drie andere variabelen zijn geanalyseerd. Hierdoor kan een meer gedetailleerde beschrijving van de subjectiviteitsprofielen van *jiran* en *youyu* worden gegeven.

Ten tweede bleken enkele van de introspectieve observaties van Xing (2001) bevestigd te worden. De huidige corpusonderzoeken ondersteunen Xing's (2001) claim dat *kejian* inferentieel wordt gebruikt, terwijl *yin'er* beschrijvend is. Door dit inferentiële gebruik komt *kejian* vooral voor in relaties die subjectiever zijn dan relaties gemarkeerd door *yin'er*.

Ten derde komen de bevindingen over de twee vormen van *yinwei* overeen met de functionele benadering van Song en Tao (2009). Volgens hen heeft een achteropgeplaatste *yinwei*-zin een interactionele functie om te bemiddelen bij potentiële problemen of deze te verzachten, terwijl vooropgeplaatste *yinwei*-zinnen worden gebruikt om informatie te delen. Het verschil in mate van subjectiviteit kan verklaren waarom de twee vormen voor verschillende functies gebruikt worden. De interactionele functie past bij het subjectieve karakter van achteropgeplaatste *yinwei*-zinnen. Wanneer de spreker daarentegen informatie wil delen met een ander, geeft hij de voorkeur aan de variant met een objectieve betekenis: de zinsinitiële *yinwei*. Er zijn uitzonderingen op deze vorm-functie-interactie: soms wordt de objectieve variant bijvoorbeeld gebruikt om een subjectieve relatie uit te drukken. Dit kan verklaard worden vanuit de retorische doeleinden van de spreker (Sanders & Spooren 2013, Stukker, Sanders & Verhagen 2008, 2009; Van den Hoven 1997). De spreker kan vanwege deze retorische doeleinden kiezen voor een objectief connectief om zijn claim overtuigender over te laten komen. Het retorische gebruik van causale connectieven is een interessant onderwerp voor vervolgonderzoek.

Ten vierde laat ons corpusonderzoek zien dat *yushi* ‘dus’ over het algemeen volitionele content relaties uitdrukt, maar ook gebruikt kan worden om epistemische relaties met inferenties en conclusies uit te drukken. Deze bevinding spreekt de claim van Lu (2000) tegen, die op basis van intuïtie en generalisatie concludeert dat *yushi* niet gebruikt kan worden om inferentiële relaties uit te drukken. Dit onderstreept het belang om talige intuïties te combineren met natuurlijke data. Dat een connectief een voorkeur heeft voor volitionele relaties, is overigens niet specifiek voor het Chinees. Het Nederlandse *daarom* lijkt op dit punt vergelijkbaar met het

Chinese *yushi*. De Chinese bevindingen onderstrepen wel de meerwaarde van Stukker et al.'s (2008) toevoeging van de notie volitionaliteit aan de domeinentheorie van Sweetser (1990).

5. Genre-gevoelige Chinese causale connectieven

Een belangrijke bevinding van dit proefschrift betreft de genre-gevoeligheid van de subjectiviteitsprofielen van connectieven (vergelijk onderzoeksvraag 5). Uit de corpusonderzoeken is gebleken dat de subjectiviteitsprofielen van twee resultaat-connectieven – *yinci* en *suoyi*, beide ‘dus’ – genre-sensitief zijn. Xing (2001) stelde dat *yinci* en *suoyi* alleen gebruikt worden in beschrijvende relaties zoals volitionele contentrelaties, maar het huidige onderzoek levert hier aan aanvulling op. De twee connectieven laten in het narratieve genre inderdaad een voorkeur voor beschrijvende relaties zien, maar in het argumentatieve en informatieve genre worden ze vaak gebuikt voor epistemische relaties met oordelen en met impliciete-spreker-SoC's.

Deze resultaten hebben twee implicaties. Ten eerste laat dit onderzoek het belang zien van genre-gevoeligheid. Xing (2001) onderzocht alleen het narratieve genre, waardoor zijn studie slechts een deel van het gebruik van *yinci* en *suoyi* liet zien. Dit onderstreept het belang van het meewegen van genre in onderzoek naar het gebruik van linguïstische elementen zoals connectieven.

Ten tweede leiden de bevindingen over genre-gevoeligheid tot een vraag: maken de subjectiviteitsprofielen deel uit van de semantische kenmerken van connectieven of worden ze (gedeeltelijk) bepaald door de pragmatiek van de context? Aangezien de resultaten zowel genre-gevoelige als robuuste profielen aantonen, is het waarschijnlijk dat in sommige connectieven de mate van subjectiviteit is vastgelegd in het lexicon, terwijl dat voor andere connectieven niet geldt. De vijf connectieven die hetzelfde profiel behouden over genres, de *gespecificeerde causale connectieven*, hebben waarschijnlijk een vaste subjectiviteitswaarde die gespecificeerd is in het lexicon. Ervan uitgaand dat genre-kennis en de specificatie van connectieven beide invloed hebben op het interpretatieproces van coherentierelaties, kunnen we ervan uitgaan dat gespecificeerde causale connectieven een sterke invloed hebben op dit interpretatieproces, en dat genre-kennis daardoor een kleinere rol speelt. Bij de twee genre-gevoelige connectieven *yinci* en *suoyi* ligt dit anders. De subjectiviteit van deze connectieven is ondergespecificeerd in het lexicon, waardoor de mate van subjectiviteit van de coherentierelaties waarin ze voorkomen vooral bepaald wordt door het genre.

6. De invloed van subjectiviteitsprofiel op verwerking

Traxler et al. (1997a, 1997b) hebben onderzoek gedaan naar de verwerking van objectieve en subjectieve causale relaties verbonden door het Engelse connectief *because* ‘omdat/want’. Dit connectief is ondergespecificeerd voor de mate van subjectiviteit en kan dus gebruikt worden om zowel objectieve als subjectieve relaties uit te drukken. De resultaten lieten zien dat subjectieve relaties met een ondergespecificeerd connectief langzamer worden gelezen dan objectieve relaties met hetzelfde connectief. Canestrelli (2013) en Canestrelli et al. (2013) hebben onderzoek gedaan naar de verwerking van gespecificeerde objectieve en subjectieve causale connectieven in het Nederlands, respectievelijk *omdat* en *want*. Hieruit bleek dat relaties met een subjectief connectief langzamer worden gelezen dan relaties met een objectief connectief. Deze bevindingen ondersteunen de subjectieve complexiteitshypothese (Noordman & De Blijzer 2000; Sanders 2005). Deze theorie voorspelt dat het verwerken van subjectieve informatie meer cognitieve inspanning vereist – en dus meer verwerkingstijd kost – dan het verwerken van objectieve informatie.

In hoofdstuk 5 is onderzocht of dit verschil in verwerking van objectieven en subjectieve relaties ook voorkomt bij Chinese teksten. Door de aanwezigheid van zowel gespecificeerde als ongespecificeerde connectieven in de Chinese taal, maakt deze taal het mogelijk om effecten van beide typen connectieven in één studie te vergelijken. Via een oogbewegingsexperiment is dan ook antwoord gezocht op de volgende onderzoeksvraag.

Vraag 6 Hoe worden verschillen in mate van subjectiviteit tussen Chinese causale connectieven weerspiegeld in online verwerking?

De causale connectieven *kejian*, *suoyi* en *yin'er* (alle te vertalen als ‘dus’) zijn onderzocht, omdat ze respectievelijk subjectieve, ondergespecificeerde en objectieve causale connectieven zijn. Er zijn vier condities gecreëerd: epistemische relaties met het subjectieve connectief *kejian*; epistemische relaties met het ondergespecificeerde connectief *suoyi*; contentrelaties met het objectieve connectief *yin'er*; en contentrelaties met *suoyi*. Zin (11) tot en met (14) zijn voorbeelden van deze condities.

- (11) *Meng Na na tiao kuzi xianzai xian de hen fei, **kejian** ta bi yiqian shou le bushao. Ta shi liang ge haizi de muqin*
 ‘Die (oude) broek ziet er nu erg los uit bij Meng Na, **kejian** ze is nu veel dunner dan eerst. Ze heeft twee kinderen.’

- (12) *Meng Na na tiao kuzi xianzai xian de hen fei, suoyi ta bi yiqian shou le bushao. Ta shi liang ge haizi de muqin.*
 ‘Die (oude) broek ziet er nu erg los uit bij Meg Na, **suoyi** ze is nu veel dunner dan eerst. Ze heeft twee kinderen.’
- (13) *Meng Na yi nian lai baoshou weibing de zhemo, yin’er ta bi yiqian shou le bushao. Ta shi liang ge haizi de muqin.*
 ‘Meng Na heeft al een jaar last van maagproblemen, **yin’er** ze is nu veel dunner dan eerst. Ze heeft twee kinderen.’
- (14) *Meng Na yi nian lai baoshou weibing de zhemo, suoyi ta bi yiqian shou le bushao. Ta shi liang ge haizi de muqin.*
 ‘Meng Na heeft al een jaar last van maagproblemen, **suoyi** ze is nu veel dunner dan eerst. Ze heeft twee kinderen.’

Tabel 3 illustreert hoe de zinnen opgedeeld zijn in regio’s voor de analyse. Zo konden we precies nagaan of en waar de effecten van het type verbindingswoord optreden.

Tabel 3. Het onderscheid tussen de drie regio’s.

	Connectief-regio	Onderwerp-regio	Predikaat-regio
Chinese zin:	<i>suoyi/yin’er/kejian</i>	<i>ta bi yiqian</i>	<i>shou le bushao.</i>
Vertaling:	dus	ze (nu) dan eerst	(is) dunner veel.

De resultaten van dit experiment laten een verwerkingsverschil zien tussen de relaties met ondergespecificeerde connectieven: in epistemische relaties met het ondergespecificeerde connectief *suoyi* (zoals (12)) worden het connectief, de onderwerp-regio en de predikaat-regio langzamer gelezen dan dezelfde gebieden in contentrelaties met het connectief *suoyi* (zoals (14)). Deze uitkomst repliceert eerder onderzoek met het ondergespecificeerde Engelse *because* (Traxler et al. 1997a, 1997b) en bevestigt de subjectieve complexiteitshypothese: een subjectieve relatie is complexer dan een objectieve relatie, waardoor de eerste meer verwerkingstijd kost.

De resultaten voor de epistemische relaties laten ook zien dat het subjectieve connectief *kejian* langere leestijden vergt bij de connectiefregio dan het ondergespecificeerde connectief *suoyi*, terwijl *kejian* bij de predikaat-regio juist kortere leestijden oplevert dan *suoyi*. Dit wijst erop dat de hoge mate van subjectiviteit die *kejian* kenmerkt het verwerken van epistemische relatie in eerste instantie vertraagt, maar vervolgens vergemakkelijkt. Het subjectieve connectief activeert direct bepaalde cognitieve processen, al tijdens het lezen van het connectief. Al voordat lezers het predikaat gelezen hebben, weten zij dat er sprake is van een epistemische relatie. Bij *suoyi* krijgen lezers vanuit het connectief geen indicatie dat er sprake is van een

subjectieve relatie. De verwerking van het connectief zelf gaat daardoor relatief snel, maar als gaandeweg, tijdens het lezen van het predikaat, duidelijk wordt dat er sprake is van een subjectieve relatie, treedt de vertraging alsnog op. Zowel de vertraging op de connectiefregio (in het geval van *kejian*) als de vertraging op de predikaatregio (in het geval van *suoyi*) zijn te verklaren vanuit de subjectieve complexiteitshypothese.

Een onverwacht resultaat trad op bij de vergelijking van de twee gespecificeerde condities, (11) en (13): het objectieve causale connectief *yin'er* zorgde voor langere leestijden bij de connectief-regio dan het subjectieve causale connectief *kejian*. Een mogelijke verklaring hiervoor is dat, hoewel *yin'er* frequenter is dan *kejian* (respectievelijk 1,50 per 10.000 woorden en 0,63 per 10.000 woorden), *yin'er* niet frequent is in gesproken conversaties, terwijl *suoyi* and *kejian* dat wel zijn. Deze bevinding komt overeen met de intuïtieve aanname dat *yin'er* bij formeel taalgebruik hoort. Het materiaal in dit experiment werd echter gekenmerkt door een informele stijl. Het is mogelijk dat de formele stijl van *yin'er* niet past bij de informele stijl van het materiaal, waardoor het langer duurde om *yin'er* te verwerken. Vervolgonderzoek kan hier meer duidelijkheid over geven.

De resultaten toonden geen verschil tussen leestijden van de predikaat-regio van causale relaties met *yin'er* en objectieve relaties met *suoyi*, vergelijk (13) en (14). Dit impliceert dat het objectieve connectief *yin'er* de verwerking van contentrelaties niet meer faciliteert dan het ondergespecificeerde connectief *suoyi*. Een mogelijke verklaring hiervoor is dat mensen eerder een objectieve dan een subjectieve relatie verwachten (oftewel de meest eenvoudige relatie) wanneer de tekst geen indicatoren bevat dat de relatie subjectief is.

Er was echter wel een verschil tussen leestijden van de connectief-regio van objectieve relaties met *yin'er* en *suoyi*. Het lezen van het objectieve connectief *yin'er* duurde langer dan het lezen van het ondergespecificeerde *suoyi*. Dit kan verklaard worden door de hierboven beschreven discrepantie tussen de formele aard van *yin'er* en de informele aard van het materiaal.

7. Conclusie

Het onderzoek waarover in dit proefschrift gerapporteerd wordt, toont aan dat de mate van subjectiviteit bepalend is voor het gebruik van Chinese causale connectieven. Dit komt overeen met eerdere bevindingen over subjectiviteit in Europese talen, ondanks verschillen tussen de manier waarop het Chinees en Europese talen connectieven categoriseren. De huidige subjectiviteitsbenadering bevestigt niet alleen de bevindingen van voorgaande studies, maar leidt ook tot nieuwe inzichten over het gebruik van causale connectieven in het Chinees. Dit proefschrift laat daarmee zien hoe een theoriegedreven en corpusgebaseerde benadering van de categorisering

van causale connectieven interessante resultaten oplevert voor verdere crosslinguïstische vergelijkingen.

De organisatie van het lexicon van causale connectieven kan gezien worden als een venster op conceptuele categorisatie. De soms subtiele verschillen tussen objectief en subjectief corresponderen met verwerkingsverschillen aangetoond in psycholinguïstisch onderzoek, en laten zien hoe subjectiviteit de manier waarop we connectieven gebruiken en interpreteren beïnvloedt. Een betere integratie van theoretisch en corpuslinguïstisch werk met psycholinguïstische studies vergroot ons inzicht in het gebruik van connectieven en andere coherentiemarkeringen. Op een hoger niveau zal deze integratie leiden tot nieuwe inzichten in het verband tussen linguïstische structuur en conceptuele structuur.

ACKNOWLEDGEMENTS

There are definitely enormous causal relations between the finalization of this dissertation and my many sources of assistance, support, and guidance. In the terminology of this dissertation, these causal relations are objective and undeniable. I would like to dedicate this section to the many people who have jointly contributed to the appearance of this book.

The first *thank you* goes to my supervisors, Ted Sanders and Jacqueline Evers-Vermeul. I clearly remember the first class I had, when I was enrolled in the Mphil program in Linguistics six years ago. The class was taught by a young, cheerful, and humorous professor. That was Ted, who later became my supervisor (MA) and promotor (PhD). Over the past six years, Ted has guided me into the amazing world of ‘causality’ and ‘subjectivity’, and equipped me for doing research in this field. *Heel erg bedankt*, Ted. Without you, this book would not have even been started. I would also like to apologize to Ted, because I feel guilty about many things at times. For example, from the plenty of jokes told by Ted, I have never caught the points to laugh. I (stubbornly) believe that Ted is indeed humorous, but that I always took too seriously whatever he said (believe it or not). Nevertheless, the humorous sparkles in his eyes have a magical power, always cheering me up on the gloomiest days.

I am lucky to have Jacqueline as my daily supervisor. She is quick-witted, critical, and amazingly efficient. Gradually, I learned to keep up with her pace, and also learned a great variety of other things from her, including how to write good journal papers, how to bake good cakes, and even how to play cards in the Dutch way. Without her careful and strict supervision on a truly daily basis, this dissertation would no doubt have taken much longer to be completed.

I would like to express the depth of my gratitude to Pim Mak. To Pim, I owe all my knowledge on everything concerning eye-tracking. Pim, thank you for your detailed guidance and insightful ideas on how to design and conduct the eye-tracking experiment, on how to analyze and interpret the output data, and eventually on how to turn the experiment into Chapter 5. Thank you for sharing your profound and original understanding about the eyes of Chinese people, the most interesting one of which concerns the ‘causal’ relationship between contact lenses and the color of a Chinese girl’s eyes.

I owe special thanks to Professor TAO Hongyin, whom I met at *the 1st International Symposium on Chinese Language and Discourse* at UCLA. He supported my research throughout the project, providing me valuable suggestions and comments especially on Chapter 3 and the project as a whole.

This dissertation combines different research methods – literature review, corpus-based quantitative study, and eye-tracking experimental study. Each type of study is conducted with the assistance and help from many people. I would like to thank ZHANG Jing (Shanghai Normal University) for helping me obtain access to the Chinese journal papers for the literature review. I wish to thank CHEN Gang (Anhui University), CHEN Weiyong (Zhejiang University), NI Yueru (Suzhou University), and SUN Yuning (UiL OTS) for assisting me with annotating the vast amount of Chinese fragments taken from linguistic corpora. I would also like to thank Huub van den Bergh for his statistical tutorial on General Log-linear Analysis. I thank the staff at the eye-tracking lab of UiL OTS, Iris Mulders and Theo Veenker, for their help with the programming and trouble-shooting. I thank CHEN Gang, CHEN Weiyong, GU Yan (Tilburg University), HAN Mengru (UiL OTS), and ZHEN Xueyan (Hebei University) for helping me to distribute questionnaires to native speakers of Chinese. I would like to thank all my Chinese colleagues at UiL OTS – CHEN Ao, DONG Xiaoli, HAN Mengru, HOU Luying, LI Xin, LIU Liquan, LIU Zenghui, WEI Yipu, YANG Anqi, and ZHANG Jingwei – not only for sharing their linguistic knowledge and intuitions on Mandarin Chinese but also for being available whenever I need them. I wish to thank the people who attended the DRP-meetings for the discussions and feedback that contributed to my research. Special thanks to Ninke Stukker, Pim Mak, Rosie van Veen, Anneloes Canestrelli, Ingrid Persoon. They are members of Ted's VICI project 'Causality and subjectivity as cognitive principles of discourse representation: Converging evidence from language use', whose achievements have laid a good foundation for the present dissertation.

I wish to thank our research coordinator, Maaike Schoorlemmer, for her valuable advice on research and practical help of various sorts. Many thanks to the secretary of the international office HUM, Julie Leijtens; the secretary of LOT, Martine Paulissen; the prior and present secretaries of UiL OTS, Mariette Bonnënkamp, Martien Camphuis, Yvonne van Adrichem; and the secretary of TLC, Chantal Verhoef. I owe to these nice and helpful people my library card, office key, residence permit, travelling documents, registrations for research-related/unrelated activities, arrangements for meetings, and a lot more.

Assaf Toledo, thank you very much for helping me take the beautiful picture for the book cover. It is among the few elements of the book that passed Jacqueline's examination in the first round. LIU Zenghui, thank you very much for your brilliant idea of inviting Assaf to take that picture. I would also like to thank HOU Luying, HU Wenyong, and ZHANG Jingwei for offering me many alternative pictures to choose from, which I will treasure forever. I wish to thank Rachel Gargiulo for proofreading my dissertation. I am very grateful to Merel Scholman, who provided a nice

summary of this dissertation in Dutch. I would like to give my special thanks to Marta Castella and WEI Yipu for being my paranymphs, standing behind me at an extraordinary moment of great tension and great happiness.

I am very grateful to my prior and present roommates in Room 1.52 (Trans 10): Anna Chernilovskaya, Anne-France Pinget, Alexia Guerra Rivera, Marko Hladnik, Mike Olson, and HOU Luying. Thank you all for your support and company; and for the strawberries, cherries, cookies, and hand-made cakes; and for translating the variety of Dutch mail and forms (although you are not Dutch, either). I will miss you all, and our sunny and cozy office that has been captured in the cover picture and also permanently in my mind.

I wish to thank all other colleagues at UiL OTS. Together, you make *Trans 10* my second home. I would also like to thank all my friends in the Netherlands and all acquaintances I met during winter/spring/summer-schools and international conferences. Every one of you have become the most precious parts of my memory: the Linguistics Spring Fest Party, the Workshop on Chinese Linguistics, cycling to Kröller Müller Museum, walking in the Amsterdamse Bos, watching fireworks on Scheveningen beach, barbecue at UCU, taking super long waterslides at Duinrell, the Cocktail and Candlelight party, playing cards at Keukenhof, picnic at Giethoorn, beach combing at Vlissingen, the Spring and Onion Party, swimming in Zwembad Dijnseburg, walking along the Seine in Paris, feeding the doves in Venice, the city tour in Los Angeles, watching dog-racing in London, excursion at Cambridge, drinking beers during the Munich Oktoberfest...

Most importantly, I am indebted to Huygens Scholarship and Nuffic. Without the Huygens Scholarship granted by Nuffic, the above experiences, together with this dissertation, would have not taken place.

Finally, I would like to express my thanks to my parents, my husband, and all other family members, in Chinese.

爸爸妈妈:

感谢你们为我营造的轻松、和谐的家，让我无忧无虑地长大。在你们温暖的羽翼下，我看到的世界充满了美好，从而积蓄了足以受益一生的正能量，支撑我一次次从挫折中爬起来，拍拍土，继续前行。感谢你们给予我无私的爱。做你们的女儿，我感到很幸运。

老胡:

那些日子，我在你的办公室里写作业到凌晨两点。你打完球，就在我对面的沙发上打瞌睡直到我说“回家”。我觉得很感动。

那一天，我把标有优等毕业的硕士证书摆在你面前，你笑着说炫耀。我觉得很幸福。

就在去年，你回国到艰苦的地方工作，而我留在荷兰作博士毕业论文。我一直很牵挂。。。
这一切，我都：很珍惜。

我远在天堂的奶奶：

感谢您十九年的日夜陪伴，加上十一年聚少离多的牵挂。感谢您一路上给我的点点滴滴，每一顿家常饭，我最爱的豆角焖面和我最恨的猪油烙饼，每一件过冬的棉衣，每一声称赞，每一句唠叨。。。
这一切，我已然无法回报。只求，将爱传递。

最后感谢我所有的挚友亲朋：

你们给予我的亲情、友情汇成了一股最温暖的力量。

The following is a poem written by my mother, who wished to be a writer since young. My father, using his philosophical eyes, had critically reviewed the previous versions of the poem, and finally gave an approving nod. My parents want me to put the poem at the end of the ACKNOWLEDGEMENTS, to look back to my hard and happy times in the Netherlands and to express our gratitude to my supervisors and all my teachers at Utrecht University.

I would like to pass on their plain and simple message faithfully in Chinese.

下面是从小立志当作家的妈妈帮我写的一首诗。爸爸用他哲学的眼光进行了严格的审查。他们嘱咐我把这首诗附在致谢的后面，来表达我对荷兰过往的缅怀和对我的导师和乌特列支大学所有老师的感谢。他们朴素的心意，我如实传达。

鲜花之国次故乡，
勤奋攻读乌寒窗，
无涯学海苦中乐，
修成正果喜泪淌。

异国六载苦短长，
乌特列支情难忘，
北海之水深千米，
不及恩师授渔香。

CURRICULUM VITAE

Fang Li was born on the 4th of August 1977 in Hebei Province, China. She studied English language and literature at Hebei University (China), where she received a BA in 2000 and an MA in 2003. She became a teacher in the Faculty of Foreign Languages Studies at Hebei University in 2003, and was promoted to lecturer in 2007. She started an international Mphil program in linguistics in 2008 at Utrecht University, where she obtained a research MA (cum laude, specializing in language use and psycholinguistics) in 2009. In 2010, she started her PhD program on subjectivity and causal connectives in Mandarin Chinese, enabled through Huygens Scholarship awarded by Nuffic. In 2013, she published a corpus-based study on subjectivity and result marking in Mandarin in the international journal of *Chinese Language and Discourse*, in joint authorship. This dissertation is the ultimate result of her PhD research.