CHAPTER 3. DATA AND METHODOLOGY

3.1 Data Description

The data of this study consist of 100 hours of video and one-million-words (1,129,437) of transcripts of face-to-face spontaneous conversations in Mandarin Chinese drawn from 300 episodes of a famous talk show in China – *Qiang Qiang San Ren Xing* 铿锵三人行 ‘Three Companions.’ The talk show, which is on the Phoenix Chinese Channel, is produced by the Phoenix Television and is broadcasted in China. This popular talk show has been airing on every weekday for 18 years since 1998. Each episode features a conversation that lasts between 20 to 25 minutes. The conversations in this talk show are unscripted and basically unedited (detailed descriptions will be given later in this chapter).

![Figure 3-1: Snapshot of the talk show Three Companions (episode of January 10, 2014)](image)

The talk show features a three-person conversation in a casual setting (Figure 3-1). The long-time host Wentao Dow (窦文涛), born in northern China, is a native speaker of Mandarin Chinese. Almost all the guests are Chinese native speakers. Only a few guests are non-native...
speakers who possess native-like Chinese proficiency. These non-native speakers were on the show for only a few episodes, and these episodes were excluded in my data. “The show invites a variety of guests from mainland China, Hong Kong, and Taiwan who are media professionals, academics, cultural critics, novelists, filmmakers, economists, and reporters.” The most frequently invited guests are Zidong Xu (许子东) and Wendao Liang (梁文道). Since 2006, the talk show has been recorded both in Hong Kong and in Beijing, with about half of time being recorded in Beijing.

3.2 Justification of Data

3.2.1 Suitable for the methodological orientation

The talk show has been broadcasted on television since April 1, 1998. The format of the talk show remains unchanged during these past 18 years: approximately 22 minutes per episode, one episode per weekday, and five days a week. Starting from 2015, approximately three more minutes were added to each episode. For 18 years, the talk show has maintained its basic structure of a three-party spontaneous conversation that is meant to resemble the casual chatting style among friends.

What makes this talk show particularly valuable to the current study is that the participants often comment on the same social event with different points of view. In other words, “participants [are] juggling different versions of past events and different images of certain people’s involvement” (Duranti 1990: 662). This feature has made the talk show suitable for my research purpose: how the same event is presented differently with the use of different grammatical constructions.

---

8 An example of these exceptions is Run Hu (胡润, English name Rupert Hoogewerf), whose is Chinese is native-like. He has studied Chinese language since 1990 and has been living in China for about twenty years.

On a macro level, the large-scale data allowed me to conduct corpus analysis to obtain quantitative patterns and tendencies. On a micro level, the conversational data allowed me to use discourse analysis methodologies to analyze the interactional context of the grammatical constructions in question. The videotaped data has made it possible for analysis of not only the verbal language but also the multimodal semiotic resources (such as prosody, gaze, and gesture) used by the participants (See Goodwin 1979, 2009, 2013 for the importance of using videotaped data), which is useful for investigating how the grammatical constructions in question are actually used.

3.2.2 Transitive especially causative events are abundant

The talk show covers a wide range of topics – from news to entertainment, from sports to politics, from history to vacations, from popular topics in the contemporary Chinese society to less popular academic subjects, and from everyday personal matters to philosophical and cultural subjects.

An important characteristic of this talk show makes it particularly useful for the study of the grammatical constructions related to transitive (especially causative) events: the participants often talk about the cause and effect of some recent events in China or in the world. For example, the topic for the episode on March 19, 2014 is regarding who/what has caused the disappearance of the Malaysia Airlines Flight MH370. In this single 21-minute episode, as many as 40 cases of analytic causatives covering 10 major types and 18 sub-types of Chinese grammatical constructions were identified.

3.2.3 Constant setting in 18 years constitutes a consistent genre

In recent years, more and more linguists have come to the understanding that research on grammar needs to be based on genre-specific data (e.g., Chafe 1982; Miller & Weinert 1998; Tao
The conversational settings, style, and content of this talk show have remained consistent during the past 18 years, making the features of the data rather consistent.

In order to create a casual chatting environment, this talk show maintains a simple setting: one desk, three chairs, and three cups. There is no background music or any other audio effects. There are no incoming calls or any audience on the recording site. “Nothing but the table and chairs, props, and the three speakers are actually real.” According to the director of this talk show Jinhui Liang, the desk is specially designed to have a trapezoidal shape, so that the faces of the three interlocutors can be better captured by the camera.

**3.2.4 Intended nature of this talk show: chatting with friends**

This talk show is a successful show that has attracted billions of viewers. The great success of this talk show can in part be attributed to its intended nature of conversation – chatting with friends. There are two common participant layouts: 1) the host and two old friends (i.e., two regular guests); 2) the host, an old friend (i.e., a regular guest), and a relatively new friend (i.e., a less regular guest). Either one, there are old friends in the conversation.

This feature reflects a deliberate thought of the host Wentao Dou. According to Dou,

---

10 胡尧熙 2008《锵锵三人行》的节目形式与技巧 In 《锵锵三人行》之窦文涛：一个话痨的十年, 《新周刊》第 269 期 2008 年 2 月 15 日上市 http://blog.sina.com.cn/s/blog_4900756601008ng2.html

11 Commented in the blog Three Torches. https://threetorches.wordpress.com/tag/%E9%94%B5%E9%94%B5%E4%B8%82%E4%B8%89%E4%BA%BA%E8%A1%8C/ Accessed on March 22, 2014.


13 “文涛说那天的电视新闻正好播放了一则关于飞机失事的新闻, 自己就假想如果把这个问题放到节目里自己要怎么说呢, 一下子他恍然大悟, 其实生活里每天都有这样的事情在发生, 我们每天都会和同事们和朋友们聊天, 大家聊得都是热点的话题, 在聊的时候是那么享受, 这不就是聊天嘛, 大家为的就是享受聊天的这个过程。仔细想来, 其实我们和其他朋友、同事在平时的聊天中并不是为了真的得到什么真理或者去解决什么问题, 就是为了乐趣而聊, 享受谈话。想明白这一点后, 真的就是一通百通了, 节目的思路就这样出来了。其实在节目中就找合适聊天的朋友来聊天就好了, 也不一定非得就请什么专家学者, 我们在平时约朋友畅聊时, 也不事先就定好话题, 定好谁开头, 谁结尾, 谁去总结, 就那么自然地发生了。虽然在节目中做不到像生活中那么仿真的聊天, 但是有了这个感觉, 其他事情就水到渠成的了。” 窦文涛：锵锵 15 年 我的那些事 http://phtv.ifeng.com/star/douwentao/detail_2013_05/29/25840838_0.shtml
people talk about things happening in the world and in life with their friends every day. People enjoy exchanging thoughts and feelings through chatting. When people are chatting with friends, they do not set a fixed topic for a conversation in advance, nor do they prescribe who should do the opening and who should do the ending or summary. The conversation simply flows naturally. The talk show uses this format of talking to resemble everyday conversation with friends. There was even a time when the host Dou talked a lot about his own personal matters in the show. For instance, who he met earlier today; what kinds of flowers he grew at home; where he went last night, etc. The entire episode could be all about these kinds of personal subjects.15

A most regular guest Zidong Xu commented that “while participants in other talk shows may be talking to the audience, in our talk show we simply chat for ourselves.” The director at Phoenix Television who oversees this talk show also said that “there isn’t much preparation for this talk show because the goal is to create an environment that allows casual chatting among old friends – What needs to be prepared before you talk with old friends?”17

This feature has been welcomed by both viewers and reviewers. The host Wentao Dou has a well-received reputation for the person who “makes people on television speak like real human beings.”18 The Wikipedia entry for this talk show describes that “the atmosphere is very casual

---


18 他让中国电视开始说“人话”。——2004 年“新锐 200”窦文涛评语. 《锵锵三人行》之窦文涛: 一个话痨的十年,
and enjoyable, like old friends having a fireside chat at night.” A viewer of this talk show observes that “there is very little scheming and manipulating in this show. Also, since it is five days a week, there is not much time for planning ahead. Therefore, in the host Dou’s own words, this talk show ‘is like chatting with friends in everyday life.’” A reviewer comments: “It (this talk show) wants to change the way how the TV media talk, i.e., not performing for the audience by following what is in the script – instead, chatting naturally in a way that resembles authentic conversation in everyday life. Therefore, this talk show often wanders off the topic, changing topics constantly in the conversation. But this is not a problem at all. In fact, this is what it intends to be.”

3.2.5 Unscripted and unedited unless violates a ban

“Three Companions is almost like a live show.” “It is this kind of improvisation and the quasi-live way that has made Three Companions unique and successful.”

To find out how much scripting and editing was involved in this talk show, I emailed one of the two most regular guests on the talk show in the past 18 years – Dr. Zidong Xu. 

---

19 “现场气氛轻松愉快，如老友围炉夜话。”

http://zh.wikipedia.org/wiki/%E9%94%B5%E9%94%B5%E4%B8%89%E4%B8%89%E5%9B%9E/zh

20 “在处理谈话氛围时，《锵锵三人行》很少进行过多策划与包装，而对于每星期五次的节目播出量，对每一期谈话内容事前都进行周详的策划也不太可能。因此，用主持人窦文涛的话来说，这个节目“就是我们平常朋友聊天这一类，我们就是用本色，把生活中的聊天尽量本色地搬上电视”。“聊天往往是强调一个聊的过程，……大家都是情中人，畅所欲言，享受的是这个”……”

http://wenku.baidu.com/view/d63a74c48bd63186bdebcc04.html

21 “它想要改变一种电视说话的方式，不要按照台本来演给观众看，而是很真实、很自然的聊天。所以这个节目经常跑题跑不停，也无所谓，要的就是这个状态。” 胡斐 2008 《锵锵三人行》十年发展史.

http://phtv.ifeng.com/star/douwentao/detail_2013_05/29/25841168_0.shtml

22 “锵锵是一个准直播的聊天节目”.

http://phtv.ifeng.com/star/douwentao/detail_2013_05/29/25841168_0.shtml

23 “正是这种即兴性和准直播方式给予了《锵锵三人行》独特的生命价值。” 窦文涛：做任何事都要像做一件艺术品

professor of modern and contemporary Chinese literature at Hong Kong Lingnan University. Dr. Xu told me that: “The conversation is pretty much spontaneous and almost no cutting or editing is done unless in some very rare situations where the opinions violate a ban. Guests do not prepare in advance either. The guests are usually told a general topic one day ahead or even a few hours before the show.” This response confirms what I have learned from other sources.

Dr. Xu has also given similar descriptions of the talk show on some other occasions. For instance, he notes that “while other shows all have planned what to say in the show, our show has no scripts. We don’t know what to say before we actually sit down and talk.” The other most frequent guest, Wendao Liang [梁文道], also said that: “We do not have any scripts. All we have is a broad sense. This is actually an advantage of our show.” The staff costs of this famous talk show are rather low. According to the host Dou, there are only two and “a half” full-time staff members for this show: the host, an assistant that sends invitations to the guests in Beijing, and a third person who also has to work for another TV show. Other people are interns. This is not surprising because this talk show does not require the kind of heavy editing typically needed for television shows.

Here is another piece of evidence showing that the talk show is lightly or not edited. In the episode on March 17, 2015, the host apologized for having shown a movie trail in a prior episode,

---

24 Original response (Mar 22, 2014): “基本上照录照播,一般不剪辑，除非有别言论犯禁。嘉宾事先也不做什么准备。提早一天或数小时知道大概话题。”


which made that episode look like an embedded advertisement for the movie. The host said that the reason they needed to play that movie trailer was that they had recorded that prior episode ahead of time, but the talk show had a new change in format – adding three more minutes.

Therefore, the episode they had already recorded was three minutes short for this new format.

This was why the host had to show the movie trailer to fill up that three additional minutes.  

3.2.6 Linguistic characteristics resemble that of everyday conversation

1) Constantly running off the topic

According to the host Dou, the talk show usually starts with a news topic, and then the topic can change freely into any other topics. The conversation is unrestrained and moves naturally. The talk show even earned such a reputation – “Three Companions; Constantly Running off the Topic” – soon after it was launched. A famous magazine in China even considers that “the high viewing rates of this talk show can be attributed to the host’s ‘idle chatting style,’” a feature that is welcomed by the viewers.

28窦文涛：访徐静蕾像宣传电影 我向观众道歉 http://phtv.ifeng.com/a/20150318/41015402_0.shtml
窦文涛：《锵锵三人行》, 今天我们可以谈谈主旋律了。主旋律的电视剧《平凡的世界》, 当然, 我们可不是宣传, 又做定影宣传, 这个我得跟大家有件事说在前头, 就是最近我们又犯错误了。我虽然隔了春节一个这么久远, 但是我还要欠大家一个道歉, 这个道歉也对了解我们节目, 对我也很启发。为什么呢? 我就是有两点感触, 一点感触就是我们这个节目虽然说小, 但是甚至我自己都没有意识到我们有什么特色, 可是在观众心目中, 特别在咱们老观众心目中, 我们是有我们的气质的, 我们是有我们的性格的。我们有些事儿是不适合干的。

周轶君：是有节操的。
窦文涛：对。为什么呢? 因为好家伙, 那个微博上, 我先看到很多人就骂, 骂什么呢, 骂徐静蕾, 不是骂徐静蕾, 徐静蕾并没有错, 徐静蕾前一阵拍的那个电影在我们这儿不是播了几集吗? 很多人就骂, 说你们已经沦为给徐静蕾的电影做这样的宣传。所以, 我就说第一就是要道歉, 第二告诉大家这个事儿是怎么来的, 其实没有, 我们今年《锵锵三人行》 这个节目承蒙领导重视, 一个是时间提前到十一点播出, 再一个整整给加长了三分钟。但是加长了三分钟。去年年底 12 月录的徐静蕾的三集访问但是还没播, 电影人家情人节才上映, 人家不让播。人家说咱们答应你们采访的条件就是, 你们必须在那个时候才播, OK, 好, 跨了今年这个节目时间不就短了三分钟, 于是这个编导就跟我说, 怎么办呢, 怎么办呢, 你录了也不能不播啊。好, 他说咱们我把这三集并一集, 反正是原来短, 现在需要更长的时间, 怎么办呢, 他说我这三集, 可是我可以加一些电影片段进去。


30“锵锵的收视率源于窦文涛的 ‘掰扯’ .” 胡赳赳 柯勇 2008 谁在看《锵锵三人行》《新周刊》2008 年 02 月 14 日 http://blog.sina.com.cn/s/blog_4900756601008ng0.html
2) Speech errors remain uncut

In this talk show, sometimes either the host or the guests may misspeak a name (see the episode on February 24, 2014, #20140224). The host may even say a guest’s name wrong for multiple times (see example #20140121 and #20140404). This kind of speech errors is kept unedited in the show.

(#20140224)

160 Host: 哦杨沫，

161 不是杨绛。

162 Li: 唉哟口误了，

163 杨沫，

164 对不起，

---

By 寂静的手指 2010-05-05: 我看《锵锵三人行》这档节目, 这节目还有规定的主题嘛, 看了这么久还真不知道。

Answers:
河阳新人: 跑题是特色。
新手中的老手: 跑题跑不停, 是他们的一大特色。
脸上写着帅: 这节目还有规定的主题嘛, 看了这么久还真不知道。
逍遥灵圣: 这节目就这个特色。
cp1154: 这就是为什么能办 16 年的原因, 人家办了 16 年你才刚关注。不喜欢就换台。
二丫讲楚: 锵锵三人行，跑题跑不停。
自由画匠: 跑题是聊天的本真，不跑题那是作报告，是官样文章，是 CCTV。
max2101983: 一本正经的聊，谁看啊。
太虚观看门人: 不跑题不锵锵。
中速铃声天籁: 你就当是朋友聊天 真要每天聊点东西出来耶不现实 他们也只是懂得稍微多的普通人。
武当山刘海鹏: 锵锵十五周年那里面就特别注明，这是一档永远跑题的谈话节目。
散花: 没法儿，事先没有稿子。
蓬蒿人: 跑题也挺有意思的~突然发现现在自己讲话也有点这样，经常讲着讲着就跑了。
艺文爱桥边: 本来就是跑题节目啊。
我得道歉，
I have to apologize (to you).

杨沫，
Yang, Mo.

对不起。
Sorry.

Host: 没事，
It doesn’t matter.

(#20140121)

Host: 你要不要跟幼婷说一下，
Do you want to tell Youting.

按说道德上讲我要不跟幼婷
Morally speaking, if I don’t tell

说呢，
Youting.

不是幼婷，
Not Youting.

玮婕，
Weijie.

玮婕。
Weijie.

Weijie: 你看！
See!

我都露了三公分的腿还叫我
I have already (put on a short skit

别的名字，
that) shows three inches of my legs

和 you still call me the wrong

name.

我要哭了。
I want to cry.

Host: 你头一次露，
This is the first time you show your

腿（on the show）。

头一次露。
The first time (you) show.
Ma: 你回去重看锵锵，
If you watch the previous (episodes of) *Three Companions*,

人家是露五公分的。
(you can see that on this show)

Youting showed five inches of her legs.

Weijie: 可能第一次总是比较紧张一点嘛，
Maybe it is because this is the first time (I show my legs), I am relatively nervous.

你给我个机会吧。
Please give me a chance.

Host: 开关还没转过来，
I haven’t gotten used to (your showing your legs).

我就说你要不要跟玮婕说呢，
I was saying that do you want to tell Weijie.

(#20140404)

Host: 我跟你说幼婷。
I tell you, Youting.

Weijie: 你看！
See!

Host: @@@
@@@

Weijie

Weijie: 然后一直叫错，
Always call me the wrong name.

第八百两千万次这样。
(This is) the 800th or the 20,000,000th time (you say my name wrong) like this.

Host: 完了完了完了。
I’m done. I’m done. I’m done.
Weijie: 完蛋了记恨了。 You’re done. I’ve held grudges.

Host: 诶，算了，算了。 Never mind. Never mind.

Ma: 玮婕 Weijie,

她是玮婕。 She is Weijie.

Weijie: 我下次自备一个名牌。 Next time I will prepare a name card myself.

3.2.7 Excluded data

These two types of data, although rare, were excluded from the analysis of the current study. 1) Lines that I could not hear clearly while transcribing the conversation. 2) Written language. Reading off from a text rarely happens in this talk show. But when it did happen\(^\text{32}\), the written lines were excluded from the analysis of the data.

3.3 Studying Grammatical Construction through Adjacent Alternation

My method is to study grammatical constructions through adjacent alternation, namely, alternative grammatical constructions used to describe the same event. I focus on cases of high adjacency – utterances produced within a single spontaneous natural conversation that lasts no longer than 30 minutes. Below is an example.

(1) UP => ba alternation

1 Dou: [Unmarked 垃圾都随身带走了。(Even) the trash (was) all taken away.]

\(^\text{32}\) For example, (#20140314)

窦文涛: 冯小刚说今天李总理的报告当中，他说，
他说，今年要再取消和下放行政审批事项两百项以上，深化投资审批制度改革，取消或者简化前置性审批，
就是之前的这个审批，((Written texts are excluded from the data analysis))
2 太有趣了。 It’s quite interesting.

3 Zhou: 我觉得这一点做得还挺 I think (they) did a good job
好， on this.

4 [Ba] 你最后把垃圾都带走。 In the end, you (even) took
all the trash away.

(#20150514\(^{33}\))

In example (1), the speakers are talking about a tourist group of 6,000 Chinese people who went to France in May 2015 and took all the trash away as they left France. The first speaker, Dou, uses an unmarked passive construction to describe the event of taking all the trash away (line 1). Immediately following Dou’s comment, the second speaker, Zhou, uses a *ba*-construction to describe the same event (line 4). The two grammatical constructions appear in close proximity in the same conversation and are describing the same event in real life.

An adjacent alternation, or sometimes referred to as “alternation” in this study for short, is counted when there are at least two alternative grammatical constructions commenting on the same event. Example (1) has one alternation; it is a 2-form alternation that contains two alternative forms: the unmarked passive construction and the *ba*-construction. Because these two constructions both appear once in this alternation, there are counted as two alternative uses. An alternative use is an occurrence of a grammatical construction in an adjacent alternation. If the speaker Zhou in example (1) had used the *ba*-construction twice, there would be three alternative uses in this alternation, but the number of alternative forms would still be two: the unmarked passive construction and the *ba*-construction.

An alternation can be notated with either a path-specified notation or a path-unspecified
notation. Whereas the symbol of the path-specified notation “=>” indicates the temporal order of the constructions used in an alternation, the symbol of the two-way notation “<=>” does not indicate the temporal order. Using the path-specified notation, example (1) can be noted as an unmarked passive => ba alternation, which means that the speaker(s) first use(s) an unmarked passive construction to describe an event and then switch(es) to using a ba-construction to describe the same event. Using the path-unspecified notation, example (1) can be noted as an unmarked passive <=> ba alternation or a ba <=> unmarked passive alternation, which notes that speaker(s) use(s) an unmarked passive construction and a ba-construction to describe the same event without specifying the temporal order of the grammatical constructions used.

3.4 Three Datasets and Research Design

My first dataset consists of a total of 1,000 minutes of conversation from 50 episodes of the talk show, which aired from January 1 to March 27, 2014. I further transcribed the conversations using the incomplete transcripts provided on the website of Phoenix Television. Based on the transcripts, I manually and exclusively coded all the actual grammatical structures that are used by the speakers to describe a transitive event that involves a causer, an affectee, a cause, and an effect. I ended up having 1,583 examples that involved 22 major types of grammatical constructions (such as the bei-passive construction) and 44 subtypes of grammatical constructions (such as long bei-passives and short bei-passives). The four most frequent grammatical constructions turned out to be: the ba-construction, the unmarked passive construction, the rang-construction, and the bei-passive construction. The total occurrences of these four grammatical constructions account for 70.1% of all the 1,583 cases. These four grammatical constructions thus became the main subjects of my investigation.

I then watched all the 300 videos along with their transcripts and identified 191
alternations involving 470 alternative uses – occurrences of these four grammatical constructions in the same conversation commenting on the same event. These 470 alternative uses, which were identified from the entire 100-hour database, constitute my second dataset, which is the main dataset.

My third dataset consists of 5,679 single uses of these four grammatical constructions, which were identified from the 100-hour database. They include 1) 5,431 single forms involving all the uses of the ba-construction (2,526), the rang-construction (1,507), and the bei-passive construction (1,397) exclusively found in the 100-hour database. This task was conducted using the corpus tool AntConc 3.4.434 and was manually checked. The raw data were 5,593, but 163 cases in which the target words were not used as special particles for syntactic constructions (such as 被子 beizi, ‘blanket’) were excluded. 2) 249 unmarked passive manually and randomly identified from the 100-hour database. This third dataset was used to quantitatively capture the alternation tendencies of these four major grammatical constructions and some semantic features of them.

For data analysis, I combine discourse analysis with corpus linguistics. One of the common features of both methodologies is that they do not focus on individual words or sentences isolated from context but instead focus on real communicative contexts and language use beyond the sentence level in natural discourse. Discourse analysis is a qualitative analytical tool that can be used to study why certain grammatical structures are used in certain contexts. Corpus linguistics is a quantitative research tool for analyzing large collections of language data.

The combination of these methodologies is a reflection of a methodological trend in linguistic research. Tao (2003c) is among the earliest advocators for utilizing the advantages of

34 Software developed by Laurence Anthony, Waseda University, Japan. http://www.laurenceanthony.net/
both corpus approach and conversation analysis (CA) in linguistics study. Tao (2003c) notes that “it has been amply demonstrated that in investigating matters of interaction and grammar, detailed analyses of single episodes of interaction as exemplified by many classic studies in the CA tradition have distinctive advantages” and that “computer-assisted analysis of large amounts of data can complement CA to some extent.” This methodology of combining computer-assisted analysis of large amounts of data and fine-grained conversation and discourse analysis has been proven fruitful in some recent publications (e.g., Tao 2003c; Sohn & Kim 2008; Thompson & Tao 2010; Sohn 2010; Couper-Kuhlen 2014).

In this study, the research question of how native speakers make the choice among a range of options that are grammatically correct and semantically similar is addressed both quantitatively and qualitatively. On a macro scale, a corpus-based analysis was applied to two datasets: 1) the first dataset that contains the transcripts of the 1,000-minute conversations. All the occurrences of the 22 major types and 44 sub-types of Chinese grammatical constructions were coded. Statistical analysis was conducted to retrieve distributions of the grammatical constructions that were actually occurring. 2) The second dataset that contains one-million-word transcripts of the 100-hour conversations. Statistical analysis was conducted to retrieve alternation rates of the four grammatical construction as well as some semantic features of the bei-passive construction and the rang-construction. On a micro scale, a question was investigated using discourse analysis: Within a single conversation, when speakers alternative different grammatical constructions to describe the same event, how and why do they switch among different forms?