Minimal Response Token *en* in Mandarin Conversation

by

Gaisha Oralova

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Arts

in

East Asian Interdisciplinary Studies

Department of East Asian Studies

University of Alberta

Abstract

This thesis focuses on one of the most frequently used response tokens in Mandarin - *en* “mm”. Through examining 6 hours of everyday spontaneous Mandarin conversation, this paper explores the interactional functions of the response token *en* in different sequential and situational environments, as well as its prosodic and visual features. The analysis shows that *en* has several uses in current conversational data. When placed in the middle of an extended turn, *en* can serve as a continuer. When it appears in the end of an extended turn, *en* functions as an acknowledgement token. *En* may be a confirmation token when it confirms a previous turn’s assertion or claim. It can also register the receipt of the listener's responses at sequence-closing third positions. In addition, this study describes the prosodic features of *en* and the body movements concurrent with the production of *en*. This thesis hopes to shed some light on the usage of the minimal response token *en* in Mandarin conversation, as well as on response tokens at large. Many learners of Mandarin find it difficult to distinguish different usages of Mandarin response token *en* in conversation. Therefore, this study on the minimal response token *en* also has implications to Chinese language teaching and learning.

Key words: minimal response token *en*, interactional functions, Conversation Analysis, Interactional Linguistics, Multimodal Analysis, Mandarin conversation

Acknowledgement

I would first like to thank the Department of East Asian Studies at the University of Alberta for giving me the opportunity to pursue my MA studies in their department.

I especially want to thank my thesis supervisor, Dr. Li Xiaoting, for the continuous support she has given me during my study in the department. Whenever I had a problem or a question about my research or writing, her office doors was always open. Also, I want to thank Professor Li for providing me with the data I needed and giving insightful comments without which I could not finish writing this thesis.

I would also like to thank Professor Tsuyoshi Ono, from East Asian Studies Department at University of Alberta, and Professor Juhani Järvikivi, from the Department of Linguistics at University of Alberta, for being the second reader and giving me valuable comments on this thesis.

Finally, I must express my gratitude to my parents, Takhir and Gulnara, my younger brother, Mansur, and to my spouse, Dmitriy, for providing me with unfailing support and encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

Table of Contents

Chapter 1 Introduction 1

1.1 The minimal response token *en* 1

1.2 Phonetic variations of the minimal response token *en* 5

1.3 Participation framework in conversation 5

1.4 Sequential structure of conversation 6

1.5 Research questions 8

1.6 The organization of the thesis 9

Chapter 2 Literature review 10

2.1. The definition of response tokens 10

2.2 Previous studies on response tokens in Mandarin 12

2.3 Studies on response tokens in other languages 16

2.4 Summary 19

Chapter 3 Methodology 20

3.1 Data 20

3.2 Methodology 22

3.2.1 Conversation Analysis 22

3.2.2 Interactional Linguistics 23

3.3 Summary 25

Chapter 4 Results: the interactional functions of the minimal response token *en* 26

4.1 Continuer 28

4.2 Acknowledgement 46
4.3 Confirmation 68
4.4 Register of receipt 79
4.5 Summary 89
Chapter 5 Conclusion 90
5.1 Findings 90
5.2 Significance and implications of the study 92
Bibliography 94
Appendix A: Transcript symbols 102
Appendix B: Glossing conventions 103

List of Tables

Table 1 Distribution of single and multiple *en* tokens in the data 26

Table 2 The four functions of *en* in the data 27

List of Figures

Figure 1 Pitch trace of the continuer *en* in line 16 of Fragment 4 37

Figure 2 Rui (left) gazes away during *juhua* (line 4)..... 44

Figure 3 Rui (left) gazes at Liang(right) during *jiu shuo* (line 4) 44

Figure 4 Pitch trace of the continuer token *en* in line 17 of Fragment 5 45

Figure 5 Body position of Liang (right) during Rui's (left) *haiyou* in line 20..... 56

Figure 6 Body position upon Liang's (right) production of acknowledgement *en* (line 21)..... 56

Figure 7 Body position after Liang's (right) production of acknowledgement token *en* (in line 21) 56

Figure 8 Pitch trace of the acknowledgement token *en* in Fragment 6 line 32..... 58

Figure 9 Lei (left) gazes at Wei (right) during *peiyang* at line 12 (Fragment 7) 66

Figure 10 Lei (left) gazes downwards during *en* at line 13 (fragment 7) 66

Figure 11 Pitch trace of the acknowledgement token *en* in line 25 of Fragment 7 67

Figure 12 Liu (left) gazes downwards during *zai* 73

Figure 13 Liu (left) gazes back at Hu (right) during *yici* 73

Figure 14 Liu (left) gazes at Hu (right) during repair 73

Figure 15 Liu (left) gazes at Hu (right) during *en* 73

Figure 16 Pitch trace of the confirmation token *en* at line 8 of Fragment 8 74

Figure 17 Pitch trace of the confirmation token *en* in line 4 of Fragment 9..... 78

Figure 18 Bai (middle) gazes downwards during the production of register of receipt *en* in line 8 in Fragment 10 82

Figure 19 Pitch trace of register of receipt token *en* at line 8 in Fragment 10..... 83

Figure 20 Pitch trace of register of receipt token *en* in Fragment 11 88

Chapter 1 Introduction

In this chapter, I will introduce Mandarin minimal response token *en*, its phonetic variations in the current data, and some relevant notions, such as participation framework and sequential structure in conversation, which will be used in data analysis. Also, this chapter will pose the research questions for this study.

1.1 The minimal response token *en*

Brief responses are frequently produced by conversational participants during spontaneous conversation, and sometimes these items occur more than a thousand times in an hour (Gardner 1998:205). These brief utterances, such as *yeah, mmhm* and *right* in English, support the progressivity of the conversation without taking the main speaker's floor (Schegloff 1982). When telling a story or simply talking to someone, we can notice the listener's reaction to our talk through these minimal responses. Imagine if the listener gave no reaction when you were speaking to him or her, you may have thought that he or she is not interested in what you were saying, disagrees with your statement, or was not even listening to you at all. The speaker may design the subsequent communication accordingly. Thus, the listener is not just a passive recipient, but an active participant who also does as much work during the interaction as the speaker. Response tokens provide information to speakers on how the information is received by the listener. They also provide information on how the utterer of a response token projects further activities into the conversation (Gardner 1991). Thus, these responses are relevant to the trajectory of the conversation. One typical minimal response token in Mandarin is *en*. One example of this minimal response token is shown in Fragment 1:

Fragment 1: the movie

01 Wei: 欸(0.3) 我们 看 过 以前 一 部 电影

ai women kan guo yiqian yi bu dianying

hey we watch EXP before one CL movie

'Hey! You remember that movie we watched?'

02 什么 (0.2) 什么 就 是 黑色 的 星期五 那 首 歌

shenme shenme jiu shi heise de xingqiwu na shou ge

what what just be black ASSC Friday that CL song

'What was that? That song is called Black Friday...'

03 Min: 嗯$_1$

en$_1$

mm

'Mm'

04 Wei: 那 首 叫 什么

na shou jiao shenme

that CL call what

'What was that song called?'

05 Min： 那 个 电影 就 是 叫 星期五 吧

nage dianying jiu shi jiao xingqiwu ba

that CL movie just be call Friday PRT

'That film was called Black Friday too, right?'

06 Wei: 不 是

bu shi

NEG be

'No'

In Fragment 1, two university students, Min (male) and Wei (female), are talking about a film they watched in class. In lines 1 and 2, Wei tries to recall the name of the film. The only thing she remembers is that the song from the movie was called Black Friday (line 2). Of particular interest here is that after noting the name of the song (line 2), Min produced a minimal response token *en* "mm" (line 3). *En* in this example signals to Wei that Min is listening to what is being said. Also, it acknowledges that the message was adequately received, and there was no problem in understanding the previous message. The continuation of Wei's turn in this sequence in line 4 shows that Min's minimal response *en* does not impede the progressivity of the speaker's turn. This example shows that minimal responses are relevant to the progressivity of conversation, and this relativity makes them an important matter for conducting a systematic study on such tokens.

There is a large body of research done on response tokens in English (Schegloff 1982; Jefferson 1984; Goodwin 1986; Gardner 2001), Japanese (Iwasaki 1990; Aoki 2008), German (Barth-

Weingarten 2011), and Finnish (Sorjonen 2001). In contrast, not much research exists describing response tokens in Mandarin. The majority of studies in Mandarin Chinese focus on the frequency of such tokens and their classification (Tao & Thompson 1991; Clancy et al. 1996; Deng 2008; Han 2007). This study aims to contribute to a growing body of research on response tokens in Mandarin Chinese by conducting a multimodal analysis of one of the most frequent occurring tokens in Mandarin conversation, the minimal response token *en*.

According to *Xiandai Hanyu Cidian* (Contemporary Chinese Dictionary 2005:990-991), the response token *en* 嗯 has the following meanings:

(1) 嗯 én[1]

Interjection: displays doubt, interrogation (used in questioning)

(2) 嗯 ěn

Interjection: used to show surprise or disapproval

(3) 嗯 èn

Interjection: expresses agreement or assent

But the actual use of *en* in natural Mandarin conversation may (or may not) differ from the

[1] This is the pinyin of the response token 嗯. Pinyin is an official romanization system for Chinese characters. In this thesis, I will use pinyin *en* to represent the Mandarin response token 嗯. Mandarin has 4 main tones, which are represented by 4 tone marks placed above the syllable. First tone is high and level, and represented by a macron (¯). Second tone is high-rising tone, and it is represented by an acute accent (´). Third tone is fall-rising tone, and marked by a caron (ˇ). The fourth tone is a high-falling tone, it is represented by a grave accent (`). The symbol in this pinyin syllable is a mark of second (high-rising) tone.

dictionary meaning. In this study, I will focus on the interactional functions of *en* in spontaneous Mandarin conversation.

1.2 Phonetic variations of the minimal response token *en*

En has two phonetic variations in the current data: bilabial nasal [m] and mid central nasalized vowel [ə̃]. Current research focuses on both variations. The two phonetic variations in the data are also indicated in the transcript: 嗯 $_1$ *(en*$_1$) for bilabial nasal [m] and 嗯 $_2$ (*en*$_2$) for mid central nasalized vowel [ə̃].

1.3 Participation framework in conversation

The notion of participation was first proposed by Goffman (1981), which is about the relationship between conversational participants. In his accounts of participation (Goffman 1981), Goffman deconstructs the speaker into several kinds. They are Animator (the person producing the talk), Author (the entity for constructing sentences at issue), Principal (the party who is socially responsible for what is said), and Figure (a character depicted in the talk). Goodwin (2006) critically comments on Goffman's model that his analysis mostly focuses on the talk of the speaker in isolation from the actions performed by the listener. Goodwin (1981) proposes his own model of participation framework which is based on an analysis of the practices through which conversational participants, both speakers and listeners, build actions together (Goodwin 1981:225). Speakers and listeners during a communication participate in a common course of action. Speakers initiate an action, while listeners display their orientation to the speaker through different verbal and non-verbal means, one of which is the production of brief vocalizations, such as continuers. Speakers while speaking can modify their speech to adapt to the listener's (dis)engagement in the talk in progress. Thus, participation is an action which demonstrates

participants' involvement in communication. In this study, I will use Goodwin's notion of participation to analyze the interaction between the speaker and the listener, because it addresses not only the speaker, but also the listener.

1.4 Sequential structure of conversation

Conversational studies revealed that interactions consist of sequences of pairs of actions, which are called "adjacency pairs" (Schegloff 2007). An adjacency pair consists of two turns performing two actions, in which first pair part (also known as a "first position") performs an initiating action, such as questions, offers, announcements, and second pair part (also known as a "second position" or "responsive position") responds to this action by giving a pair-type related responses, such as answers, acceptances or declinations. Pair-type related response refers to a second pair part that is immediately relevant and expectable in relation to a first pair part. In literature, this notion is called conditional relevance (Schegloff 1972). Fragment 2 below illustrates an example of an adjacency pair.

Fragment 2: (modified from Schegloff (2007:22))

01	Ton:	How are you.
02	Mar:	Fine.

In Fragment 2, in line 1, Ton initiates an action, i.e. asks a question, which is the first pair part of the adjacency pair. In line 2, Mar responds to this action with *Fine* which is the pair-type related response, and is the second pair part of the adjacency pair.

These minimal pairs of actions can be expanded before first pair part, between first and second pair parts, and also after the second pair part (Schegloff 2007). Of particular interest in this study is the minimal expansion after the second pair part, which is called "third position". Minimal third position expansion is an important place at which speaker A displays that a prior speaker B has or has not adequately produced a second pair part conditionally relevant to the speaker A's first pair part (Schegloff 1992).

Fragment 3: (modified from Schegloff (2007:119))

01	Nan:	hhh Does he have his own apart [ment?]
02	Hyl:	[°hhhh] Yea:h,
03	Nan:	Oh:,
		(1.0)
04	Nan:	How did you get his number,

Let us analyze the sequential structure of Fragment 3 above. In line 1, Nan initiated an action. i.e. a request for information, which is in the first position in this sequence. It is followed by a second pair part which is the delivery of information, a positive response with the response token *Yeah* in line 2. In line 3, Nan deploys a "change of state token" *oh* (see Heritage 1984) to register the receipt of the information in Hyl's preceding utterance in line 2, and to register it as new. This *oh* is placed in the third position of the sequence. As it can be seen from line 4, Nan starts a new sequence, thus *oh* in third position here also closes the sequence (lines 1 to 3).

In the current data, a free-standing token *en* can be placed in a second position as a response to the action in the first pair part (see detailed discussion in Sections 4.1, 4.2 and 4.3). In addition, *en* can be placed at "third position" in response to the second pair part. This minimal expansion of the second pair part does not project any further talk in the sequence, it is rather designed to propose its closing (Schegloff 2007). For this reason, third position object in the literature is called "sequence closing-third". This use of *en* will be discussed in detail in Section 4.4.

1.5 Research questions

In recent years, conversation analysts have paid a great deal of attention to the study of the listeners' responses in different languages, because these responses play an important role in the perception and interpretation of the output produced by the speaker (Goodwin 1986). However, little attention has been paid to the response tokens in Mandarin Chinese. To the best of my knowledge, no study has been done on the usage and functions of the response token *en*, one of the most typical response tokens in Mandarin Chinese (Zheng 2007:56). Thus, this study intends to examine the interactional functions of this token in natural conversation. Specifically, this study will focus on the following research questions:

- What are the interactional functions of response token *en*?
- What is the sequential position of response token *en* in a larger sequence of interaction?
- What are the prosodic features of response token *en*?
- What are the bodily-visual movements associated with the production of the response token *en*?

By answering these questions, this study will contribute to our understanding of how the response tokens are used in Mandarin conversation.

1.6 The organization of the thesis

This thesis is structured as follows. Chapter 2 gives an overview of the previous studies conducted on the use of response tokens in Mandarin and other languages. Chapter 3 presents an overview of the data and the methodology used in this study. Chapter 4 explores the typical interactional functions of the Mandarin *en* in different linguistic environments: such as a continuer, a confirmation and an acknowledgement token, as well as a token that registers the receipt of the listener's responses. Chapter 5 discusses the findings, and implications of this research in Chinese language teaching.

Chapter 2 Literature review

This chapter aims to review the existing literature on response tokens in Mandarin Chinese and other languages, such as English, Japanese, and German. First, I will discuss the definition of response token (Section 2.1). Second, I will introduce previous studies on response tokens in Mandarin (Section 2.2). Finally, I will review existing studies on response tokens conducted in other languages (Section 2.3).

2.1. The definition of response tokens

There have been many discussions concerning the classification of response tokens. The most widely used term for these items is “minimal responses” (Gardner 2002; Fishman 1983). However, some researchers call them “backchannels” or “backchannel responses” (Yngve 1970), and others call them “reactive tokens” (Clancy et al. 1996), “receipt tokens” (e.g. Atkinson 1992), or "hearer signals" (Bublitz 1988), etc. In what follows, I will provide an overview of some of these terms, and account for the use of the term “minimal response token’ in this study.

There is no agreement on how to call those brief utterances. The most frequently used term for these brief utterances is ‘backchannel’, which was first used by Yngve (1970). According to Yngve’s definition, backchannel responses include a large variety of response types (brief questions, collaborative finishes, head nods, brief vocalizations such as *yeah* and *mm hm*), the main function of which is to show listenership or a receipt of input. Duncan (1974) extended the notion of a backchannel term by including repetitions, requests for clarification, word supplies,

and brief restatements. The term backchannel implies the existence of two channel interactions, during which one of the speakers takes the predominant channel by holding the floor and conveying the message, while the other speaker takes the "back" channel and plays a role of a listener.

The term "reactive tokens" was introduced in the study conducted by Clancy et al. (1996). They defined reactive tokens as "a short utterance made by an interlocutor who is playing a listener's role and who does not claim the floor during the other interlocutor speakership" (Clancy et al. 1996:356).

According to Bublitz, "hearer signals" are "linguistic and non-verbal forms which the hearer employs to signal… and to confirm that he is the hearer" (Bublitz 1988:169). Such signals do not state a position or manifest an attitude, but rather take note of what have been said and meant.

Although some forms and functions of the Mandarin *en* suite the aforementioned terms, these definitions do not capture all of its fundamental features. Gardner (2001) in his studies on listener responses, used the term "response token". According to him, response tokens are "one class of conversational objects whose primary functions are not to make reference to the world, but to provide some information on the course the talk is taking" (Gardner 2001:14). Many studies which focused on the analysis of interactional functions adopted the term "response token" (e.g. Aoki 2008; Gorish 2010; Yang 2013). This research follows this definition and also uses this term. However, taking into account the short form of *en*, this study defines it as a "minimal response token", which is a monosyllabic utterance produced by conversational participants in a responsive position.

2.2 Previous studies on response tokens in Mandarin

Over the past decades, there have been studies conducted on some of the listener responses in Mandarin conversation. One of the most influential papers on response tokens is Clancy et al. (1996), which classifies all response tokens in Mandarin "based partly on its form partly on its sequential function" (Clancy et al. 1996:354). They categorized "reactive tokens" in their data into several types: backchannels, reactive expressions, collaborative finishes, repetitions, and resumptive openers. Backchannels are non-lexical vocalic forms that serve as continuers and include utterances such as *uhm, a, ao, en,* and *eh*. Reactive expressions are short non-floor taking lexical phrases or words, such as *zheyang hao, shi a,* and *dui*. Collaborative finishes occur when the non-primary speaker finishes the previous speaker's utterance. Repetitions are used when a non-primary speaker repeats a portion of a primary speaker's speech. Resumptive openers are non-lexical elements which are used at turn-initial positions, such as *ai* or *e*. Many subsequent studies based their analysis on this classification (e.g. Deng 2008: Xu 2009). In this study, they listed *en* as an example of a broad category 'backchannels', however, they did not show how this token is different from other 'backchannels'.

Yin (2010) made a full survey of listener responses, and divided them into 14 distinct categories in terms of their pragmatic functions. For example, confirmation or agreement tokens, such as *a, shi, jiushi, yeshi*, etc; negation tokens, such as *bu, bukeneng, cuole*, etc; tokens displaying listenership (*en, ai, a, o, e*, etc.); tokens displaying the listener's awareness of the message (*o, ou, wa, zhidao, dongle, mingbai*, etc.), and others. In his study, the Mandarin response token *en* is classified into three types based on its pitch movement. The three types are *en* with falling pitch movement, e*n* with rising intonation contour, and *en* with rise-fall intonation contour. The first type of *en* is a token for agreement; *en* with rising pitch movement is a token for doubting and

questioning the previous statement; *en* with rise-fall pitch movement is a token for expressing surprise. Yin's research includes all possible response tokens in Mandarin conversation, which make his study very unique and useful. In this regard, under the same group of interactional functions there were many response tokens, which include monosyllabic and bisyllabic utterances, as well as whole phrases.

In their study on Mandarin backchannels (such as *ao*, *ai* and *dui*) Tao and Thompson (1991) presented the following findings. First, Mandarin speakers rarely use backchannel responses (only 10 out of 119 speaker changes were backchannel responses). Second, out of 10 backchannels there were no continuers, and none of them occurred in overlap. In their data, the Mandarin backchannels functioned as claims of understanding, signals of confirmation, or acknowledgment of agreement. Although this study attempted to assign distinct interactional functions to backchannel responses, they did not analyze sequential features of each token.

The above listed research on Mandarin response tokens mainly focused on categorization. As it can be noticed from the previous studies, all response tokens, including *en*, were lumped together as one homogenous group. None of the tokens were analyzed individually in terms of their pragmatic and interactional uses in Mandarin conversation. Thus, research on the interactional functions of each different response token is sorely lacking.

There are a few studies, which investigated the pragmatic functions and environments in which some Mandarin response tokens are analyzed individually (e.g. Tsai 2001; Xu 2009; Wang et al. 2010). For example, there is a body of research on interactional and discourse functions of the response tokens of *hao* and *dui*. Miracle (1989) observes that *hao* has three distinctive functions in an interaction: it may play a role in the development and closure of a request; it can serve as

an assertion marker and it can mark the transition to a new topic. Shao and Zhu (2005) divide the discourse functions of *hao* into three types: active answering function which includes appraisal and affirmation, passive answering function such as acceptance of indirect refusal, and discourse cohesion function such as transition and closure. There are also a number of studies on the Mandarin token *dui*, such as the one conducted by Tsai (2001). She observes that the receipt token *dui* can be a positive answer to interrogative forms which imply a proposition; as well as an agreement token which affirms the other speaker's utterance. Yu (2004) looked at different sequential positioning of *dui*, and proposed that apart from being an agreement token, this token signals the transition of interactional sequences. In a comparative study on the discourse functions of *hao* and *dui*, in combinations with small particles (such as *le, a, ba*), Wang et al. (2010) found that *hao* is used to express acceptance of the other speaker's act during interactions; while, *dui* acknowledges the propositional content of the utterance produced by the other speaker (Wang et al. 2010:242).

Studies on multiple sayings of Mandarin response tokens are very scarce in literature. One of the valuable exceptions is Yang (2013). She investigated multiple sayings of *dui dui dui* (right right right) from a multimodal perspective. She found two types of *dui dui dui* in her data: the first type displaying affiliation with the speaker's previous assertion; and the second type functioning as a confirmation of the recipient's collaborative completion of the speaker's turn. One of the merits of this study is that the analysis was done from a multimodal perspective, which considers sequential, prosodic and bodily-visual features. Yang's research findings suggest that each type of *dui dui dui* has distinct prosodic and visual features, and also carries different interactional functions.

To the best of my knowledge, there have only been a few studies that explore the interactional and discourse functions of the minimal response token *en* in Mandarin. Gao (2007), basing her research on everyday conversation, considered *en* to be a "discourse marker". The discourse marker *en* supports the interlocutor during his/her speakership while maintaining a smooth flow of conversation. She describes three roles of the discourse marker *en*: "turn holding *en",* which fills the pause during the turn and lets the turn holder to keep the turn; "turn controlling *en"*, which signals the speaker to stop his speakership and yield the floor to the *en* utterer; and "supporting *en*", which is produced by a listener to support the speaker during his turn production. Gao (2007) also analyzed *en* being an interjection. She argues that the interjection *en* with different lexical tones has distinct pragmatic functions. When appearing in a turn initial position and followed by further talk by its utterer, *en* can display agreement (flat tone), surprise (high-rising tone), blame (high-rising tone), begging (fall-rising tone), and approval (high-falling tone). When used as a single standing token in its own turn, *en* can represent consent (flat tone), doubt (high-rising tone), refusal to do something (fall-rising tone) and agreement with the previous statement (high-falling tone). Although Gao's study focused on *en* as an interjection and discourse marker, she did not point out that *en* may also serve as a response token. Also, her analysis was not clear in describing how interactants orient to *en*, or to which action this token may respond.

Zheng (2007) pointed out two pragmatic functions of the token *en*. The first function is a response token; the second is a discourse marker. When serving as a response token, *en* maintains the continuity of a discourse. Being a discourse marker, *en* represents the speaker's emotional state. When produced as a discourse marker, it can also represent the process of thinking; express hesitation or doubt; or display self-confirmation. She also found evidence for

en being a turn holder which is when a speaker pauses during the production of a turn but intends to keep holding the floor. In such a case, the speaker while planning the production of his/her next utterance usually produces *en*. Zheng's study is not based on everyday spontaneous conversation and she uses dialogues from modern Chinese literature as data to support her argument. Although the study focuses on *en* as a discourse marker and response token, there is no detailed analysis of *en* as a response token having different interactional functions.

Studies by Gao (2007) and Zheng (2007) analyzed pragmatic functions of the token *en* and its role in the turn taking system. However, although they mentioned the role of the Mandarin *en* as a response token, they did not analyze what exact interactional functions it can fulfill in different sequential environments. Thus, a systematic study on the interactional functions of the response token *en* based on its sequential placement is needed in Mandarin studies.

2.3 Studies on response tokens in other languages

There have been a great number of studies conducted on different types of response tokens in terms of their interactional functions, prosodic features and body behavior. There have also been many studies examining nonverbal responses such as head nods. In this section I will review how listener responses in other languages, such as *mm* in English, *nn* in Japanese, *mm* in Finnish, and *ja* in German, have been studied.

These response tokens can be distinguished from one another via their interactional functions during a conversation. For instance, Schegloff (1982) and Sacks (1992) labelled such items as *uh huh, mm hm* and *yes* "continuers", the primary function of which is to display the evidence of attention and display understanding that the extended turn is still in progress and not yet

complete. Jefferson (1984) conducted a research on a class of tokens called "acknowledgement tokens", referring to them as a pre-shift objects. She also pointed out the difference between *yeah/yes* and *mm hm* in terms of their readiness to speaker shift. "*yeah* can exhibit a preparedness to shift from recipiency to speakership, while *mm hm* exhibits what I will call "Passive Recipiency", where "Passive Recipiency" is that its user is proposing that his co-participant is still in the midst of some course of talk" (Jefferson 1984:200). Nevertheless, Gardner (2001) argues that *yeah,* depending on its sequential position and prosodic shape, may serve as an acknowledgement token and as a continuer. Gardner (2001) also made a survey of 8 response tokens in English, and grouped them into 4 categories: continuers (*mh hm* and *uh huh*), acknowledgement tokens (*yeah* and *mm*), newsmarkers (*oh* and *right*), and "change of activity" tokens (*okay* and *alright*).

Heritage (1984) distinguished the response token *oh* among others referring to it as a "change-of-state token", which occurs when a listener is informed of something. This response particle "is used to propose that its producer has undergone some kind of change in their current state of knowledge, information, orientation or awareness" (Heritage 1984:299). Siitonen & Wahlberg (2015) found that in response to a proposal, the Finnish response particle *mm* is the most frequently occurred particle (in comparison with *jaa* and *joo*) in their data. They observed that the speaker of *mm* acknowledges the prior turn as heard and understood, and thus, unproblematic. Also, this particle does not imply the recognition of willingness to elaborate the interpretation proposal further, and therefore, was interpreted as the least encouraging response token to a proposal (Siitonen and Wahlberg 2015:78). Golato and Fagyal (2008) studied single and double sayings of the German response token *ja*. They pointed out that their analysis of single response token *ja* showed this token can be a response to a yes-no question, can serve as a continuer or as

an acknowledgement token. Whether *ja* functions as a continuer or an acknowledgement token depends on its sequential placement in the dialogue.

Gardner (2001) examined the intonation contour of the English token *mm*. He claimed *mm* to be a flexible response token in English. With a fall-rising intonation contour, it can be treated as a continuer. The response token *mm* with a falling intonation contour can also function as an acknowledgement token. The acknowledgement *mm* claims an adequate receipt in understanding of the speaker's turn. Golato and Fagyal (2006) found that double sayings of German response token *jaja* having two distinct nuclear contours accomplish separate interactional goals. The above findings of two studies suggest that participants use different prosodic cues in producing the response tokens to accomplish different interactional tasks.

There has been a growing body of research on body behavior which may also can function as listener responses. Knight et al. (2006) analyzed the relationship between the length of head nods and their functions. Their findings show that short head nods function as response tokens, whereas long head nods besides functioning as response tokens also tend to have an additional discourse meaning. Aoki (2008) conducted a multimodal analysis of the Japanese response token *nn* and head nods made by recipients during extended talk. She characterized this token into two groups: plain tokens, consisting of a single *nn* and a head nod, and complex tokens, typically consisting of a succession of *nn*s and head nods. The study found that when plain tokens are placed in TRPs[2], it serves as a backchannel, and participants find it problematic because it shows lack of affiliation. Therefore, plain tokens are often placed at non-TRPs to display the receipt of the message, which encourages the speaker to continue. Studies on head nods show that they also

[2] TRP (Transition Relevance Place) refers to the place at the possible completion of a current turn constructional unit (Sacks et al. 1974). At this place, turn transition may, but need not, occur. Turn constructional units (TCUs) are building blocks of a turn in conversation. TCUs include sentential, clausal, phrasal and lexical constructions (Sacks et al. 1974).

can function as response tokens, and their interactional uses depend on its length and the placement in the sequence of talk.

The above studies all show that response tokens can perform different interactional functions in conversation depending on their sequential position, prosodic features and bodily-visual features. Face-to-face interaction is by default multimodal, in which gestures, body postures, head movements, prosody, lexico-syntactic constructions, sequential position, and other modalities work together. The communicative work that is achieved by one modality may be supported by another modality (Stivers and Sidnell 2005). Therefore, to understand how Mandarin *en* is used in face-to-face interaction, it is essential to perform a multimodal analysis of different practices associated with its production.

2.4 Summary

In this chapter, I have introduced different definitions of response tokens which have been used by previous studies. Furthermore, I also presented the research done on response tokens in Mandarin conversation. Finally, I have reviewed studies on response tokens in languages other than Mandarin, including English, German, Japanese, etc. It is evident from the literature review that there is no systematic study of the Mandarin response token *en* and its usages in interaction. In the next chapter, I will be discussing the methodology employed in this study.

Chapter 3 Methodology

3.1 Data

The data for this study consist of 6 hours of everyday spontaneous Mandarin conversation between 19 native speakers of Mandarin, forming 8 dyads and 1 triad. Each dialogue was conducted in Mandarin Chinese and lasted from 30 to 60 minutes. The conversations were collected in the city of Edmonton, Canada. The real names of the participants were replaced by the pseudonyms. All the participants are university students, aged between 20 to 25. 12 of them are female, and 7 of them are male. The topics of their conversations included university life, study, food, travel, social relationships, career, etc. Video recordings mainly took place in classrooms and study halls. All of the data were transcribed according to the GAT-2 transcription system (Selting et al. 2009) with modifications adapted to Mandarin. Arrows in the transcript above the utterance indicate the placement of nods and their direction (up arrow ↑ for upward component of a nod, and down arrow ↓ for the downward component of a nod (transcription symbols for nods borrowed from Whitehead (2011)).

A sample of transcribed data is presented as follows:

Sample

06 Rui： 啊 有 一 个 什么 技术;

a you yi ge shenme jishu

ah have one CL what technique

'There is one technique'

Su head: ↓↑↓↑

07⟶ Su： [嗯$_1$.

[en$_1$

*[*mm

‘Mm’

08 Rui ： [或者 是 什么 领域;

[huozhe shi shenme lingyu

[or be what field

‘Or (there is) a field’

The number on the left of the dialogue represents the sequence of intonation units. The first line is the Chinese version of the turn, and the second line provides the pinyin (Romanized pronunciation of characters) for each corresponding word. The third line outlines a word-for-word translation. The last line presents the free translation.

The occurrences of the minimal response token *en* in the data are annotated through ELAN Version 4.9.4 (Brugman and Russel 2004), a professional software program for making annotations on video recordings. The prosodic features of each instance of *en* were analyzed with PRAAT Version 6.0.18 (Boersma and Weenink 2016).

3.2 Methodology

The methodology used in this study includes Conversation Analysis (CA) (Ten Have 1999) and Interactional Linguistics (IL) (Couper-Kuhlen and Selting 2001). With the help of CA, I will be able to analyze the organization of sequences in given dialogues, and how actions within these sequences are performed by participants. IL will be used to examine what interactional functions are formed by particular linguistics forms.

3.2.1 Conversation Analysis

Conversation analysis (CA) is an approach that studies conversation through analyzing sequential patterns or organizations to find the practices which make mutually comprehensible communication and action possible in interaction. CA is strictly rendered on recorded, spontaneously occurring interactions that are subsequently transcribed and analyzed with respect to any tiny details which occur in stretch of talk. There are four basic concepts that underpin CA analysis: turn-taking, turn design, social action and sequence organization (Drew 2005).

Turn-taking is the most fundamental form of conversation, with speakers literally taking turns between speeches. When analyzing the sequential positions of the Mandarin *en* in a defined action type within which it appeared, it is important to note to *what* it is oriented and *how* it is in turn responded by the interlocutor.

When designing a turn, the speaker selects what action they will perform and what details (words, syntactic structures, prosody, body behavior) will go into that turn. In my analysis, I will be looking at verbal and non-verbal means of a turn design which solicit the production of the Mandarin *en* in a subsequent turn.

When people converse, they are constructing their turns to perform a course of actions. CA research is mostly concerned with how participants understand each other's actions. Specifically, it analyzes the actions performed during speech. In this study, I also explore what actions *en* performs and what actions it responds to.

Turns are connected with one another in a systematic sequence. The most basic sequence organization is an adjacency pair, or pair of actions, in which one speaker initiates an action, and the other responds with an action paired with that first action. Utterances become intelligible when they are "tied to their particular sequential location in a stretch of talk" (Pomerantz and Fehr 1997:68). Therefore, the function of each response token at the beginning, in the middle and the end of a larger sequence of talk can be different.

The basic concepts of CA analysis described above will help us to see the differences of Mandarin *en* in conversation, explicate the actions to which it responds to, and how it is understood by its recipients.

3.2.2 Interactional Linguistics

The methods of Interactional Linguistics (IL) have also been adopted in this study. IL takes an interdisciplinary and cross-linguistic perspective on language. Its goal is to understand how languages are shaped by interaction and how specific interactional practices are molded in different languages. The key questions IL is trying to find an answer for are: "What linguistic resources are used to articulate particular conversational structures and fulfil interactional functions? And what interactional function or conversational structure is furthered by particular linguistic forms and ways of using them?" (Couper-Kuhlen and Selting 2001:3). Therefore, by

answering the fundamental questions of IL, we will be able to explore what interactional functions the minimal response token *en* can accomplish in Mandarin conversation.

Based on the methodology of CA and IL, we analyze how the minimal response token *en* is shaped by interaction and what interactional goals are achieved through its production. Based on these notions, in this study, I will look at:

1. The interactional functions of the Mandarin *en* in conversation.
2. The placement of *en* within a larger sequence (e.g., in the middle or at the end of the larger sequence).
3. The placement of *en* within a turn. Was the utterance produced at a point of possible grammatical and intonational completion, or did it occur simultaneously with the production of a previous turn, in overlap? Previous studies mention that backchannels which occur in overlap are more likely to serve as continuers while backchannels produced at points where utterances are grammatically, intonationally and pragmatically complete are more likely to be acknowledgment tokens.
4. Prosodic features include voice quality, intonation pattern, and duration. Phonetic details and social actions, as mentioned in Ford and Couper-Kuhlen (2004), are interconnected. That is, social activity can be influenced by the change of phonetic features deployed during an utterance. For instance, Yin (2010) found that the Mandarin *en* with different intonation contours has different conversational functions and meanings. Besides intonation contour, other prosodic features also play an important role in the understanding of the produced utterance. Gardner (1997), in his study on the conversational object *mm*, noticed that the receipt token *mm* produced immediately after the previous turn without any pause reveals that there is no problem in articulation and understanding. I will analyze prosodic features of Mandarin token *en* to see if

there are any noticeable differences between each other.

5. Body-visual features which accompany the production of the response token *en* including gaze direction, body posture and head nods. The visuospatial modality also has a communicative import for interaction (Stivers and Sidnell 2005). For instance, Heath (1992) showed that listeners use gestures to display a co-participation; and C. Goodwin (1979) observed that gaze accomplish an important role of selecting a recipient in multiparty conversation. In this study, I also analyze the body behavior that accompanies the production of response token *en*.

When examining the phonetic and bodily-visual features of *en*, I only considered those features that are relevant to the use of the minimal response token *en* as oriented to by participants. The prosodic features relevant to the production of *en* include voice pitch contour and duration. Regarding bodily-visual features, gaze direction, head movements, and body posture seem to be relevant to the use of *en* in the data.

3.3 Summary

In this chapter, I introduced the data, data transcription method, and the basic notions of Conversation Analysis and Interactional Linguistics as the methodology used in this study. As face-to-face interaction is a multimodal interaction, this study also seeks to analyze the different modalities associated with the production of the minimal response token *en*, including vocal-aural modality (spoken language with prosody) and the visuospatial modality (gaze, head movements, and body posture). In addition, the interactional functions of the response token *en* are identified.

Chapter 4 Results: the interactional functions of the minimal response token *en*

In the current data, there are 757 instances of *en* that are used as response tokens, including single and multiple tokens. However, a big portion of the data consists of single token *en* occurrences. Table 1 illustrates the frequency of all *en* tokens in the data. It can clearly be seen from the table that single usage of the token *en* far outnumbers the usage of multiple tokens of *en*, making up 93.6% of the whole number of tokens. Among all multiple tokens, double sayings of the *en* token are the most frequent which accounts to 4.3% of all tokens. Furthermore, it also can be noted from the table that the more tokens there are, the less frequent they become.

Multiple tokens are not within the scope of this study. This research focuses its analysis uniquely on single uses of *en*.

Table 1 Distribution of single and multiple *en* tokens in the data

Single/multiple token *en*	Occurrence	Percentage
en	709	93.6%
en en	33	4.3%
en en en	6	0.8%
en en en en	5	0.7%
en en en en en	3	0.4%
en en en en en en en	1	0.2%
Total	757	100%

The single *en* tokens in my data have four interactional functions. One function of *en* is a continuer displaying the listener's alignment with the speaker and acknowledgement that the turn is still in progress. It can be seen from Table 2 that continuer *en* is used more frequently than all other uses, making up 66.7% of all occurrences of the token. The second function of *en* is to display acknowledgement. The acknowledgement token *en* is produced by a listener to claim understanding of what is being said. Making up 17.8% of all *en* tokens, this token is the second most frequent in the data. The third and the fourth functions of the token are different from previous *en*s in that they are produced by a speaker. The third function of *en* is to indicate confirmation, which immediately follows the listeners' clarification request. There are only 31 confirmation tokens in the data, which make this use of the token the least common among the four uses in conversation. The fourth function of *en* is to register the receipt of listeners' responses. Table 2 provides an overview of the four functions of *en* in the data.

Table 2 The four functions of *en* in the data

Function of the token	Occurrences	Percentage
continuer	473	66.7%
acknowledgement	126	17.8%
confirmation	31	4.4%
register of receipt	79	11.1%
Total	709	100%

In the following sections, I provide a detailed account of the four uses of the response token *en* in the data.

4.1 Continuer

Activities such as telling a story, telling a joke, descriptions, and explanations can be considered as larger units of talk. According to Houtkoop and Mazeland (1985), these units are called Discourse Units, whereas Sacks (1969) describes these units as larger projects. Discourse Unit is a unit that involves a larger stretch of talk and performs actions that take more than one turn to complete. For instance, storytelling is considered as one type of the larger unit in interaction. It has a recognizable trajectory. As noted by Jefferson (1978), it usually starts with a "story-entry device", and by using such a device the speaker claims the floor, and signals to the listener that the larger turn is about to start. This blocks the listener from treating the completion points of further TCUs as normal TRPs (Houtkoop and Mazeland 1985). At the possible completion of storytelling, the speaker may produce a "story-ending device" (Jefferson 1978), which can be a summary assessment. In the middle of the storytelling, the listener usually does not take the floor, but reacts to the story by providing short remarks. At the possible closure of the telling, the listener may provide a comment or appreciation, which shows their understanding of the storytelling activity as completed. This provides evidence that participants orient to storytelling as an interactional unit. Thus, we can consider this type of larger stretch of talk as a Discourse Unit.

During the production of a larger stretch of talk such as storytelling, speakers and listeners have asymmetrical communicative roles. Speakers do most of the talking, and listeners produce minimal vocalizations, one common function of which is to acknowledge and encourage the speaker to continue speaking. Schegloff (1982) labelled these minimal vocalizations as continuers. The main function of such small portions of talk includes displaying attention and

showing understanding that the extended turn is still in progress. For example, continuers usually signal to the speaker that the listener understands and follows what is being said. The listener producing a continuer in the midst of a speaker's extended turn passes the opportunity to speak, and gives back the floor to the speaker. In this regard, continuers are rarely found in further talk by their producers (Gardner 2001). In English, such tokens were studied extensively by linguists, and the tokens exhibiting the function of continuer include *uh huh* and *mh mm*. In this section I will show how the Mandarin minimal response token *en* can function as a continuer. I will also introduce its sequential, prosodic, and bodily-visual features.

According to Schegloff (1982), continuers usually appear at the end of speaker's turn-constructional units (TCUs); that is, at transition relevance places (TRPs). In my data, the continuer *en* can be placed at two different positions in relation to the speaker's turn: one is at the end of a speaker's TCU, the other is in the middle of a speaker's TCU. Fragment 4 and 5 in this section exemplify how *en* can be used as continuer in Mandarin conversation.

Fragment 4 demonstrates the use of *en* as a continuer at the end of speaker's TCU. In this fragment, Shu (female) and Tan (male) talk about the languages (Russian and French) selected to record documents in the United Nations, and why other languages, like English and Chinese, were not selected. *En* is produced five times by Shu (lines 6, 16, 18, 20 and 22), during Tan's extended talk, but Shu does not take the floor, and Tan maintains primary speakership.

Fragment 4: cont1 (languages)

01 Shu: 就 说 它 是 –

jiu shuo ta shi

just say 3SG be

'Let's say it is,'

02 在 就 是 在 世 就 世界;

zai jiu shi zai shi jiu shijie

at just be at world just world

'in the world'

03 Tan: 只 有 四 从 四 大国 中间 选 一 个 –

zhi you si cong si daguo zhongjian xuan yi ge

only have four from four big countries among choose one CL

'Only one language can be chosen from four leading countries.'

04 选 选 两 门 语言;

xuan xuan liang men yuyan

choose choose two CL languages

'Two languages can be chosen,'

05 或者 是 一 到 两 门 语言;

huozhe shi yi dao liang men yuyan

or be one to two CL languages

'or from one to two languages.'

Shu head: ↓↑↓↑

06 ⟶ Shu： **嗯**$_2$.

en$_2$

mm

‘Mm.’

07 Tan： **英语 肯定 不 排 在 其中**;

yingyu kending bu pai zai qizhong

English sure NEG arrange at among

‘English is not among these languages.’

Shu head: ↓↑↓↑

08 **这么 不 完善 的 语言**;

zheme bu wanshang de yuyan

such NEG perfect ACCS language

‘Such an imperfect language.’

09 **所以 美国 去掉 了**.

suoyi Meiguo qudiao le

therefore America exclude CRS

‘That’s why America (American English) is excluded.’

10 中文 过于 完善 [已经 也 不 适合 了.

Zhongwen guoyu wanshang [yijing ye bu shihe le

Chinese too perfect [already also NEG suitable CRS

‘Chinese is too perfect, also is not suitable.’

11 Shu： [那 你 不 是 不 是;

[na ni bu shi bu shi

[Then you NEG be NEG be

‘In that case, you, no no,’

12 英语 不 是 那么 完善 –

Yingyu bu shi name wanshang

English NEG be so perfect

‘English is not as perfect,’

13 美国 确实 –

Meiguo queshi

America indeed

‘America is indeed.’

14 Tan： 你 知道 为什么 英语 的 比如 说 –

ni zhidao weishenme Yingyu de biru shuo

you know why English ASSC such as say

'You know why in English, for example,'

15 你 的 化学词 物理词;

ni de huaxueci wulici

you POSS chemical terms physical terms

'Some terms in chemistry, in physics,'

Shu head: ↓↑

16 ⟶ Shu： 嗯$_2$.

en$_2$

mm

'Mm.'

Shu head: ↑↓↑↓↑↓↑↓ ↑↓↑↓↑↓

17 Tan： 都 是 从 拉丁语 里面 直接 copy 过来 的 嘛 –

dou shi cong Ladingyu limian zhijie copy guolai de ma

all be from Latin inside straight copy come over PRT PRT

'all of them were copied from Latin language.'

18 ⟶ Shu： 嗯$_2$.

en$_2$

mm

‘Mm.’

19 Tan： 就 是 因为 那 些 词 根本 就 不 是 英语;

jiu shi yinwei na xie ci genben jiu bu shi Yingyu

just be because that CL words at all just NEG be English

‘Because these words are not originated from English.’

Shu head: ↓↑

20 ⟶ Shu： 嗯$_2$.

en$_2$

mm

‘Mm.’

21 Tan： 就 标准 的 拉丁语;

jiu biaozhun de Ladingyu

just standard ACCS Latin

‘(They were originated) from standard Latin.’

Shu head: ↑↓↑↓↑↓↑

22 ⟶ Shu： 嗯$_{1}$.

en$_{1}$

mm

‘Mm.’

23 Tan： 所以 英语 因为 没 有 啊;

suoyi Yingyu yinwei mei you a

therefore English because NEG have PRT

‘That is why English is not included.’

24 比如 说 –

biru shuo

for example say

‘For example,’

25 俄语 比 英语 高级 就 高级 在 –

E’yu bi Yingyu gaoji jiu gaoji zai

Russian compare English superior just superior at

‘Russian is superior than English in… …’

In the sequence above, Tan provides two reasons for why languages such as English and Chinese were not selected to record documents in the United Nations. At lines 3 to 10, Tan produces his extended turn by arguing that the UN can only select one or two languages, and that English and Chinese are not among these languages due to their imperfections. During this turn, in line 6, Shu produces her first *en* after Tan's TCU in line 5. Here, after producing *en*, Shu withholds from taking the floor, and displays her understanding that Tan's turn is still in progress. Thus, *en* is used as continuer, encouraging Tan to continue his turn. Tan orients to *en* as not competing for the floor, and continues his turn starting with *ni zhidao weishenme* "do you know why…" (line 14). Beginning the turn with these words is another signal that Tan is going to continue speaking. In lines 14 to 21, Tan provides another reason for why English in particular was not chosen for documentation in the UN. In lines 16, 18, 20 and 22, Shu produces *en* another 4 times, all occurring in the middle of Tan's turn. Every time Shu produces *en*, she withholds from talking, displaying her attention and encouraging the speaker, Tan, to continue. Once again, Tan orients to these *en*s as not competing for taking the floor. Thus, these 4 response tokens are used as continuers. To conclude, in this fragment, Shu, by producing *en,* expresses her understanding of the prior speaker's turn as still in progress. She passes up the opportunity to speak and encourages Tan to continue. These *en*s display the listener's orientation to the ongoing talk as continuous. As it can be noticed from lines 6, 18, 20 and 22, all *en* continuers in the above fragment appear at the end of the speaker's TCUs (lines 5, 17, 19 and 21).

Furthermore, I will discuss the prosodic and visual cues of *en* in Fragment 4. Figure 1 illustrates the prosodic features of the continuer *en* in line 16 of Fragment 4. Although the pitch movement displayed in PRAAT appears to be a level pitch contour with a very small pitch range (from 199Hz to 179Hz), repetitive listening by Mandarin native speakers reveals that *en* in this

example has a falling pitch movement. *En* used as a continuer in this example has a total duration of 0.3 seconds. The phonation type of the sound is modal.

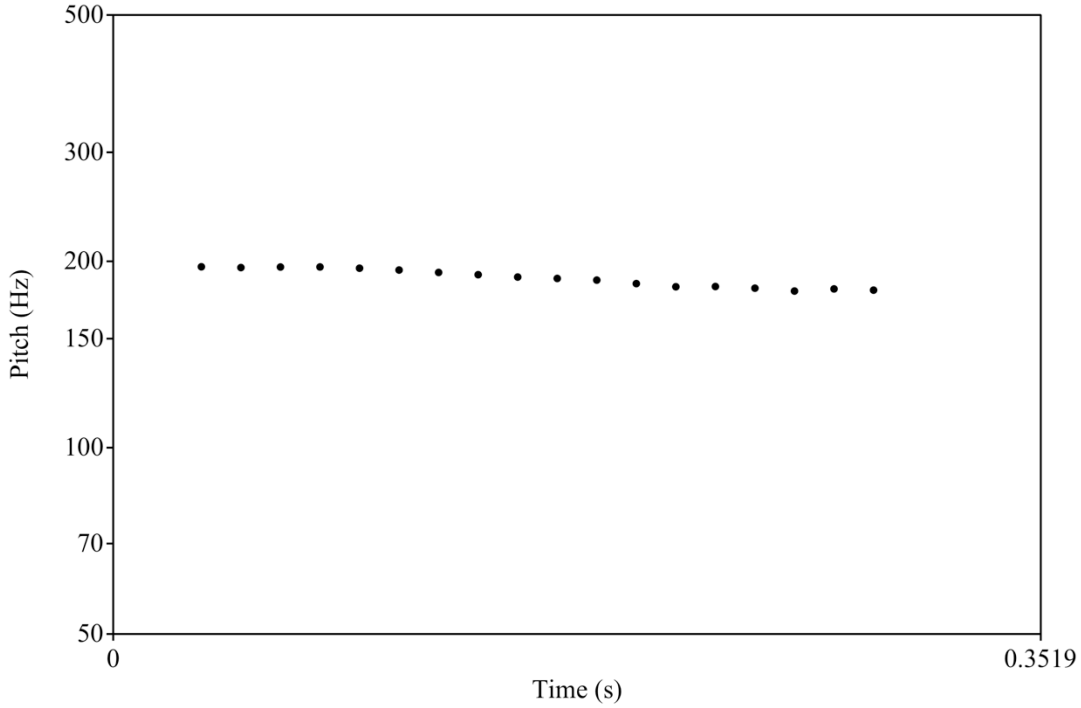

Figure 1 Pitch trace of the continuer *en* in line 16 of Fragment 4

The production of the continuer *en* is also frequently accompanied by head nods. Some of these head nods are single nods consisting of only one downward and one upward movement. Others consist of a succession of nods. In Fragment 4, four *en*s co-occurr with head nods (lines 6, 16, 20 and 22). Vocal continuers and head nods co-occurring together in this fragment display the listener's access to the speaker's stance as well as signs of attention from the listener to what is being said. Tan maintains a 0.2-second pause and looks straightly at Shu. In lines 16 and 18, upon the completion of his TCU in line 15, Shu produces the continuer *en* which is accompanied by a head nod consisting of a downward and upward movement. During the production of the response token *en*, Shu is seen gazing directly at her interlocutor Tan.

Fragment 5 is another example demonstrating the use of *en* as a continuer. In this fragment, Rui (male) describes how introductions in academic papers are usually written in his field of research, i.e., engineering. Prior to this sequence, Liang (female) gets confused between thesis statements and arguments, not knowing how to distinguish these two terms. Rui points out that writing a thesis statement is very similar to the process of writing an introduction in academic papers. The targeted minimal response token *en* appears in lines 5, 7, 9, 11, 13, 15, 17 and 19.

Fragment 5: cont2 (thesis statement)

01 Rui： 啊 你 首先 当然 -

a ni shouxian dangran

a you first of all certainly

'First of all, of course,'

02 像 我们 领域 的话 -

xiang women lingyu dehua

like we field if

'in our field, for example,'

03 一般 写作 的 风格 应该 是 这 样子.

yiban xiezuo de fengge yinggai shi zhe yangzi

usually writing ASSC style should be this sample

‘the style of writing should be like this.’

04 **就 第一 句话 就 说 -**

jiu diyi juhua jiu shuo

just first sentence just say

‘The first sentence should be,’

Liang head: ↓↑↓↑↓

05→ Liang： **嗯**$_1$;

en$_1$

mm

‘Mm.’

06 Rui： **啊 有 一 个 什么 技术;**

a you yi ge shenme jishu

a have one CL what technique

‘there is one technique,’

Liang head: ↓↑↓↑

07→Liang： **[嗯**$_1$.

[en$_1$

[mm

‘Mm.’

08 Rui： **[或者 是 什么 领域;**

[huozhe shi shenme lingyu

[or be what field

‘or (there is) a field.’

Liang head: ↓↑

09→ Liang： **[嗯**$_1$.

[en$_1$

[mm

‘Mm.’

10 Rui： **[它 很 重要;**

[ta hen zhongyao

[3SG very important

‘It is very important.’

Liang head: ↑↓↑↓↑↓

(0.7)

11 **它 为什么 很 重要 -**

ta weishenme hen zhongyao

3SG why very important

'Why is it very important?'

12 因为 在 各 个 领域 有 -

yinwei zai ge ge lingyu you

because at every CL field have

'Because in every field, it has'

Liang head: ↓↑

13⟶ Liang： [嗯$_1$,

[en$_1$

[mm

'Mm.'

Liang head: ↓↑

14 Rui： [比较 广泛 的 -

[bijiao guangfan de

[relatively broad ASSC

'relatively broad'

Liang head: ↓↑

15 ⟶Liang： [嗯$_{1}$.

[en$_{1}$

[mm

‘Mm.’

16 Rui： [implication;

‘implication.’

Liang head: ↓↑↓↑

17 ⟶Liang： 嗯$_{1}$.

en$_{1}$

mm

‘Mm.’

18 Rui： 但 然后 有 很 多 的 研究 -

dan ranhou you hen duo de yanjiu

but then have very many ASSC studies

‘But, then, there are many studies.’

Liang head: ↑↓

19 ⟶Liang： 嗯$_{1}$,

en$_{1}$

mm

'Mm.'

20 Rui： 但是 还有 什么 问题 没有 解决 -

DANSHI haiyou shenme wenti meiyou jiejue

BUT else what problem NEG resolve

'But what are the problems which were not solved.'

Rui starts the sequence by describing the contents of an introduction in academic papers in his field of research. He uses *shouxian* "first of all" in line 1, which marks the start of a list construction (Schegloff 1982). During Rui's turn in lines 1 to 4, Liang refrains from taking the turn, displaying her understanding that the turn is still not complete. Her first *en* is produced in line 5, in the middle of Rui's TCU (lines 4 and 6). Liang's *en* at this point signals to Rui that Liang attends to what he is saying, and does not pretend to take the floor, therefore encouraging Rui to continue speaking. What is notable here is that Rui shifts his gaze towards Liang at the end of line 4 while producing *jiu shuo*… "that is", and Liang produces *en* immediately after Rui's gaze shift. Gaze has been documented as mobilizing recipient's responses (Stivers and Rossano 2010). Thus, Liang's production of the response *en* immediately after the gaze shift displays her orientation to the speaker's gaze shift as seeking a response.

Figure 2 Rui (left) gazes away during *juhua* (line 4)

Figure 3 Rui (left) gazes at Liang(right) during *jiu shuo* (line 4)

Furthermore, Rui describes what contents the introduction of an academic paper must include (lines 6 to 20). During this activity, Liang produced *en* another six times, all instances serving as continuers (lines 7, 9, 13, 15, 17 and 19). Four of these *en*s occur at TRPs (7, 9, 17 and 19), whereas the other two occur in the middle of Rui's incomplete TCUs (lines 13 and 15). We will take a closer look at the two *en*s appearing in lines 13 and 15. Liang produces these two tokens in the middle of Rui's TCU (lines 12 to 16). At line 12, in the *yinwei*-clause, the Direct Object of *you* "have" is missing. Thus, the syntactic structure is incomplete. At line 14, the ongoing TCU is still not complete grammatically due to its ending with the incomplete NP with modifier+*de*, which is followed by a noun *implication* in line 16. The two uses of the continuer *en* at lines 13 and 15 show Liang's understanding of what is being said by Rui. They also display her understanding that the ongoing TCU is in progress, thus, encouraging the speaker, Rui, to continue speaking.

In Fragment 5, all continuers are also concurrent with head nods. Throughout Rui's turn in lines 14 to 17, Liang produces multiple head nods concurrent with *en*s. Maynard (1987) found that head nods which may or may not co-occur with brief vocalizations such as *uh huh* and *right* function as backchannel continuers.

Next, I will analyze the prosodic features of *en* as a continuer in Fragment 5. During Rui's explanation in lines 12 to 17, Liang produces 3 continuers, two of which (lines 13 and 15) occur in overlap with Rui's subsequent TCUs. It is difficult to conduct an acoustic analysis of sounds in overlap. Thus, I will analyze the prosodic features of the last *en* token in line 17 which is more clearly produced. Figure 4 summarizes the prosodic features of the continuer *en* in Fragment 5 (line 17). As indicated in the figure, the F0 trace illustrates the level pitch movement. The pitch of the syllable starts at 88Hz and stays more or less the same until the end of the syllable.

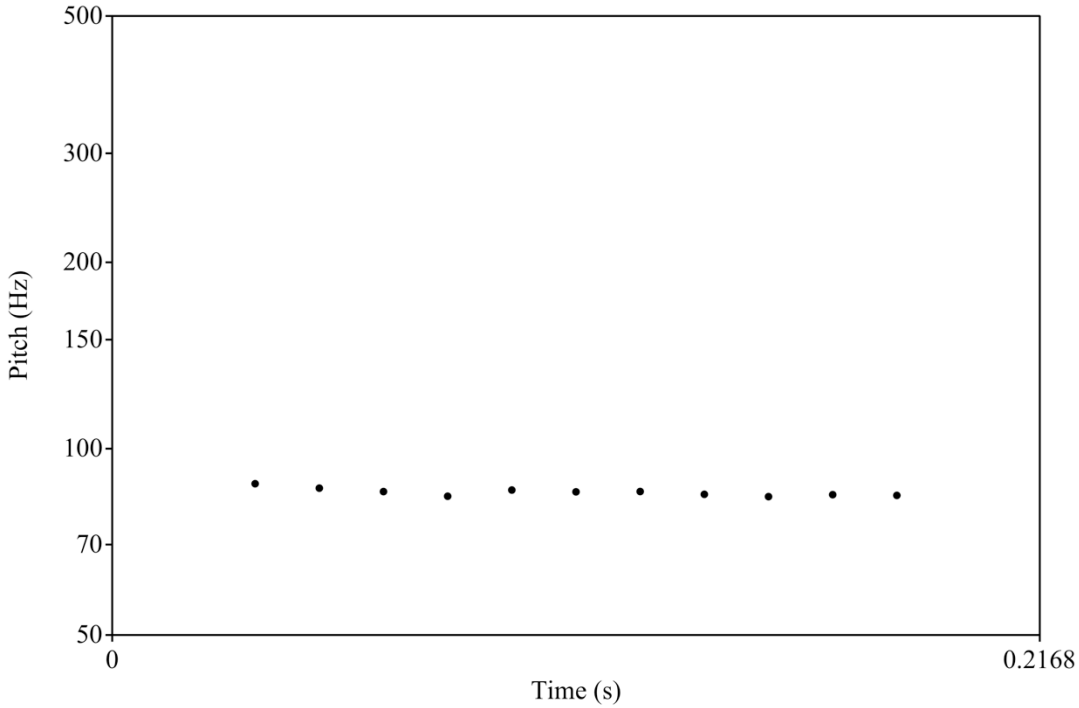

Figure 4 Pitch trace of the continuer token *en* in line 17 of Fragment 5

This section shows the use of *en* as a continuer in Mandarin conversation. The producer of an *en* as a continuer never takes the floor from the speaker. The function of *en* as a continuer is to display the listener's understanding that the speaker's turn is still in progress while showing the speaker that there is "no problem" with the prior speaker's turn. When used as a continuer, *en*

frequently occurs during larger units of speech such as storytelling, explanations or descriptions, where one of the speakers takes the floor for the majority of the discussion. In particular, *en* continuers may appear in the middle or at the end of speaker's TCU.

In a nutshell, this section examined the function of *en* as a continuer, as well as its sequential environment and its prosodic and visual features. The major findings are summarized as follows.

The interactional function of the continuer *en*:

- To display an understanding that the turn is still in progress

The sequential features of continuer *en*:

- In the midst of a larger sequence

The prosodic features of continuer *en*:

- Slightly falling or level pitch movement

The visual features of continuer *en*:

- Often accompanied with a single nod, or with a succession of nods
- Gaze of a speaker directed at the listener

4.2 Acknowledgement

In research on English response tokens, the term "acknowledgement" is usually used for the token *yeah.* The main function of an acknowledgement token is a retrospective receipt, which claims understanding, agreement, or hearing (Gardner 2001). In my data *en* has been shown to be similar to the English token *yeah* as an acknowledgement token. *En* used for acknowledgement in the current data claims that that the prior turn was received, and also displays an understanding that the speaker's turn is complete. Just like with continuers, acknowledgement

tokens can occur during activities such as storytelling, explanations, descriptions, discussions and arguments, etc. *En* used for acknowledgement in my data is often placed at the end of an extended turn. *En* in Mandarin, just like with the English *yeah*, has a higher probability that its producer will continue with further speech and move away from a listener's role (Drummond and Hopper 1993). Further talk following the acknowledging token *en* can be brief or extended.

Fragment 6 shows the typical use of *en* as acknowledgement token and its sequential environment. Fragment 6 is from the same conversation as Fragment 5. In this fragment, two speakers, Liang and Rui, discuss how to get tenure while working as professors in the fields of science and engineering.

Fragment 6: ackn1 (tenure)

01	Rui:	不 同 的 不 一样;
		bu tong de bu yiyang
		NEG similar ASSC NEG same
		'It is different (in every field).'
02		像 我们 那 个 我们 理 那 个 理工科 的话 -
		xiang women na ge women li na ge ligongke dehua
		such as we that CL we science that CL science if
		'As in our field, in engineering, then,'
Liang head:		↓↑↓

03 Liang： 嗯$_1$.

en$_1$

mm

‘Mm.’

04 Rui: 有 发 文章 的 要求;

you fa wenzhang de yaoqiu

have publish paper ASSC request

‘there’s a requirement of publishing papers.’

05 第一 是 paper ;

diyi shi paper

first be paper

‘First, is a paper.’

06 Liang**:** 嗯$_1$.

en$_1$

mm

‘Mm.’

07 (1.3)

08 Rui： 第二 是 你 拿 的 funding;

dier shi ni na de funding

second be you take ASSC funding

'Second, is the funding you got.'

Liang head: ↓↑

09 Liang**:** 嗯$_1$.

en$_1$

mm

'Mm.'

10 (1.4)

11 就 是 你 可以 申请 到 的: -

jiu shi ni keyi shenqing dao de

just say you can apply to ASSC

'The funding you can apply for…'

12 Rui： 你 拿 到 的 钱;

ni na dao de qian

you get arrive ASSC money

'The money you got.'

13 Liang： 哦 哦 [我 拿 到 的 钱.

o o [wo na dao de qian

oh oh [I get arrive ASSC money

'Oh oh, the money that I got.'

14 Rui： [你 -

[ni

[you

'You…'

15 对.

dui

yes

'Yes.'

16 还 有 就 是 你 带 的 学生 -

hai you jiu shi ni dai de xuesheng

still have just be you bear ASSC student

'Also the students you are supervising'

17 Liang: [哦.

[o

[oh

'Oh.'

18 Rui: [有 的 地方 带 的 学生 也 有 要求.

[you de difang dai de xueshng ye you yaoqiu

[have ASSC place bear ASSC student also have request

'In some places there's a requirement of supervising students too.'

19 Liang: [嗯$_1$ -

[en$_1$

[mm

'Mm.'

20 Rui: [还 有 可能 是 要 编书 啊 什么 的.

[HAI you keneng shi yao BIANshu a shenme de

[still have likely be need write book PRT things like that

'Also, it is possible that you need to publish a book, or something.'

Liang head: ↓↑

21 → Liang： 嗯$_1$.

en$_1$

mm

‘Mm.’

22 *(1.3)*

23 Rui： 就 这 [主要 是 看 这 几 个;

jiu zhe [zhuyao shi kan zhe ji ge

just this [mainly be look this several CL

‘Mainly, they look at all these.’

24 Liang： [ＸＸＸ

25 Rui： 对.

dui

yes

‘Yes.’

26 Liang： 嗯$_1$.

en1

mm

'Mm.'

27 Rui： 这 些 条件 你 都 要 满足 了-

zhe xie tiaojian ni dou yao manzu le

this CL condition you all need satisfy CRS

'If you satisfied all these requirements,'

28 你 才 可以 拿 到 那 个 tenure;

ni cai keyi na dao na ge tenure

you still can get arrive that CL tenure

'you can get tenured.'

Liang head: ↓↑↓↑

29 Liang： [嗯1,

[*en1*

[mm

'Mm.'

30 Rui： [要 不然 的话 -

[YAOburan dehua

[otherwise if

'Otherwise.'

31 **你 就 必须 跑.**

ni jiu bixu pao

you just must run

'you have to 'run'.'

Liang head: ↓↑↓↑

32 ⟶ Liang: **嗯**$_1$.

en$_1$

mm

'Mm.'

33 (0.7)

34 **那 拿 到 了 就 是 说 -**

na na dao le jiu shi shuo

then get arrive PFV just be say

'If I get tenured, it means'

35 我 永远 可以 在 这 个 学校 工作 ;

wo yongyuan keyi zai zhe ge xuexiao gongzuo

I always can at this CL school work

"I can always work in this school, right?'

36 Rui: [啊:: -

[a

[PRT

'A…'

37 [好像 也 不 是 吧;

[haoxiang ye bu shi ba

[like also NEG be PRT

'It seems like it is not.'

At the beginning of this sequence, Rui makes a claim that the requirements of getting tenure depend on the research field. Furthermore, from lines 3 to 20, he lists the requirements that are common in his field, such as publishing papers (lines 4 and 5), getting funding (line 8), and supervising students (line 16). Then, he lists the last requirement for achieving tenure which is writing a book (line 20). Rui ends this TCU with the non-exhaustive list ending phrase *shenme de* 'something like that' at line 20 (Liu 2009). The single standing token *en* in line 21 produced by Liang displays her understanding that Rui's extended turn is possibly complete. Her

orientation to Rui's turn as possibly complete also can be evidenced from Liang's body movements. That is, when Liang starts to lean back on her chair upon the production of *en* (Figure 5, 6 and 7).

Figure 5 Body position of Liang (right) during Rui's (left) *haiyou* in line 20

Figure 6 Body position upon Liang's (right) production of acknowledgement *en* (line 21)

Figure 7 Body position after Liang's (right) production of acknowledgement token *en* (in line 21)

The token *en* (line 21) in this fragment acknowledges that Rui's turn was received and understood. Acknowledgement tokens are retrospective in nature as the listener receives and acknowledges the information produced by the speaker in previous turns. In this regard, they are distinctive from continuers, which are prospective in that they signal the speaker to continue his/her speech.

In lines 30 and 31, Rui announces the consequences which occur if the requirements of getting tenure are not fulfilled. Here again, the acknowledgement token *en* in line 32 displays her orientation to Rui's turn as possibly complete. When Liang produced this *en* (line 32), she begins a new yet brief turn (line 34) stealing the floor away from the speaker. In lines 34-35, Liang's turn consists of an understanding check in lines 34-35, and a mitigated disagreement phrase - *haoxiang ye bushi ba* "it seems like it is not" in line 37. In this example, there are two acknowledgement tokens, *en* in line 21 which is a single-standing token that constitutes a turn in itself, and *en* in line 32, which is followed by a brief talk.

Figure 8 shows the pitch movement of *en* as an acknowledgement token in Fragment 6 (line 32). This *en* has a falling pitch movement and is produced in a modal voice. The pitch begins at the highest point of 225Hz while its lowest point towards the end reaches 180Hz. Although it is not evident from the pitch trace in Figure 8 that *en* has a falling pitch movement, repetitive listening by several native speakers concluded that it actually has a gradual falling pitch movement. This token is one of the longest among other acknowledgement *en*s in my data and lasts 0.7 seconds.

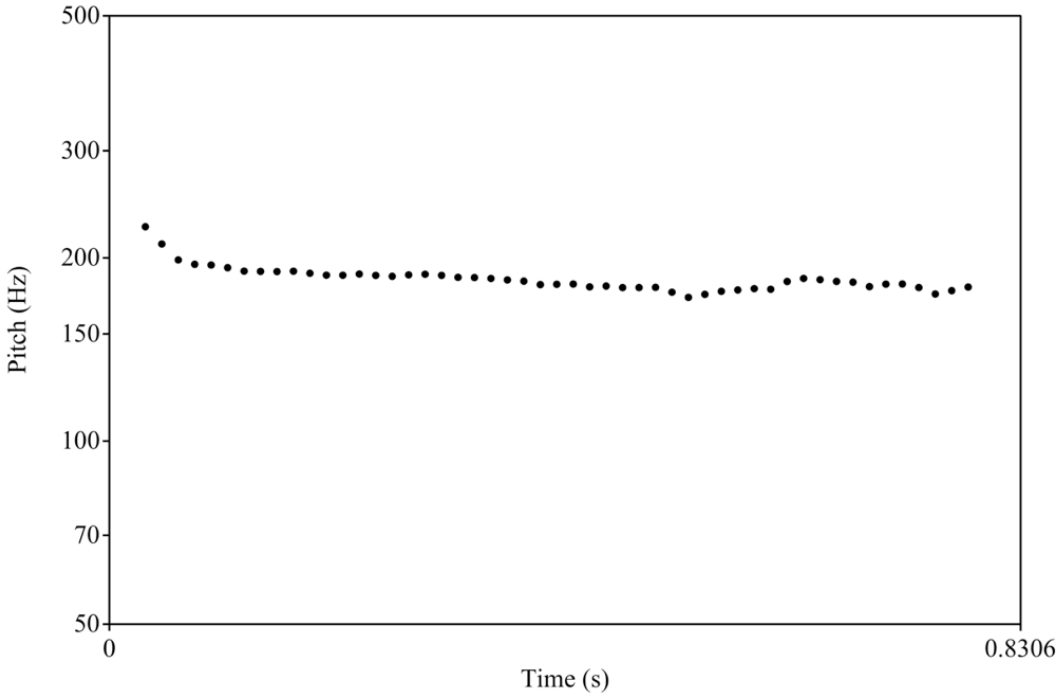

Figure 8 Pitch trace of the acknowledgement token *en* in Fragment 6 line 32

The next fragment shows another two examples of the acknowledgement *en* in a slightly different situational environment. In the above sequence, two participants, Liang and Rui, were involved in the explanation activity, where the roles of a speaker and listener are easily distinguishable.

In Fragment 7 below, two speakers, Lei (female) and Wei (male), are engaged in a discussion on the benefits of reading books and why young generations do not often read them. The conversation does not have an overt primary speaker, because the two participants contribute equally to the development of the topic performing the act of telling in equal time. Prior to the sequence, Lei states that young generations are very impulsive, and do not have enough time to read books.

Fragment 7: ackn2 (reading books)

01 Lei: 我 觉得 -

wo juede

I think

'I think,'

02 你 通过 看书 的话 -

ni tongguo kanshu dehua

you through reading book if

'through reading books,'

03 你 这 个 人 的 贪图 啊 ;

ni zhe ge ren de tantu a

you this CL person ASSC covet PRT

'your greediness,'

04 还 有 你 对 一 些 事情 的 见解 ;

hai you ni dui yi xie shiqing de jianjie

still have you to one CL matter ASSC understanding

'also your understanding of matters,'

05 是 跟 一般 人 是 要 不 一样 的.

shi gen yiban ren shi yao bu yiyang de

be with oridinary person be need NEG same PRT

'is not as same as other's.'

Wei head: ↓↑ ↓↑

06 Wei： 对 对.

dui dui

yes yes

'Yes yes.'

07 (0.7)

08 我 觉得 在 一 个 人生 就 是 -

wo juede zai yi ge rensheng jiu shi

I think in one CL life just be

"I think, in one's life,

09 看书 留下 的 是一 个 积淀 ;

kanshu liuxia de shi yi ge jidian

read book leave ASSC be one CL accumulated wisdom

‘reading books leaves accumulated knowledge to us.’

Lei head: ↓↑↓↑↓↑↓↑↓

10 Lei: **嗯**$_1$ **嗯**$_1$.

en$_1$ en$_1$

mm mm

‘Mm mm.’

11 Wei： **是 一 只** -

shi yi zhi

be one CL

‘It,’

12 **是一 个 对 这 个 人 气质 的 一 个 培养** .

shi yi ge dui zhe ge ren QIzhi de yi ge peiyang

be one CL to this CL person temperament ASSC one CL cultivation

‘it cultivates one’s temperament.’

Lei head: ↓↑↓↑

13 ⟶ Lei： **嗯** $_1$.

en$_1$

mm

‘Mm.’

14 我 觉得 像 伟世佳 可能 -

wo juede xiang WeiShijia keneng

I think like WeiShijia maybe

‘I think, WeiShijia maybe,’

15 现在 因为 就 是 学习 -

xianzai yinwei jiu shi xuexi

now because just be study

‘because s/he studies now.’

16 可能 就 是 在 国外 学习 还是 -

keneng jiu shi zai guowai xuexi haishi

maybe just be at foreign study still

‘Maybe (because s/he) studies in the foreign country,’

17 你 好多 时候 适应 语言 这 个 -

ni haoduo shihou shiying yuyan zhe ge

you many time accommodate language this CL

“(because s/he) adapts to the language’

18 花 的 时间 比较 多 -

hua de shijian bijiao duo

spend ASSC time relatively many

‘and spends a lot of time on that.’

Wei head: ↑↓↑↓↑↓↑↓↑↓

19 Wei: [嗯$_{1}$.

[en$_{1}$

[mm

‘Mm.’

20 Lei: [所以 我 觉得 -

[suoyi wo juede

[therefore I think

‘That’s why, I think,’

21 就 是 把 时间 放 在 阅读 课外书 的 这 个 上头 ;

jiu shi ba shijian fang zai yuedu kewaishu de zhe ge shangtou

just be PRT time put at reading extracur. book ASSC this CL on

'(s/he) spending time on reading extracurricular books,'

23 可能 会 有 一点 少.

keneng hui you yidian shao

possibly may have a little less

'maybe it's a bit less.'

24 Wei： 嗯 我 觉得 现代人 可能 也 比较 功利 吧;

en wo juede xiandairen keneng ye bijiao gongli ba

mm I think modern people possibly also relatively materialistic PRT

'I think people nowadays are materialistic'

25⟶ Lei: 嗯 $_1$.

en$_1$

mm

'Mm.'

26 [而且 -

[erqie

[moreover

'Moreover,'

27 Wei: [我 想 ;

[wo xiang

[I think

'I think,'

28 Lei: [((lip smack))

29 Wei: [可能 看书 就 是 对 我 的 收益 -

[keneng kanshu jiu shi dui wo de shouyi

maybe reading books just be to I POSS gain

'maybe readings books to me……'

Lei starts the sequence by stating her opinion on the benefits of reading books (lines 1 to 5), and Wei affiliates with her by producing the multiple response token *dui dui* "yes yes" (line 6). After a 0.7 pause in line 7, Wei continues the sequence by displaying his own stance on how important books are. He argues that books leave accumulated knowledge for young generations and cultivate temperament in people (lines 8 to 12). Wei's turn comes to an end in line 12. In line 13, Lei produces *en* showing her orientation to Wei's turn in line 12 as possibly complete. Next, she takes the floor and moves on to her explanation of why young people do not read books anymore (lines 14 to 23). One thing to note here is that when Lei produces the acknowledgement token *en* in line 13, we can see a gaze aversion, in which Lei withdraws her gaze from Wei and looks downwards. The arrows in Figures 9 and 10 indicate Lei's gaze direction. She withdraws her

gaze, treating Wei's action of stating his own opinion as possibly complete, while Wei continues to maintain his gaze towards Lei. Gaze withdrawal of one of the speakers at points of possible completion displays an orientation toward the possibility of ending the sequence, while the other party expects its continuation by maintaining gaze on the other (Rossano 2012). In addition, during the delivery of *en*, two slight head nods also were observed. These head nods occur with *en* and signal to the speaker that information is being received and that there is no problem with understanding.

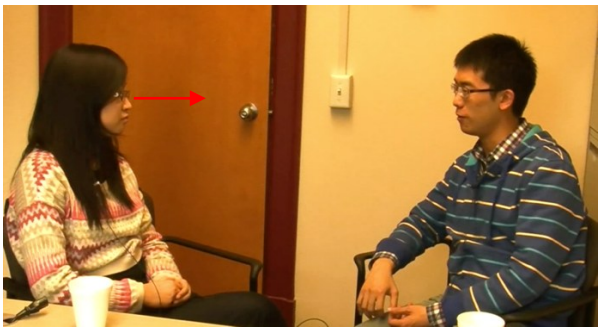

Figure 9 Lei (left) gazes at Wei (right) during *peiyang* at line 12 (Fragment 7)

Figure 10 Lei (left) gazes downwards during *en* at line 13 (fragment 7)

After Lei completes her turn in line 23, Wei continues to suggest another possible reason for why young people do not often read books. After Wei's turn in line 24, Lei produces the second *en* acknowledgement token, showing that she understands Wei's turn as finished. In line 26, Lei wants to start her turn with *erqie* "moreover" to continue elaborating on the current topic. This is further evidence of her understanding that Wei's turn is finished.

Generally, the acknowledgment token *en* displays that the listener treats the speakers' turn as unproblematic, and thus shows that the speaker's turn was received and understood by the listener. *En* also shows understanding that the speaker's turn is possibly complete and often

appears at the end of extended turns. This is similar to the usage of the English token *mm* with falling contour (see Gardner 2001). *En* can be a single standing token, or it can be followed by further speech.

When *en* is used as an acknowledgement token, it is often produced with certain prosodic features, which are illustrated in Figure 11 below. First of all, it can be seen from Figure 11 that the acknowledgement token *en* has a gradual falling pitch movement with its highest point at 184Hz and its lowest point at 155Hz. Second, this token is longer than a continuer *en* with a length of 0.5 seconds. Finally, in terms of its segmental features, this token is produced in a modal voice.

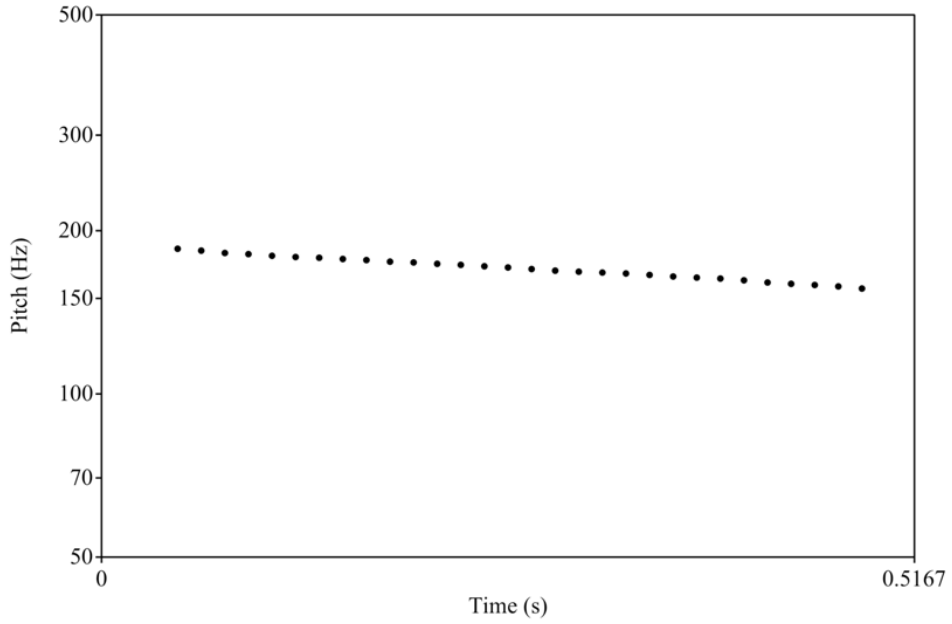

Figure 11 Pitch trace of the acknowledgement token *en* in line 25 of Fragment 7

In summary, in this section I examined the interactional function, sequential features, prosodic features and body movements performed during the delivery of the acknowledgement token *en* produced by listeners. The findings are summarized as follows.

The interactional function of the acknowledgement token *en*:

- Claims understanding that the previous turn to be complete

The sequential features of the acknowledgement token *en* are:

- Utterance initial position
- At the end of the speaker's extended turn

The prosodic features of the acknowledgement token *en* are:

- Overall falling pitch movement
- Longer duration than with continuers (up to 0.7 seconds)

The body movements associated with the production of the acknowledgement token *en* are:

- Gaze aversion
- Leaning backwards
- One or two head nods

4.3 Confirmation

The Mandarin minimal response token *en,* apart from being a continuer and an acknowledgement token, can also function as a confirmation token. *En* can confirm the information in the previous turn. It is different from the previous uses in that it is not usually produced by the listener, but by the speaker. Confirmation *en* tokens in the current data usually follow and confirm an understanding check initiated by a listener. Understanding checks/requests for clarification usually appear during a speaker's extended turn. A typical example of the confirmation token *en*

is shown in Fragment 7.

In Fragment 8, Hu (male) asks Liu (female) if her friend plans to visit Prague with her. Liu said that her friend had already visited the city, and that this friend liked the place so much that he plans to go there again (lines 2 to 6). In line 7, Hu initiates an understanding check by asking if the particular place her friend had visited was Prague. Liu confirms that this place was indeed Prague with *en* (line 8).

Fragment 8: conf1 (Prague):

01 Hu: 他 也 要 去 啊，-

ta ye yao qu a

3SG also need go PRT

'He also wants to go there?'

02 Liu： 他 去 过 了-

ta qu guo le

3SG go EXP PFV

'He already went there'

03 然后 告诉 我 –

ranhou gaosu wo

then tell I

‘And told me’

04 那 边 特别 特别 好玩 –

na bian tebie tebie haowan

that side very very interesting

‘This place is a lot of fun’

05 他 说 他 一定 还 要 再 去 一 次；

ta shuo ta yiding hai yao zai qu yi ci

3SG say 3SG definitely still need again go one time

‘He told me he definitely will go (there) again’

06 就 专门 再 去 一次 [布拉格；

jiu zhuanmen zai qu yi ci [Bulage

just expecially again go one time [Prague

‘Specially go to Prague again’

07 Hu： [说 布拉格 吧 他；

[shuo Bulage ba ta

[say Prague PRT 3SG

‘He is talking about Prague, right?

Liu head: ↓↑

08 ⟶ Liu: **嗯**$_1$ –

en$_1$

mm

'Mm'

09 Hu: **是** –

shi

yes

'Yes'

10 **布拉格 确实 很 出名**.

Bulage queshi hen chuming

Prague indeed very famous

'Prague is very famous'

11 **然后** –

ranhou

then

'Then,'

Prior to the sequence, Liu mentions that one of their mutual friends would like to visit Prague. At this point, Hu starts this sequence by asking if their friend is also going to visit the Czech Republic (line 1). Liu produces an extended response to Hu's question (lines 2 to 6). During her extended turn, Liu becomes a primary speaker in this sequence. Liu informs Hu that the friend *ta* would like to visit the place again (line 5), but the NP representing the place in the VP phrase is omitted: *zai qu* + (NP) + *yici* "Go to + (Prague) + again" (NP being the Goal). It is unclear what city their friend visited until line 5.

In line 7, Hu initiates a repair by using *shuo Bulage ba ta…* "He is talking about Prague, right?". As a response to Hu's repair, Liu confirms the candidate's understanding with *en*. Although Liu states which city their friend went to in line 6, this statement is in overlap with Hu's repair initiation. During her extended turn, Liu gazes downwards (Figure 12), and only when uttering the last words in line 5 does she return her gaze back to Hu. (Figure 13). Speakers tend to look away from listeners during and near the beginning of long utterances and they usually redirect their gaze back to the listener near the end of their utterance in order to seek a response from their interlocutor (Kendon 1967; Bavelas, Coates and Johnson 2002). When Liu gazes back to Hu, Hu withdraws his gaze from Liu (Figure 14) and initiates a repair, displaying his understanding that Liu seeks a response from him. Liu maintains her gaze during the whole production of Hu's repair. When producing *en*, Liu continues gazing at Hu and produces one expansive nod (Figure 15).

Figure 12 Liu (left) gazes downwards during *zai*

Figure 13 Liu (left) gazes back at Hu (right) during *yici*

Figure 14 Liu (left) gazes at Hu (right) during repair

Figure 15 Liu (left) gazes at Hu (right) during *en*

It should be noted here that the turn prior to a confirmation token *en* tends to be an action, which makes confirmation conditionally relevant. In other words, the previous turn to which a confirmation token responds tends to be a question or a statement that does not need an elaborated response. Mandarin has a variety of devices which can be used to seek a confirmation, e.g. tag questions, the statement suffixed with final particles. In this example, Hu produces a question using the sentence final particle *ba*. In Mandarin, the statement to which *ba* is attached elicits the approval or agreement of the listener (Li and Thompson 1981).

The token is produced with a falling pitch movement, with the contour reaching its peak at 213Hz and falling down to 173Hz (Figure 16). It is notable that the pitch register of confirmation tokens is higher than with continuers and acknowledgement tokens. The pitch range of the token is also relatively broader in comparison with continuers. The confirmation token *en* in this excerpt has a duration of 0.3 seconds (Figure 16).

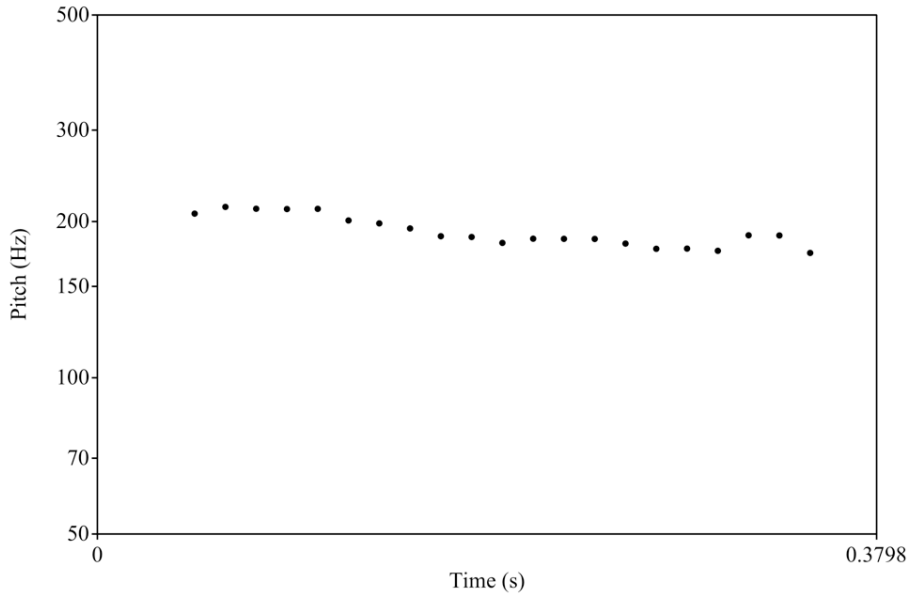

Figure 16 Pitch trace of the confirmation token *en* at line 8 of Fragment 8

The following is another instance of the confirmation token *en*. In the sequence below, three female participants, Ru, Anni, and Bai, are discussing the discount Ru receives in the store where she works. Prior to this interaction, Ru had just informed Bai that she works at a clothing store in Edmonton.

Fragment 9: conf2 (discount)

01 Bai： 从 那 之后 你 开始 打工 啊;

cong na zhihou ni kaishi dagong a

from that after you start work PRT

'You started to work on that day?'

02 Ru： 啊.

A

PRT

'Yes.'

03 Bai: 那 你 买 东西 都 有 折扣 的 对 吧 –

na ni mai dongxi dou you zhekou de dui ba

then you buy things all have discount NOM yes PRT

'You have a discount to buy clothes, right?'

Ru head: ↓↑

04 → Ru: 嗯$_1$.

en$_1$

mm

'Mm.'

05 就 七 折;

jiu qi zhe

just seven discount

‘30% discount.’

06 买 那 个 ERF 是 九 折 .

mai na ge ERF shi jiu zhe

buy that CL ERF be nine discount

‘10% discount on ERF.’

In line 1, Bai asks Ru if she receives any discounts when buying clothes at the place where she works. This question consists of a statement and a tag question *dui ba* “is it right?”. This tag question serves to seek confirmation of the statement that occurs before the tag (Li and Thompson 1981:546). Thus, this tag question implies that Bai believes that Ru has a discount, and seeks confirmation of this discount from Ru. Ru uses *en* to respond to Bai’s question in line 3. Then, in line 6, Ru informs Bai that her discount differs from store to store. In this fragment, we found that confirmation tokens follow confirmation requests (line 3). During the production of these actions, Bai maintains her gaze towards Ru, whereas Ru gazes at Bai only during the production of the confirmation token (line 4). As mentioned in Rossano (2009), in requests for repair or confirmation, it is speakers rather than listeners who are more likely to be gazing. However, because participants are involved in other activities such as eating, this could possibly explain why Ru looked away during those actions. Bai’s gaze at Ru during her confirmation requests functions as a way to mobilize a response from Ru (Stivers and Rossano 2010).

In this section, I introduced another usage of *en* in Mandarin conversation. It can serve to confirm a recipient's understanding check through repair initiation (e.g. at line 8 of Fragment 8). It also can be a confirmation of the direct request for confirmation (e.g. at line 4 of Fragment 9). In terms of its sequential position, *en* as a confirmation token tend to occur at the responsive position after the listener's repair initiation or request for confirmation. One difference between confirmation and acknowledgement *en* tokens is that confirmation tokens are usually produced by speakers, while acknowledgement tokens are usually produced by listeners. Another crucial difference is the epistemic status/authority of the participants. The producer of the confirmation token *en* has an epistemic authority over a referent event, whereas the producer of an acknowledgement *en* does not have such authority. Epistemic status may be the most important factor that distinguishes these two uses of the *en* token: confirmation and acknowledgement. Another factor, in which these two tokens could be differentiated, is their positions in the sequence. The acknowledgement token *en* occurs at the end of a larger sequence, whereas confirmation tokens following a repair initiation and request for confirmation usually occur in the middle of a larger sequence.

Figure 17 shows the prosodic features of the confirmation token which appeared in line 4 of Fragment 9. *En* in this instance has a falling pitch movement and a broader pitch range than with continuers and acknowledgements. Its peak reaches a height of 257HZ and remains level until the middle of the syllable, when it sharply falls to its lowest point of 175HZ (Figure 20). *En* produced by Ru in Figure 17 has a short duration of 0.15 seconds. In a nutshell, the prosodic features of a confirmation token *en* can be summarized as follows: modal voice, falling pitch movement, broader pitch range, and short duration.

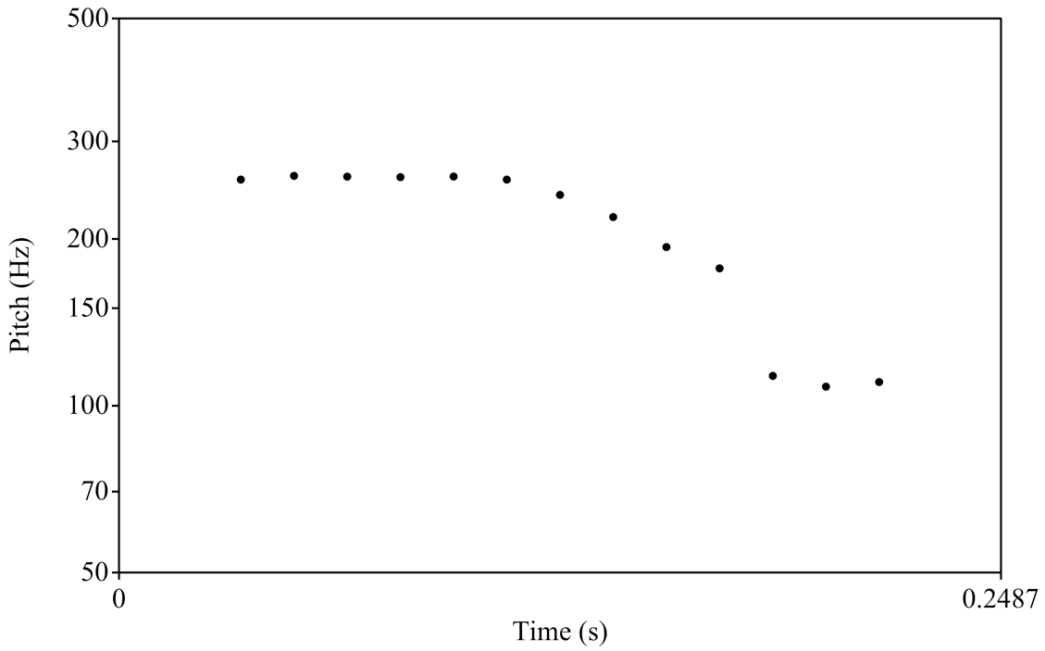

Figure 17 Pitch trace of the confirmation token *en* in line 4 of Fragment 9

This section analyzed the third interactional function of *en* and the sequential environment in which it appears. The analysis of prosodic features and visual cues relating to the production of the confirmative *en* were also introduced. The main findings are summarized below.

The interactional function of the confirmation token *en*:

- To confirm a previous turn's information validity or adequacy

The sequential features of the confirmation token *en*:

- Turn initial position
- Responsive position after a repair/request for confirmation

The prosodic features of the confirmation token *en*:

- Modal voice
- Falling pitch contour

- Relatively short duration

The visual features of the confirmation token *en*:

- Gaze during the production of *en* is usually directed at the listener
- Accompanied by a single nod or by a succession of nods

4.4 Register of receipt

In this section, we will examine the function of *en* to register the receipt of the listener's response, and its sequential, prosodic and visual features. In the interaction below, three girls, Bai, Anni and Ru, discuss if Bai's husband will find a job in Canada. Prior to the sequence, Bai (female) talks about her husband who plans to come to Edmonton. Bai wants her husband to be employed at the company where she currently works. She worries that he will not be hired because he is over-qualified for the position that the company offers. However, she believes her husband's PhD will help him find a job, and that this degree also could help him climb the career and promotion ladder later on.

Fragment 10: reg1 (husband)

01	Bai：	因为	我	觉得	有	博士	的	这	个	过程 ;
		yinwei	*wo*	*juede*	*you*	*boshi*	*de*	*zhe*	*ge*	*guocheng*
		because I	think	have	PhD	ASSC	this	CL	process	
		'Because I think having PhD,'								
02		他	思维	能力	还	是	不	一样	的	嘛;

ta siwei nengli hai shi bu yiyang de ma

3SG thinking ability still be NEG similar PRT PRT

'his thinking ability is different.'

Anni head: ↓↑↓↑↓

03 Anni： [对 对 对.

[*dui dui dui*

[yes yes yes

'Yes yes yes.'

04 Bai： [这 样 再 往上 走 可能 会 -

zhe yang zai wangshang zou keneng hui

this case again forward go possible can

'In terms of promotion,'

05 Anni: 确实.

queshi

indeed

'Indeed.'

06 Bai： 帮助 他 更 多.

bangzhu ta geng duo

help 3SG more many

'this will help him a lot.'

07 Anni： 嗯$_1$.

en$_1$

mm

'Mm.'

Bai head: ↓↑↓↑↓↑

08 → Bai： 嗯$_1$.

en$_1$

mm

'Mm.'

09 Ru: 我 说 博士 那 个 非 人类 .

wo shuo boshi na ge fei renlei

I say PhD that CL not humanity

'PhD graduates are non-humans…'

In the above sequence, Bai provides us with the knowledge that her husband will be employed in the company where she works. During Bai's speakership, Anni produces several multiple

agreement tokens (*dui dui dui* 'right, right, right') at line 3 and one brief agreement *queshi* 'indeed' at line 5 to display her affiliation with the speaker. The closing of the sequence begins with the last utterance of Bai's extended turn in line 6. Anni displays her understanding that Bai's accounts have possibly come to an end, and produces the acknowledgement token *en* in response to Bai's turn. In line 8, Bai also produces *en*. This *en* is distinct from the token *en* in line 7 in that it registers the recipiency of Anni's response, that is, that it can be found at the third position in the sequence. In other words, Bai, by producing this token registers that she received the listener's response. After *en* at line 9, Ru starts a new sequence, with the token *en* in third position concluding the sequence.

Upon the production of the register of receipt token *en*, Bai gazes down at her plate (Figure 18). The speaker's gaze withdrawal at the point of possible sequence completion displays an orientation toward the possibility of ending the sequence (Rossano 2012). She also begins to produce a series of nods with three downward and two upward head movements. These nods are not expansive in terms of their amplitude and total duration. Such nods in Whitehead (2011) are referred to as "acknowledgement nods". In this sequential position they register the receipt of listener's answer without treating it as news (Whitehead 2011).

Figure 18 Bai (middle) gazes downwards during the production of register of receipt *en* in line 8 in Fragment 10

Figure 19 illustrates the prosodic features of the token *en* at line 8 of Fragment 10. *En* is produced in a modal voice and has a gradually falling pitch contour, and a broad pitch range with the peak of the contour starting at 203Hz and falling down continuously to the point of 153 Hz. This token had one of the longest durations among other register of receipt tokens in my data, at exactly 0.4 seconds (Figure 19).

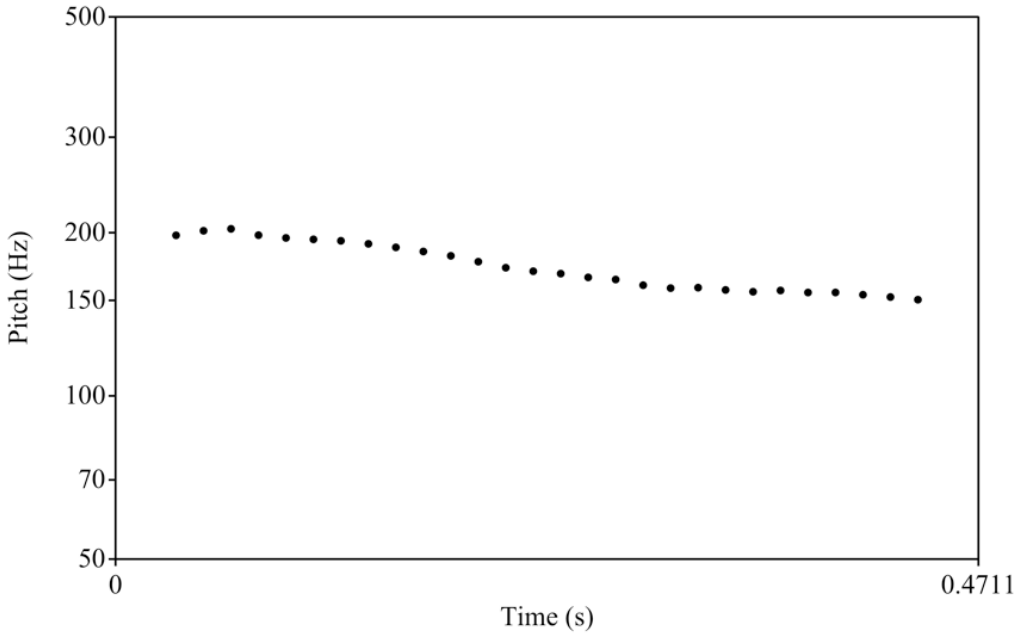

Figure 19 Pitch trace of register of receipt token *en* at line 8 in Fragment 10

Below is another example of *en* used as a register of receipt token. In the following exchange, Lei (female) and Rui (male) discuss the benefits of having extra-curricular activities during a study at a university (here, a choir). The register of receipt token *en* in line 11 in this excerpt is another example which portrays its interactional function similarly to the previous one in Fragment 10.

Fragment 11: reg2 (friends)

01 Lei: 那 [觉得 反正 –

na *[juede* *fanzheng*

in that case [think anyways

'Anyway, I think,'

02 Rui: [mhm]

'Mh mm.'

03 Lei： 不 同 [专业 背景 的 人 在 一起 ；

bu *tong* *[zhuanye* *beijing* *de* *ren* *zai* *yiqi*

NEG same [profession background ASSC person at together

'together with people of different professional backgrounds'

04 Rui： [认识 –

[renshi

[know

'Know'

05 Lei： [交流 挺 不 错 的.

[jiaoliu *ting* *bu* *cuo* *de*

[communicate quite NEG bad PRT

'communicating is very good.'

Rui head: ↓↑

06 Rui： [对 对 对.

[dui dui dui

[yes yes yes

'Yes yes yes.'

07 是 认识 的 人 都 不 一样.

shi renshi de ren dou bu yiyang

be know ASSC person all NEG similar

'You get to know different people.'

Lei head: ↓↑↓↑↓↑

08 Lei： [嗯 $_1$.

[en$_1$

[mm

'Mm.'

09 Rui： [然后 交流圈 也 都 扩大 了 很 多.

[ranhou jiaoliuquan ye dou kuoda le hen duo

[then social circle also all broaden CRS very much

'And social circle also gets expanded.'

Lei head: ↓↑↓↑↓↑

10 Lei： 嗯$_{1}$.

en$_{1}$

mm

'Mm.'

Rui head: ↓↑

11 ⟶ Rui： 嗯$_{1}$.

en$_{1}$

mm

'Mm'

12 (0.5)

13 Lei: [那 -

[na

[then

‘Then,’

14 Rui： [我 觉得 是 一 个 很 好 的 契机；

[wo juede shi yi ge hen hao de qiji

[I think be one CL very good ASSC opportunity

'I think, this is a very good opportunity.'

15 Lei: 嗯$_1$.

en$_1$

mm

'Mm.'

Lei with the discourse marker *na* "in that case" initiates the sequence and provides her reasons for why activities outside the university environment are beneficial. *Na* in Mandarin Chinese is often used for topic shifting or introducing of a new aspect of the topic, as well as for starting a new turn (Liu 2009). The listener of Lei's turn, (lines 1 to 5), Rui, aligns with the speaker's stance by deploying multiple sayings of the agreement token *dui dui dui* as his response (Yang 2013). Then Rui, in lines 7 to 10, provides his own accounts for why extra-curricular activities are important. The beginning of the closure of this sequence starts at line 9, with Rui finishing his turn with the last TCU. Lei, by producing the acknowledgement token *en* at line 10 shows her understanding that Rui's turn is complete. Once again, we can see that there is another *en* token following the acknowledgement token in line 10. As in the previous case, this *en* appearing in this sequential position, that is, in third position, registers the adequacy of the listener's receipt of the prior message. Note that third position response token proposes a sequence closure as observed by Schegloff (2007). After Rui's production of the third position *en* in line 11, Lei

displays her understanding of his extended turn to be finished as she attempts to start a new sequence in line 13, again with the use of the discourse marker *na*.

Figures 20 demonstrates the pitch movement of *en* at line 11 in Fragment 11. This *en* is produced in a modal voice with a level pitch movement. The pitch contour stays more or less same at 182Hz. This token is one of the shortest among other register of receipt tokens. It has a duration of only 0.15 seconds.

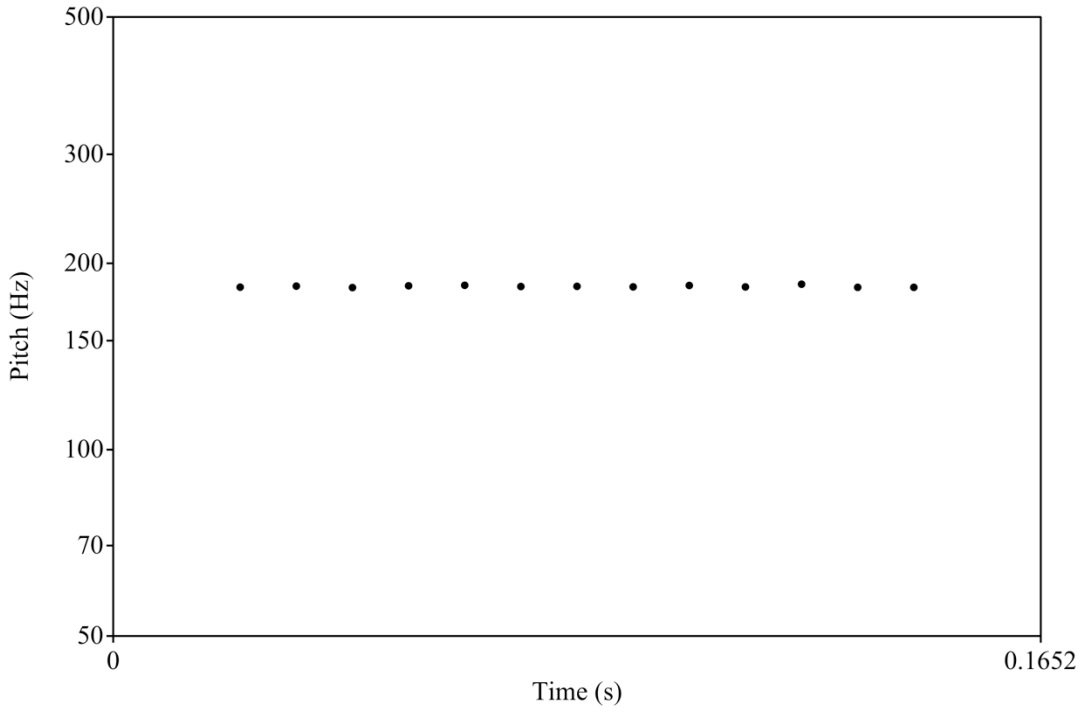

Figure 20 Pitch trace of register of receipt token *en* in Fragment 11

In this section, we analyzed the interactional function of the register of receipt token *en* in Mandarin conversation and the sequential, prosodic and visual aspects of its production. The findings are summarized below.

The interactional function of the register of receipt token *en* is:

- To register the receipt of the listener's response

The sequential features of the register of receipt token *en* are:

- Utterance initial position
- Sequence-closing third position

The prosodic features of the register of receipt token *en* are:

- Falling or level pitch movement
- Duration variance from 0.15 seconds to 0.4 seconds
- Use of the modal voice

The head movements involved during the production of the register of receipt token *en* are:

- One or several head nods in succession.
- Gaze withdrawal

4.5 Summary

This chapter reports on the interactional functions of the minimal response token *en* in Mandarin conversation. The analysis of the sequential environment, the prosodic and body-visual features of four types of *en* were also presented. The analysis showed that *en* can have 4 functions in Mandarin conversation. First, *en* produced by a listener can serve as a continuer showing his/her understanding that the speaker's turn is still in progress. Second, *en* produced by listeners can be an acknowledgement token, which displays their understanding that the speaker's turn is complete. Third, *en* can function as a confirmation token, which confirms that the previous turn's information is valid and correct. Finally, *en* can register the reception of listener's response in the sequence closing third position.

Chapter 5 Conclusion

This chapter summarizes the findings of this study (Section 5.1) and discusses the significance of this research on the minimal response token *en,* as well as its implications to second language teaching and learning (Section 5.2).

5.1 Findings

The current study investigates the interactional functions of the minimal response token *en* in everyday spontaneous Mandarin conversation. Adopting the methodology of Conversational Analysis, Interactional Linguistics, and Multimodal Analysis this research analyzed the sequential, prosodic and bodily-visual features produced in conjunction with the token. Based on its sequential position in the conversation, four interactional functions of the response token *en* were detected.

The continuer *en*. The continuer *en* is frequently produced by a listener during extended turns by another interlocutor. Extended turns include storytelling, extended explanations or descriptions, where one of the interlocutors takes the floor most of the time. Sequential analysis showed that the continuer *en* appears in the middle of the primary speaker's extended talk. Its positioning at the level of the turn may occur in the middle of the TCU or at the end of TCU. The utterer of the continuer *en* never takes the floor of the primary speaker. The main function of the token is to display understanding that the speaker's turn is still in progress. The prosodic analysis showed that the continuer *en* can have slightly falling or level pitch contour, and usually pronounced in a modal voice. Also, *en* has a relatively short duration, and can vary in loudness. Body behavior which accompanies continuer *en* includes head nods (single or a succession of nods) and gazes which are typically directed at the speaker.

Acknowledgement *en*. This token occurs in extended turns during another interlocutor's speakership, or during 'turn-by-turn' talk, where all participants contribute equally to the development of the topic. The acknowledgement token *en* only orients to the turn that has been possibly completed. In terms of its position in the sequence, the acknowledgement *en* often appear at the end of an extended turn, particularly at points of possible turn transition. It can be a single stand-alone token, or it can be followed with further talk by its producer. The prosodic analysis of the token showed that *en* is produced with the modal voice and have a falling pitch contour with a larger pitch range than in continuers. It has a longer duration, with an average of 0.5 seconds. Body behaviors associated with the production of the acknowledgement token *en* include leaning backwards upon uttering the token, and withdrawing the gaze from the speaker. One or two head nods also often occur.

Confirmation *en*. This token is used to confirm that the previous turn's information is valid and adequate. It also serves as a confirmative response to an other-initiated-repair, or to a direct request for information. In terms of its sequential position, it usually appears at a responsive position in the middle of a sequence. This token is produced in a modal voice, and has a falling intonation contour as well as the same duration as a continuer, with the average 0.2 seconds. The confirmation token *en* is always accompanied by a single nod or by a succession of nods. The speaker's gaze during the production of *en* is usually directed at the listener.

Register of receipt *en*. This token is used to register the reception of the listener's responses. Sequentially, it usually appears in the sequence closing third position, wherein the speaker proposes an imminent sequence closure. The prosodic features of register of receipt token *en* can be characterized by its average duration of 0.3 seconds, and a falling or level pitch contour. All

tokens are pronounced in a modal voice. The register of receipt *en* often is accompanied by a single or several head nods.

It has been mentioned in Chapter 1 that minimal response token *en* has two phonetic variations, bilabial nasal [m] and mid central nasalized vowel [ə̃]. An examination of the data shows that there is no one-to-one correspondence between the two phonetic variations and their interactional functions. Both varieties are used in performing different interactional functions.

5.2 Significance and implications of the study

This study is significant due to three aspects. To begin with, it is the first study that focuses on interactional functions of the Mandarin minimal response token *en*. Previous studies either lumped it with other tokens as one homogeneous group (Clancy et al. 1996; Yin 2010), or mainly considered it as discourse marker (Zheng 2007; Gao 2007); neither of the studies build their analysis based on its sequential position. Secondly, it is one of the few studies that investigates response tokens from a multimodal perspective, in terms of including prosodic and bodily-visual practices into the analysis (see another of Yang 2013). Finally, by giving a systematic description of the interactional functions of using the Mandarin *en* in everyday spontaneous conversation, this article also contributes to the literature of cross-linguistic conversational studies. Previous studies on response tokens of different languages show that "interactional goals may be achieved by very different means by members of different groups" (Stubbe 1998:263). This study explicates the typical usages of one of the most frequently used response tokens in Mandarin conversation and sheds light on how the minimal tokens of other languages are employed differently in the sequence of speech.

The results of this study also have implications in second language pedagogy**.** Gardner (1998:205) once mentioned, “as minimal responses are perhaps archetypical exponents of receipt-response, and are so pervasive, it is worth asking why they are not widely taught.” Being so massively used during interaction, it is obvious that response tokens should be taught to learners of foreign languages. However, despite the need to teach response tokens to second language learners, in most cases they are neglected and not included in the teaching curriculum. The general reason for this can be that minimal responses are being frequently "lumped together as a more or less homogenous group" (Sinclair and Coulthard 1975; Coates 1986). Nevertheless, in fact, each particular response token has its place of distinction among others (Beach 1993, Drummond and Hopper 1993, Gardner 1995). There is no systematic knowledge of these conversational items. As Drummond and Hopper (1993) once mentioned, they are so varied and vague that it is sometimes hard to describe them and present them pedagogically.

Perhaps, some teachers will question if it is necessary to teach response tokens to second language learners. There is little evidence showing their importance. However, according to Gardner’s (1998) survey among L2 teachers in Australia, a majority of them expressed a need to teach response tokens to students. In addition, what was notable is that all teachers showed a lack of awareness of the distinctive functions of these feedback tokens. This study hopes to contribute to the knowledge of the interactional functions of the Mandarin *en* in everyday spontaneous conversation and, thus, explain its distinctive functions from other response tokens.

Bibliography

Aoki, H. (2008). *Hearership as Interactive Practice: A multi-modal analysis of the response token Nn and head nods in Japanese casual conversation*. PhD Dissertation.

Atkinson, J. M. (1992). Displaying neutrality: Formal aspects of informal court proceedings. In Drew and Heritage (eds.) *Talk at work: Interaction in institution settings*: 199-211.

Barth-Weingarten, D. (2011). Double sayings of German *JA*—More observations on their phonetic form and alignment function. *Research on Language and Social Interaction* 44(2): 157–185.

Bavelas, J. B., Coates, L., and Johnson, T. (2002). Listener responses as a collaborative process: The role of gaze. *Journal of Communication* 52 (3): 566–580.

Beach, W. (1993). Transitional regularities for 'casual' "Okay" usages. *Journal of Pragmatics* 19: 325–352.

Boersma, Paul & Weenink, David (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.18, retrieved 23 May 2016 from http://www.praat.org/.

Brugman, H., Russel, A. (2004). Annotating Multimedia/ Multi-modal resources with ELAN. In *Proceedings of LREC 2004, Fourth International Conference on Language Resources and Evaluation.*

Bublitz, W. (1988). *Supportive Fellow-Speakers and Cooperative Conversations*. Amsterdam and Philadelphia: Benjamins.

Clancy, P. M., Sandra A., Thompson, Suzuki, R. and Tao, H. (1996). The conversational use of reactive tokens in Japanese, Mandarin and English. *Journal of Pragmatics* 26.1:

355-387.

Coates, J. (1986). *Women, Men and Language: A Sociolinguistic Account of Sex Differences in Language*. London: Longman.

Couper-Kuhlen, E. and Selting, M. (2001). Introducing interactional linguistics. Selting & Couper-Kuhlen 2001b. 1-24. *Amsterdam & Philadelphia: John Benjamins.*

Couper-Kuhlen, E. and Barth-Weingarten, D. (2011) A system for transcribing talk-in interaction: GAT 2', *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* 12:1-51 [http://www.gespraechsforschung-ozs.de].

Deng, X. (2008). The use of listener responses in mandarin Chinese and Australian English conversations. *Pragmatics* 18(2).303-328.

Drummond, K. and Hopper, R. (1993). Back channels revised: Acknowledgement tokens and speakership recipiency. *Research on Language and Social Interaction*. 26 (2): 157-177.

Drew, P. (2005). Conversation analysis. In Handbook of Language and Social Interaction. (K. Fitch, R. Sanders, eds.). Lawrence Erlbaum: 71-102.

Duncan, S. (1974). On the structure of speaker–auditor interaction during speaking turns. *Language in Society* 3:161-180.

Fishman, P. (1983). Interaction: The work women do. In *Language, Gender and Society*. Thorne, B., Kramarae C. and Henley, N. (eds), 89–101. Rowley: Newbury House.

Ford, C. E. and Couper-Kuhlen, E. (2004). Conversation and phonetics. Essential connections. In *Sound patterns in interaction*, Elizabeth Couper-Kuhlen and Cecilia E. Ford (eds.), 3–25. Amsterdam: Benjamins.

Gao, L. (2007). Cong gongnengyuyanxue jiaodu tanxi “en” de shengdiao moshi yu qi biaoyi

gongneng zhi guanglian. (From pragmatics point of view explore the relation between *en*'s tone and its semantic functions). *Anhui literature*, 2007:10.

Gardner, R. (1995). *On some uses of the conversational token 'Mm'*. Unpublished PhD thesis. University of Melbourne.

Gardner, R. (1997). The conversation object mm: A weak and variable acknowledging token. *Research on Language and Social Interaction* 30 (2): 131–156.

Gardner, R. (1998). Between listening and speaking: The vocalisation of understandings. *Applied Linguistics* 19 (2): 204–224.

Gardner, R. (2001). When Listeners Talk: response tokens and recipient stance with special reference to Mm. *Pragmatics and Beyond Series*. Amsterdam: J. Benjamins Publishing.

Gardner, R. (2002). *When listeners Talk*. Amsterdam and Philadelphia: John Benjamins.

Goffman, E. (1981). Footing. In *Forms of Talk*. E. Goffman, ed. Pp. 124-159. Philadelphia: University of Pennsylvania Press.

Golato, A., Fagyal, Z. (2006). Two contours, two meanings: the intonation of jaja in German phone conversation. *Speech Prosody*. May 2-5.

Golato, A. & Fagyal Z. (2008). Comparing Single and Double Sayings of the German Response Token ja and the Role of Prosody: A Conversation Analytic Perspective. *Research on Language and Social Interaction* 41:3, 241-270.

Goodwin, C. (1979). The Interactive Construction of a Sentence in Natural Conversation. In G. Psathas (ed.) *Everyday Language: Studies in Ethnomethodology*. New York: Irvington Publishers, pp. 97-121.

Goodwin, C. (1981). *Conversational Organization: Interaction between Speakers and Hearers*.

New York: Academic Press.

Goodwin, C. (1986). Between and Within: Alternative Treatments of Continuers and Assessments. *Human Studies* 9:205-17.

Goodwin, C. (2006). Interactive Footing. In Holt, E. and Clift, R. (ed.) *Reporting Talk: Reported Speech in Interaction*. Cambridge: Cambridge University Press.

Gorish. J., Wells, B. and Brown, G. J. (2010). The response token as a copy of the prior turn: prosodic and visual properties of 'uhu'. *Phonetics of Talk-in-Interaction.* New Frontiers: 73.

Han, L. (2007). Patterns of Backchannel Responses in Canadian-Chinese Conversations. *International Communication Association.* Annual Meeting.

Heath, C. C. (1992). Gesture's discrete tasks: multiple relevancies in visual conduct and in the contextualization of language. In *The Contextualization of Language*. Auer & di Luzuo (eds.), pp. 101-127.

Heritage, J. (1984). A "change of state" token and aspects of its sequential placement. In J. M. Atkinson & J. Heritage (eds.), *Structure of social action*: 299-345. Cambridge, England: Cambridge University Press.

Houtkoop H. & H. Mazeland. (1985). Turns and Discourse Units in everyday Conversation. *Journal of Pragmatics* 9:595-619.

Iwasaki, S. (1990). Cohesive force in Japanese conversation. *Paper presented to the Annual Meeting of the Association for Asian Studies*.

Jefferson, G. (1984). Notes on a systematic deployment of the acknowledgement tokens "Yeah" and "Mm Hm". *Linguistics* 17.2: 197-216.

Kendon, A. (1967). Some functions of gaze direction in social interaction. *Acta psychologica* 26:

22-63.

Knight, Dawn, Evans, David, Carter, Ronald, and Adolphs, Svenja. 2009. "HeadTalk, HeadTalk and the corpus: towards a framework for multi-modal, multi-media corpus development." *Corpora* 4:1–3.

Li, C. and Thompson, S. (1989). Mandarin Chinese: A Functional Grammar. University of California Press.

Liu, B. (2009). Chinese discourse markers in oral speech of Mainland Mandarin speakers. In Xiao, Smitihield and Rhodel Island (eds.) *Proceedings of the 21st North American Conference on Chinese Linguistics (NACCL-21)* 2: 358-374. Bryant University.

Maynard 1987. Interactional functions of a non-verbal sign: Head movement in Japanese dyadic casual conversation. *Journal of Pragmatics* 11(5): 589-606.

Miracle, W.C. (1989). *Hao*: A Chinese Discourse Marker', in B. Music et al.(eds). *Papers from the 25th Regional Meeting of the Chicago Linguistic Society, Part 2: Parasession on Language in Context*, pp. 213–27. Chicago, IL. University of Chicago. Linguistics department.

Pomerantz, A. and Fehr, B. J. (1997). Conversation analysis: An approach to the study of social action as sense making practices. In van Dijk (ed.), 64–91.

Rossano, F., Brown, P., and Levinson, S.C. (2009). Gaze, questioning and culture. In *Conversation Analysis: Comparative Perspectives*, Jack Sidnell (ed.), 187–249. Cambridge: Cambridge University Press.

Rossano, F. (2012). Gaze in conversation. In *The handbook of conversation analysis*. Wiley Blackwell Publishing. pp.308-329.

Sacks, H. (1969). *Unpublished lecture notes*. Winter 1969, lecture 9.

Sacks, H., Schegloff, E. and Jefferson, G (1974). A simplest systematics for the organization of turn-taking for conversation. Language 50: 696-735.

Sacks, H. (1992). (edited by Gail Jefferson). *Lectures in Conversation*. Vol. 1 Oxford: Blackwell.

Schegloff. E. A. (1972). Notes on Conversational Practice: Formulating Place. In D. N. Sudnow (ed.) *Studies in Social Interaction*. New York: Free Press, 75 – 119.

Schegloff, E. A. (1982). Discourse as an interactional achievement: some uses of 'uh huh' and other things that come between sentences. In Deborah (eds.) *Analyzing discourse: text and talk*: 71-93.

Schegloff, E. A. (1992). Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology* 97: 1295-1345.

Schegloff, 2007. *Sequence organization in Interaction: A primer in Conversation Analysis I*. Cambridge University Press.

Selting et al. (2009). Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion 10, 2009: 353-402 (www.gespraechsforschung-ozs.de) (mit vielen anderen Kolleginnen und Kollegen).

Shao, J. and Zhu, X. (2005). The Discourse Functions of Hao and its Evolution Toward Functional Usage. *Zhongguo Yuwen* [Chinese Language] 5:399–407.

Siitonen, P. & Wahlberg, K. (2015). Finnish particles *mm, jaa* and *joo* as responses to a proposal in negotiation activity. *Journal of Pragmatics* 75: 73-88.

Sinclair, J. and Coulthard, M. (1975). *Towards an Analysis of Discourse: The English Used by*

Teachers and Pupils. London: Oxford University Press.

Sorjonen, M. (2001). *Responding in conversation. A study of response particles in Finnish*. Amsterdam: John Benjamins.

Stivers T & Sidnell, J. (2005). Introduction: Multimodal interaction. *Semiotica*. 156.1/4: 1-20.

Stivers, T. and Rossano, F. (2010). Mobilizing response. *Research on Language and Social Interaction* 43:1, 3 – 31.

Streeck, J., Goodwin, C., & LeBaron, C. (2012). Embodied interaction in the material world: An introduction. In J. *Streeck*, *C. Goodwin, & C. LeBaron* (Eds.).

Stubbe, M. (1998). Are you listening? Cultural influences on the use of supportive verbal feedback in conversation. *Journal of Pragmatics* 29: 257–289.

Tao, H. Y. and Thompson, S. A. (1991). English backchannels in Mandarin conversation: A case study of superstratum pragmatic 'interference'. *Journal of Pragmatics*. 16: 209–223.

Ten Have, P. (1999). *Doing Conversational Analysis*. London: Sage Publications.

Tsai, H. C. (2001). *The discourse function of the dui receipt in Mandarin conversation*. Unpublished MA Thesis.

Wang, Y. F., Tsai P. H., Goodman, D. and Lin, M. Y. (2010). Agreement, acknowledgement and alignment: The discourse-pragmatic functions of *hao* and *dui* in Taiwan Mandarin conversation. *Discourse Studies* 12.2: 241-267.

Whitehead, K. (2011). Some uses of head nods in "third position" in talk-in-interaction. *Gesture* 11:2 (2011), 103–122.

Xu, J. (2009). *Displaying over recipiency: Reactive tokens in Mandarin task-oriented conversation*. University of Nottingham. PhD dissertation.

Yang, J. 2013. *A multimodal study of response token dui dui dui in Mandarin conversation*. University of Alberta. MA thesis.

Yin, S. (2010). Yingda ciyu de yuyi yuyong leixing ji goucheng (Categorization and structure of response tokens). Zouxiang dangdai qianyan de xiandai hanyu yufa yanjiu guoji xueshu yantaohui, Beijing (Conference paper presented at International conference of research on grammar in Modern Chinese, Beijing).

Yngve, V. (1970). On getting a word in edgewise. In *Papers from the sixth regional meeting*. Chicago Linguistic Society: 567-577. Chicago. Il: Chicago Linguistic Society.

Yu, K.-L. (2004). *The Functions and Usage of ''Dui'' in Mandarin Spoken Discourse.* Unpublished MA thesis. National Tsing Hua University, Taiwan.

Zheng, Y. (2007). 'En' de huayu gongneng fenxi. （Analysis of pragmatic functions of 'en'）. *Nanfang lunkan* (South journal). 2007.4:56-57.

Appendix A: Transcript symbols

The transcription system in this study is mainly based on GAT2 (Selting et al. 2009).

Symbol	Meaning
[ ]	Overlap
(0.5)	Pause duration in seconds and tenth seconds
:,::	Segment lengthening
XXX	Unintelligible passage with each "X" representing one syllable
↓↑	Head nod (downward and upward movement)
01	Line number
HAISHI	Stressed word
,	Rising pitch movement of intonation unit
_	Level pitch movement of intonation unit
;	Falling pitch movement of intonation unit
.	Low falling pitch movement of intonation unit
°h, °hh, °hhh	Breathing in, according to its duration

Appendix B: Glossing conventions

ASSC	Associative (*de*)
PRT	Particle
NOM	Nominalizer
CL	Classifier
NEG	Negative (*bu*)
3SG	Third person singular
POSS	Possessive (*de*)
CRS	Currently relevant state (*le*)
PFV	Perfective aspect (*le*)
EXP	Experiential aspect (*guo*)
NOM	Nominalizer (*de*)