NA: On what it encodes
- from A reference interpreting perspective

Qingqing Wang

NTNU

August, 2013

献给我的父母

Table of Contents

List of tables

List of figures

List of Abbreviations

ADJ	Adjective
ASP	aspect marker
ASSOC	associative phrase marker
CL	classifier
CS	cognitive status
INT	interrogative
Mod	modifier marker
N	noun
NEG	negatives
PART	particle
POSTP	postposition
PREP	preposition
RC	relative clause
S	sentence
SG	segment

Acknowledgement

I would like to express my deepest appreciation to my supervisor Kaja Borthern for her academic support, patience and constant encouragement during my thesis writing. I will always remember the precious comments and inspirations that she gave me through so many times of revising. My independant thoughts are respected and encouraged, because of this, writing the thesis has been a pleasure. I cannot imagine a better supervisor.

I owe a debt of gratitude to my Phonetics professor, Wim van Dommelen. Thank you for helping me with Excel. The methods that you taught me have been helpful in the statistics calculation of the data.

The Institution of Language and Communication (ISK) has been my home for the past 2 years, and I am deeply grateful for the nice personnel and daily support of ISK. I would also like to thank my friends and fellow master students, Signe Rix Berthelin, Atena Mohanmod, Cecilie Slinning Knudsen and Pia Farstad Eriksen. Thank you for the nice and friendly atmosphere in the reading room. It has been a pleasure for me to stay at Dragvoll thanks to all of you.

I would like to thank my parents who have been loving me unconditionally, thank you for supporting my study in Trondheim. You are the best parents. My thanks also go to my boyfriend Emanuele Di Santo. Grazie per essere sempre lì per me tra alti e bassi.

Chapter 1. Introduction

This thesis focuses on the study of the Mandarin word 那 nà. In The Contemporary Chinese Dictionary (one of the most authoritative Mandarin dictionaries), the word *na* is cited as corresponding to a distal demonstrative as in example (1-2) and a conjunction being used to continue the mind flow from the previous discourse, as well as proposing the new opinion accordingly, such as that in (3).

(1) 那 棵树
   na ke shu
   that CL tree
   "that tree (over there)"

(2) 那 是 1937 年
   na shi 1937 nian
   that be 1937 year
   "that is the year of 1937"

(3) 这 样 做既然 不 行, 那 你 打 算 怎 么办 呢？
   Zhe yang zuo ji ran bu   xing *na* ni da suan zen me ban ne
   This type do since NEG work *na* you plan   how   do INT
   "Since doing (it) like this did not work, what do you plan to do (with it) then?"

Na has been considered as starting to function as a definite article in Mandarin by previous linguists (e.g. Li & Thompson 1989, Hedberg 2003, etc), as the languge is not equipped with an article system like that of English. The direction of this thesis is to explore what meaning is encoded in *na* when it functions as a determiner or pronoun as in (1) and (2).

The study also attempts to test the previous claims that have been made about *na* by comparing the investigation with other similar investigations (e.g. the one by Gundel, Hedberg and Zacharski (1993), henceforth GHZ) that have been conducted previously.

This study will mainly be conducted within the framework of the Givenness Hierarchy theory (GHZ, 1993), and coding for the cognitive status (henceforth, CS) of each datum will be done with strict reference to the *Coding Protocol for Statuses on the Givenness Hierarchy* (GHZ, 2006, henceforth the Coding Manual). While the Givenness Hierarchy alone will probably not be able to capture and explain the whole reference assignment process in natural language discourse, this study also involves the theoretical accounts of Relevance Theory for a more satisfying explanation of language communication.

The rest of the thesis is organized as such: Chapter 2 gives an introduction to the *Mandarin* language whereas Chapter 3 reviews claims that have been made about *na* by previous linguistics. Chapter 4 presents the theories being used in the thesis, including Relevance Theory, the Givenness Hierarchy theory and Donnellan's (1971) distinction between attributive and Referential reading of descriptive expressions. Chapter 5 concerns issues about the methodology, it explains the architecture behind my investigation and the ideas that lead through the data collection; a sample of coding for CS will also be included in this chapter. Chapter 6 presents the investigation per se, with sections corresponding to the 7 data groups, which are distinguished according to the internal structure of the nominal. This section also includes the summary of the investigation and the conclusion about what cognitive status is encoded in *na*. Residual issues about the Givenness Hierarchy theory will be discussed in Chapter 7 while a brief conclusion about the thesis is made in Chapter 8.

Chapter 2. The Chinese language

2.1 Mandarin: Pǔ tōng huà and Guó yǔ

The word *Mandarin* is an established linguistic term in the West that commonly denotes the Chinese language. In the year of 1955, the government of People´s Republic of China established a uniformed language for the nation, being named as modern standard Chinese, it is also known as *pǔ tōng huà* (literally, common language). *Pǔ tōng huà* is based on the pronunciation of the Beijing dialect, the grammar of the Chinese spoken in the northern part of China, and the vocabulary of modern vernacular literature. It is the official language of People´s Republic of China. During the early 1950s, the government in Taiwan also proclaimed a uniform language based on the Beijing dialect, and this uniformed language in Taiwan is known as *Guó yǔ* (literally, national language). In this thesis, I will adopt the linguistic term Mandarin, in the sense of the Chinese language including both *pǔ tōng huà* and *guó yǔ.*

A total number of 56 ethnic groups live within the territory of China, where approximately 129 language varieties from 5 language families are spoken. This complex situation results in numerous mutually unintelligible dialects, thus *pǔ tōng huà* was initially established to facilitate the communication between speakers of mutually unintelligible dialects, as well as to promote universal education among the Chinese people. [1]As has been mentioned, *pǔ tōng huà* is a manually unified language being spoken across a vast geographic area, it is a rather ideal language. In reality, the usage of *pǔ tōng huà* has been inevitably influenced by the mother tongues in the local areas, therefore, the *pǔ tōng huà*s being spoken in the various areas of China are not precisely the same. The same applies to *guó yǔ* in Taiwan, *guó yǔ* has been influenced by the Taiwanese Hokkien (first language of about 70% of the population of Taiwan) and other mother tongues of the local people such as Hakka (spoken natively by about 15% of the Taiwanese).

---

[1] During this process, the signs of the writing system have been simplified, in order to enhance education among the large number of farmers and workers. Meanwhile, the writing system of *guó yǔ* (the Chinese being used in Taiwan) remains intact, which is recognised as traditional Chinese, to be distinguished from the simplified writing system in mainland China.

2.2 Word order

Languages are categorised into four groups according to their word orders. The four categories are: VSO (Verb Subject Object) languages, SVO languages, SOV languages and those for which no basic word order can describe. According to Li & Thompson (1989), Mandarin appears to fall in the last category, as they state: "in Mandarin, the positions of elements in a sentence interact with other features of the language, such as the notion of topic, and the expression of definiteness and directionality, and we have noted that Mandarin may be undergoing a change from an SVO to an SOV word order. " (Li & Thompson, 1989:26).

In this thesis, most of the presented examples come in the SVO word order. All the data are authentic examples being selected from natural language discourses, they are generally more sophisticated than created examples. In fact, most of the presented examples contain complex sentences with two to three subordinate clauses, and complex sentences tend to have an SVO order in Mandarin.

2.3 Tone, Pin yin and the Chinese characters

" A language is a ´tone language´ if the pitch of the word can change the meaning of the word. Not just its nuances, but its core meaning." (Yip, 2002:1) It is widely acknowledged that Mandarin is a tone language. For instance, the syllable consisting the consonant /m/ and the vowel /a/ can be used to mean ´mother´ when the syllable is marked with high level tone (the first tone), and its meaning can be changed to ´horse´ when it is marked with a contour tone. Except for a neutral tone that indicates an unstressed, short reading of the syllable, Mandarin has four lexical tones, namely the first tone, which is a high level tone; the second tone, which is a rising tone; the third tone, which is the only contour tone (beginning from mid-low to low and then rise to mid-high pitch), and the fourth tone, which is a high falling tone.

Notably, tone is not the only factor that determines lexical meaning, the character of the lexical (i.e. the sign in which the lexical is written) plays a significant role as well, e.g. The syllable /ma/ being marked with contour tone corresponds to more than one characters, it means ´horse´ only if the corresponding character is 马, when the corresponding character is changed to 码, for instance, the lexical is accordingly

changed into a measurement unit meaning ´yard´, regardless of the same syllable combination and tone.

Pīn yīn, literally ´spell out sound´, is the official phonetic system for transcribing the pronunciation of Mandarin characters into Latin script. Being developed in the 1950s based on earlier forms of Romanization, the *pīn yīn* system is commonly used to teach *pǔ tōng huà* (to both first language and second languages learners) and spell Chinese names in non-Chinese linguistic contexts, it is also the base of the various input methods to enter Chinese characters into mobile phones and computers. Incidentally, it is not the same system as IPA[2]. Within the *pīn yīn* system, the four lexical tones are marked by diacritics placed above vowels. Take the syllable /ma/ as an example, the four tones are marked as mā (the first tone, T1), má (T2), mǎ (T3) and mà (T4).

The selected examples in this thesis are systematically glossed in three layers below the written characters. The three layers are: The first layer of *pīn yīn*, the second layer of POS (Part of Speech for the word), and the third layer of the free translation in English. To be reader-friendly, pronunciations of the relevant referring expressions are marked in bold in both the *pīn yīn* layer and the POS layer. In order to avoid preconceptions, *na* will systematically not be glossed in the selected examples. (i.e. *na* is explained as na in the POS layer)

2.4 Aspect markers

Unlike most of the languages in the world, Mandarin does not have tense markers. That is, according to Li & Thompson (1989): " The language does not use verb affixes to signal the relation between the time of the occurrence of the situation and the time that situation is brought up in speech." The way in which Mandarin indicate time is to use aspect markers. As Li & Thompson further stated: " Aspect, on the other hand, refers, not to the time relation between a situation and the moment of its being mentioned in speech, but, rather, to how the situation itself is being viewed with

---

[2] The International Phonetic Alphabet (IPA) is an alphabetic system of phonetic notation based primarily on the Latin alphabet. It was devised by the International Phonetic Association as a standardized representation of the sounds of oral language.

respect to its own internal makeup." (Li & Thompson,1989:184). Some frequently used aspect markers in Mandarin are, for instance, the perfective aspect marker *le*, the progressive aspect marker *zhèng zaì* etc. Such aspect markers are marked with the abbreviation *ASP* in the POS layer.

2.5 Noun phrases

The internal structure of noun phrase might be the most important syntactic issue to be understood in this thesis, since all of the referring expressions in the investigation are noun phrases, and the data are grouped according to the syntactic structure of the relevant expressions. A noun phrase contains at least one noun, which can be a personal pronoun such as *wǒ* (´I´), a common noun such as *píng guǒ* (´apple´) or a compound noun such as *chuáng dān* (´bed line´). It is also possible for the noun to be preceded by other elements. According to Li & Thompson (1989), the following three types of elements typically precede nouns: classifier phrases, associative phrases and modifying phrases.

2.5.1 Classifier phrase

Mandarin has classifiers. A classifier is a form that marks a noun of a specific semantic class. The choice of the classifier depends on the ontological category of the noun, and each noun has its corresponding classifier that is fixed. For example, the common noun jìng zi (´mirror´) can only be preceded by *miàn*, a classifier that is commonly used for thin and flat objects. The classifier *toú* corresponds only to livestock, and it is grammatically incapable for nouns denoting human beings. Classifiers are required to occur together with determiners, they rarely occur alone.

2.5.2 Associative phrase

"Associative phrase denotes a type of modification where two noun phrases (NPs) are linked by the particle – de 的. Within such an NP, the first noun phrase together with the particle - de is the associative phrase, and the second noun phrase is the noun being modified " (Li & Thompson, 1989: 113).

The particle - *de* of this type is glossed as ASSOC, meaning associative phrase marker in my examples. In order to be precise, - *de* in possessive phrase is marked as POSS, e.g. wǒ de (´I POSS` - my).

2.5.3 Modifying phrase

"Modification is a type of syntactic construction in which a head is accompanied by an element typically not required by it." (Matthews, 2007:248). A modifying phrase can be either a relative clause or an attributive adjective. The phrase in bold in (1) is a relative clause in Mandarin.

（1）   他   遗失 了   那   条   我   们   一 起   买 的   手   链。

       ta   yi shi le   *na* tiao wo men yi qi   **mai de**   shou lian

       he lose ASP **that CL   we   together buy MOD** bracelet

       "He has lost the bracelet that we bought together."

The phrase in bold is a relative clause that modifies the head noun *shou lian* (´bracelet´). Similar to associative phrases, modifying phrases are usually connected to the head noun by the particle de (的[3])

2.5.4 Word order of the elements inside the noun phrase

A noun phrase contains at least a noun, in addition to that, elements such as associative phrase, classifier phrase and modifying phrases can be optionally applied to precede the noun. In some complicated examples, more than one of those elements

---

[3] Notice here, the pronunciation *de (neutral tone)* corresponds to three particles in Mandarin, namely 的, 得, 地. Only the one that corresponds to the character 的 is the modifying phrase marker that we are talking about.

的 is generally used to connect modifying elements with nouns.

地 indicates adverbs, it is used to connect modifying elements to verbs. Besides, it only occurs in preverbal position, being preceded by the modifying element. e.g. 大声 地吼 (da shen **de** hou (the Verb), shout loudly).

得 also indicates adverbs, similar to 地, it is used to connect modifying elements to verbs. However, it is only allowed to occur in post verbal position, being followed by the modifying element. e.g. 哭得好绝望 (ku (the verb) **de** hao jue wang, to cry desperately)

occur within the same nominal, the order of them are fixed according to the schema in (2) below:

(2)a. associative phrase + classifier / measure phrase + relative clause + adjective + noun

   b. associative phrase + relative clause + classifier / measure phrase + adjective + noun

(Li & Thompson, 1989: 124)

Chapter 3. Previous Claims about *na*

3.1 Hedberg (1996)

According to Hedberg (1996), the purpose of her article is to show show that the Givenness Hierarchy framework can shed new light on the claims concerning the encoding of definiteness in Mandarin proposed by Li and Thompson (1976, 1981). The article discuss the behaviours of both the distal demonstrative *nei*[4] and the proximal demonstrative *zhè*, I will only focus on the arguments about neì in this review. Li and Thompson´s (1989) well-known claim about the Chinese distal demonstrative *na* is that it is beginning to function like the English definite article *the*, and this is what Hedberg (1996) argued for.

GHZ (1993) constructed the Givenness Hierarchy for Chinese as in (3) below:

(3). THE GIVENNESS HIERARCHY: MANDARIN

In Focus	> Activated >	Familiar	> Uniquely Identifable >	Referential	> type identifiabl
{ Ø	{TA		{*nei* N}		{*yi* N ´a N´
ta ´s/he; it}	*zhe* PROX				Ø N }
	nei DISTAL				
	zhe N}				

As (3) illustrates, GHZ (1993) claim that Mandarin does not have linguistic forms to signal that a referent is Familiar or Referential. Regarding on the correlation between *nei* and a specific cognitive status (henceforth CS), they claim that the distal demonstrative *nei* requires only activation when it is used alone, and it encodes Uniquely Identifiable when in the form *nei N*.

According to Hedberg (1996), the claim that *nei* is beginning to function equivalent to a definite article in English, (i.e. *the*) is not easy to be explained in the Givenness Hierarchy due to the logical relation between the CSs. The reason is that *nei* is initially a distal demonstrative whose English counterpart is *that*, which signals Familiar in the hierarchy for English. According to GHZ (1993), the six CSs on the

---

[4] The pronunciation being recorded in the article is nei, I chose to follow the author´s intention in this chapter of literature review.

hierarchy are implicationally related, that is, if the referent is Familiar to the addressee, it is automatically also Uniquely Identifable to him. Thus, it is unclear if *nei* is supposed to be more precisely translated as *the* in a case where it is actually translated as *that*, since using the latter acknowledges using of the former.

A more promising example given by Li and Thompson (1975) is quoted in (4), whereby *nei* occurs in a relative clause.

(4). Wo diu le de nei ben shu
     I lose ASP ATT that CLS book
     "the book I lost"

Hedberg (1996) argues that the addressee is still required to know the fact that the speaker lost a pen previously to succeed in retrieving the referent. " […] the information encoded in the nominal alone is still insufficient to enable the addressee to uniquely identify the referent without previous Familiarity." (Hedberg, 1996: 185). Once the condition of Familiarity is met, *that* can be used as well. Therefore, it is still unclear if *the* is the only counterpart of *nei* in this example.

The ideal example for *nei* signalling Uniquely Identifable prosed by GHZ (1993) is (5), whereby the referent is only Uniquely Identifable to the addressee.

(5) 我 昨晚   睡 不 着。   隔壁 那 只 狗 让 我 一直   醒 着。
    wo zuowan shui bu zhao   gebi nei zhi gou rang wo yizhi xing zhe
    I last night sleep NEG PART next door that CL dog make me always awake PART
    " I cound´t sleep last night. The dog next door kept me awake."

The addressee does not necessarily need to know that the speaker´s neighbour has a dog previously for the appropriate use of *nei*. The information encoded in the nominal is adequate enough for the addressee to uniquely construct the representation, therefore the referent does not need to be Familiar to the addressee.

Hence, Hedberg(1996) agrees with the observation proposed by Li and Thompson (1989) that *nei* has one of the functions of the definite article property.

3.2 Brøseth & Jin (2008)

While GHZ (1993) proposed the correlation between each of the cognitive status and their corresponding linguistic forms in five languages, Heidi Brøseth and Fufen Jin (Heidi & Jin, 2008) focus on one particular CS Referential and its relevant referring forms in Chinese. The study was conducted on a comparative base with analysis of examples from Chinese, Norwegian and English.

Chinese does not have an article system such as the one of English, that is, indefinite article and definite article do not exist in Chinese. However, as Brøseth & Jin state: "[…] several authors (Li and Thompson 1981; Gundel et al. 1993; and Robertson 2000) maintain that the distal demonstrative determiner in Chinese "*nei* CL N" is *beginning* to take on some functions associated with the definite article." (Brøseth & Jin, 2008:115). A similar phenomenon has been observed in Norwegian where the pre-positioned definite articles *den* (´the´) is homophonous with the distal demonstrative determiner *den* ´that´. Based on the assumption proposed by previous linguists and the pattern of Norwegian homophonous articles, Brøseth & Jin proposed the hypothesis that Chinese at the current stage of development has two homophonous linguistic forms, *nei* the emerging definite article and *nei* the distal demonstrative determiner, and that they signal two different cognitive status depending on the context. Nevertheless, although the tendency that *na* starts to functions equivalently to a definite article is observed, it is not tested and exemplified further in their article.

The novelty of the proposal in Brøseth & Jin´s article is that both Chinese and Norwegian associate the cognitive status Referential with a particular linguistic form. The Chinese part of this argument contradicts the claim made by GHZ (1993), as the latter acknowledge no Chinese linguistic signalling Referential. Conversely, as Brøseth & Jin (2008) claim, some examples of their Chinese data indicate that the referent of the form ´nei CL N´ is only Referential. However, this does not lead them to conclude that the emerging definite article *nei* has the cognitive status Referential as its necessary and sufficient condition. Rather, they propose a new approach: In addition to the distal demonstrative determiner *nei* and the emerging definite article

nei that already have been observed, Chinese has an indefinite article *nei* as well. It is this third *nei* that can signal Referential.

Cconcerning the emerging definite article *nei*, Brøseth & Jin (2008) hypothesize that *nei* CL N in Chinese will continue to have a demonstrative function while developing into a definite article, and that "the demonstrative determiner nei N ´that N´ could have a higher cognitive status than the emerging definite article nei N ´the N´, which has Uniquely identifiable as its necessary and sufficient condition." (Brøseth & Jin, 2008:121). This hypothesis has not been investigated in the article, but they suggest that the necessary condition for the demonstrative *na* should be at least Familiar, following the assumptions in the Givenness Hierarchy theory.

To summarize, Brøseth& Jin (2008) propose that Chinese has three homophonous linguistic forms of *nei* that signal different cognitive statuses. The emerging definite article *nei* N ´the N´ signals Uniquely Identifiable. The demonstrative determiner *nei* N ´that N´ is suggested and supposed to have at least Familiar as its necessary condition. Last but not least, the indefinite article *nei* of *nei* CL N signals Referential. Therefore, Brøseth& Jin (2008) claim that Chinese does have a linguistic form whose necessary and sufficient condition is Referential.

In Chinese, a nominal is commonly required to be proceeded by a specific classifier. For example, the classifier 棵(ke) only precedes nominals denoting trees. Therefore, it is possible to use *na CL* (without the nominal) solely as a referring expression, since the classifier encodes information about the referent´s ontological category. It is likely that classifiers play a role in the reference assigning process by narrowing down the referent set.

Both *nei N* and *nei CL N* are possible referring forms in Chinese, but the latter is more frequently and formally used. At the beginning of the conclusion part in Brøseth & Jin (2008)´s article, they state:" In this paper, we questioned the assertion of Gundel et al (1993) that nei CL N has the cognitive status Uniquely identifiable as its necessary and sufficient condition". A closer examination of the GHZ (1993) article shows that such an assertion has not been made. More precisely, what GHZ (1993) concludes is

that nei N (**without CL**) has Uniquely identifiable as its necessary and sufficient condition. Although the examples presented in GHZ (1993)´s paper seem to indicate that by ´nei N´, they actually mean ´nei CL N´ or even more, but this is not clearly underspecified. Considering the special role of classifiers, I choose to distinguish *nei CL N* and *nei N* as two referring expression forms in my data analysis.

Lastly, the pronunciation of the form 那 should be clarified, since the notion of ´homophonous forms´ has been mentioned in the Brøseth & Jin (2008) article. The citation pronunciation of 那 in the Chinese dictionary is nà, it is also explicitly stated that 那 is read as its citation pronunciation in structures of *na* and *na* N; on the other hand, it should be read as neì when it is followed by whichever classifier or the cluster of a number and a classifier. In other words, *na* N, *nei* CL N are read differently.

Chapter 3. Theoretical Frameworks

3.1 The Givenness Hierarchy Theory

3.1.0 Introduction

Regarding the issue of reference assignment, previous linguists have attempted to investigate the process and its subtasks from various viewpoints. One of the subtasks in reference understanding is to access the form of the referring expressions, so as to unveil the speaker´s assumptions about the addressee´s knowledge and attention state of the intended referent. The Givenness Hierarchy Theory proposed by Jeanette Gundel, Nancy Hedberg and Ron Zacharski (henceforth, GHZ, 1993) offers a nice account for the distribution and understanding of various linguistic forms of referring expressions in natural language discourse. It is the main approach that the investigation of my thesis based on.

What is the major belief that the Givenness Hierarchy theory builds on? The main premise of Givenness Hierarchy Theory is that " […] different determiners and pronominal forms conventionally signal different cognitive status (information about location in memory and attention state), thereby enabling the addressee to restrict the set of possible referent." (GHZ, 1993:274). That is to say, according to my understanding, the form of the referring expression per se encodes built-in information that reveals the referent´s location in the addressee´s memory and attention state (or how to establish it), and such a mental location is assumed by the speaker for the addressee when the referring expression is formulated. The addressee makes use of this information to narrow down the possible referent range, and eventually identify the intended referent together with other factors.

3.1.1 The Givenness Hierarchy

In the introduction part, I have clarified the premise of the Givenness Hierarchy theory pointed out by GHZ (1993). In this part, I will summarize the theoretical aspects of the Givenness Hierarchy.

As mentioned in the previous section, cognitive statuses are information about location in memory and attention state. With the Givenness Hierarchy, GHZ categorize and distinguish six cognitive statuses, namely *In Focus, Activated, Familiar, Uniquely Identifable, Referential* and *Type Identifiable*. They also discuss the nature of each cognitive status (CS) while describing the condition of appropriately using the CSs. Besides, they line out the logical and empirical relations among the six CSs and point out the mapping between each CS and its corresponding linguistic form/forms across languages.

The Givenness Hierarchy is a scale of the six CSs that GHZ proposed to be relevant to the forms of referring expressions cross-linguistically. The six CSs are related and ranked according to the degree of givenness in the hierarchy shown in (6) below.

(6). The Givenness Hierarchy

In Focus	>	Activated	>	Familiar	>	Uniquely Identifable	>	Referential	>	Type Identifiable
{it}		{that this this N}		{that N}		{the N}		{indefinite this N}		{a N}

For each suggested English referring form, the corresponding cognitive status is both necessary and sufficient for its proper use. As an example of this, the referring form *it* signals that the referent is not only in the addressee's short term memory, but also is at the current center of attention. Similarly, the definite article *the* signals that the addressee is supposed to be able to uniquely identify the referent.

The logical relation between these six CS is that they are implicationally related to each other, rather than being mutually exclusive as other theoretical approaches of the same kind claim (see e.g Ariel, 1988 and Garrod & Sanford, 1982). This is explicitly stated by GHZ (1993): "[…] in the model we proposed here the statuses are implicationally related (by definition), such that each status entails (and is therefore included by) all lower statuses, but not vice versa." (GHZ, 1993: 275). For instance, the English demonstrative determiner *that* signals that the referent is not only Familiar to the addressee and thus can be retrieved from memory, but is also, obligatorily, Uniquely Identifiable, Referential and Type Identifiable to the addressee. Being able to

uniquely identify a tree naturally means that the addressee knows what kind of thing a tree is, for example.

The individual statuses are characterized and exemplified below. Except for (3a), the involved English examples are quoted from the original article of GHZ (1993).

IN FOCUS: "The referent is not only in short-term memory, but is also at the current center of attention. […] Entities In Focus generally include at least the topic of the preceding utterance, as well as any still-relevant higher-order topics" (GHZ, 1993:279). In their Coding Protocol (GHZ, 2007), this general characteristic is divided into a set of criteria, as being listed in (7) below:

(7) A referent is IN FOCUS if it meets at least one of the following criteria:

7.1   It is the referent of a DP in a syntactically prominent position (incl. non-overt subjects) in the main clause of the immediately preceding sentence. Such position include at least the following:
- Syntactic subject
- Syntactic focus (e.g. existential, cleft, identifying) of immediately preceding main clause (This excludes indefinite predicate nominal such as "an architect" in "He is an architect.")

7.2   It is the referent of a DP earlier in the same sentence

7.3   It is a higher level topic that is part of the interpretation of the preceding clause   (whether it is overtly mentioned there or not).

7.4   It is part of the interpretation of each of the two preceding clauses.

7.5   It is the event denoted by the immediately preceding sentence

(8a) is an English example whereby the referent has the CS In Focus, and (8b) its Chinese counterpart.

(8a) My neighbor has a dog. It has been barking the whole night. (English)
(8b) 我 邻居 有 一 只 狗, 叫 了 一 整 夜。(Chinese)

    wo   linju   you yi zhi gou   jiao le   yi zheng ye

    I   nabour have one CL dog   bark ASP one whole night

Judging by the context, the dog is the focus of the first segment in (3a). In the English example, the speaker uses the third person personal pronoun *it* to indicate that the dog is not only in short-term memory, but is also at the current center of attention. Moreover, it is evident that the dog is continued as the topic of the second segment. In the Chinese version, however, the pronoun is dropped, nevertheless, the dog is In Focus.

ACTIVATED**: "**The referent is represented in current short-term memory. Activated representations may have been retrieved from long term memory, or they may arise from the immediate linguistic or extralinguistic context." (Gundel et al, 1993:278). Referents that are Activated but not necessarily In Focus require more processing effort of the addressee than those are IN FOCUS, because the information of the intended referent is not located in the current center of the addressee's attention state, although it is already in his current short-term memory. GHZ (1993) conclude that activation is necessary for appropriate use of all pronominal forms and the definite demonstrative determiner *this*. In addition, it is sufficient for the demonstrative pronoun *that* and for stressed personal pronouns.

According to the Coding Manual (GHZ, 2007), a referent is ACTIVATED if it meets one of the following criteria in (9):

(9) A referent is ACTIVATED if it meets one of the following criteria.

9.1   It is mentioned in one of the immediately preceding two sentences.
9.2   It is something in the immediate spatio-temporal context that is Activated by a simultaneous gesture or eye-gaze.

9.3   It is a proposition, fact or speech act associated with the eventuality (event or state) denoted by the immediately preceding sentence(s).

(10a) and (10b) are the English and Chinese examples respectively, whereby the referents have the CS Activated.

(10a) I couldn´t sleep last night. That kept me awake.
(10b) 我 昨晚 睡 不 着。那 让 我 一直 醒 着。

> wo zuowan   shui bu   zhao *na* rang wo yizhi xing   zhe
> I last night sleep NEG   PART that make I   always awake PRAT

In (10a), the demonstrative pronoun in both English and Chinese indicates that the referent has been Activated previously either by being mentioned verbally or by being pointed out by gestures, and that the referent is still in the addressee´s short-term memory.

FAMILIAR: 'The addressee is able to uniquely identify the intended referent because he already has a representation of it in memory (in long-term memory if it has not been recently mentioned or perceived, or in short-term memory if it has) (Gundel et al, 1993: 278). The CS Familiar is necessary for all personal pronouns and definite demonstratives, and it is sufficient for appropriate use of the demonstrative determiner *that*.

According to The Coding Manual (GHZ, 2007), a referent is FAMILIAR if it meets one of the following criteria in (11):

(11). A referent is FAMILIAR if it meets one of the following criteria
11.1   It was mentioned at any time previously in the discourse.
11.2   It can be assumed to be known by the hearer through cultural/encyclopedic knowledge or shared personal experience with the speaker.

An English example of Familiar referent is found in (12a), and in (12b) its Chinese counterpart:

(12a) I couldn´t sleep last night. That dog (next door) kept me awake.
(12b) 我 昨晚   睡 不 着。 (隔壁) 那 只 狗 让 我 一直 醒   着。
> wo zuowan shui bu zhao   gebi *na* zhi gou rang wo yizhi xing   zhe
> I last night sleep NEG PART next door that CL dog make me   always awake PART

(12a) and (12b) are appropriate only if the addressee already knows that the speaker´s neighbour has a dog, in which case it is Familiar to the addressee.

UNIQUELY IDENTIFIABLE: "The addressee can identify the speaker's intended referent on the basis of the nominal alone." (GHZ, 1993). The addressee may be able to identify a specific referent because he already has an existing representation in his memory, but he shall also be able to construct one from scratch based on what is encoded in the nominal itself. GHZ (1993) conclude that this status is a necessary condition for all definite reference, and it is both necessary and sufficient for appropriate use of the definite article *the*.

According to the Coding Manual (GHZ, 2007), a referent is UNIQUELY IDENTIFIABLE if it meets one of the following criteria in (13):

(13).

13.1 The referring form contains adequate descriptive/conceptual content to create a unique referent.

13. 2 A unique referent can be created via 'bridging inference' by association with an already Activated referent (e.g. A house …. the front door)

Example (14a) and (14b) demonstrates one of the felicious conditions for the CS Uniquely Identifable.

(14a) I couldn´t sleep last night. The dog (next door) kept me awake.

(14b). 我 昨晚　　 睡　不　　着。 (隔壁) 那 只 狗　让 我 一直 醒　　 着。
　　　　wo zuowan shui bu zhao 　gebi *na* zhi gou rang wo yizhi xing 　zhe
　　　　 I last night sleep NEG PART next door that CL dog make me always awake PART

Following the argument of GHZ (1993) from above, the addressee does not necessarily need to know that the speaker´s neighbor has a dog for (14a-b) to be appropriate. He/she may be able to achieve a referent on spot when the utterance is accessed. In such a case, the information encoded in the nominal will have to be adequate enough for the addressee to construct the representation.

Regarding the distinction between the CS Uniquely Identifable and the CS Referential, GHZ (1993) state: "Thus, expressions which are Referential but not Uniquely Identifable require the addressee to construct a new representation as

determined by the content of the Referential expression along with the rest of the sentence. For expressions which are both Referential and Uniquely Identifable, on the other hand, the addressee is expected to construct or retrieve a representation on the basis of the referring expression alone (See Webber 1983 and Millikan 1984 for further discussion)." [5] (GHZ, 1993:277)

REFERENTIAL: " The speaker intends to refer to a particular object or objects. To understand such an expression, the addressee not only needs to access an appropriate type − representation, he must either retrieve an existing representation of the speaker´s intended referent or construct a new representation by the time the sentence has been processed. The status ´Referential´ is necessary for appropriate use of all definite expression, and it is both necessary and sufficient for indefinite *this* in colloquial English." (GHZ, 1993: 276).

In the Coding Manual (GHZ, 2007), the criteria for a referent to have the CS Referential is listed as those in (15) below:

(15). **A referent exists, is REFERENTIAL, if it meets one of the following criteria:**

15.1   It is mentioned subsequently in the discourse.

15.2   It is evident from the context that the speaker intends to refer to some specific entity.

For examples, see (16a) and (16b) below:

(16a)   I couldn´t sleep last night. A dog (next door) kept me awake.

(16b)   我 昨晚　 睡 不 着。(隔壁的) 一 只 狗 让 我 一直 醒　 着。

　　　 wo zuowan shui bu zhao　 gebi　　 yi zhi gou rang wo yizhi xing　 zhe

　　　 I last night sleep NEG PART next door　 one CL dog make me　 always awake PART

Examples (16a) and (16b) can be interpreted in two ways. The referent has the CS Type Identifiable when the speaker only intends to explain the cause of his sleeping problem, which means, the speaker merely intends to communicate that it is a barking dog rather than other factors that keeps him awake for the whole night. However, if

---

[5] I will make use of Webber and Millikan's argument later, because I am not completely convinced concerning GHZ's definition of the CS Uniquely Identifable.

the speaker has more specific intention, for example, if he intends to complain about a particular dog of his neighbor, the referent then has the CS Referential.

TYPE IDENTIFIABLE: " The addressee is able to access a representation of the type of object described by the expression." (GHZ, 1993:276). This use is plausible as long as the addressee is aware of the meaning of the nominal and can therefore understand what type of the thing the nominal stands for. Type identifiable is necessary for appropriate use of any nominal expressions, and it is sufficient for use of the indefinite article *a* in English, as GHZ (1993) concluded.

According to the Coding Manual (GHZ, 2007), an interpretation is TYPE IDENTIFIABLE if the sense of the phrase (the descriptive/ conceptual content it encodes) is understandable. Thererfore, the interpretations of (17a) and (17b) are Type Identifiable to the addressee if he knows what type of thing a dog is. Nothing more is required.

(17a) (repeating 16a)
I couldn´t sleep last night. A dog (next door) kept me awake.

(17b) (repeating 16b)

我	昨晚	睡	不	着。	(隔壁的)	一	只	狗	让	我	一直	醒	着。
wo	zuowan	shui	bu	zhao	gebi	yi	zhi	gou	rang	wo	yizhi	xing	zhe
I	last night	sleep	NEG	PART	next door	one	CL	dog	make	me	always	awake	PART

3.1.2 Empirical Investigation

Each status on the hierarchy is a necessary and sufficient condition for the appropriate use of a different form or forms. When a particular form is used, the speaker signals that he/she assumes the associated cognitive status is met, as well as the lower CSs. On the other hand, for instance, when the English ´the´ is used, it means that the referent is at least Uniquely Identifable to the addressee, but it can also be Familiar or In Focus to him. Departs from this prediction, GHZ (1993) hypothesize that in actual discourse, the same form can be distributed across more than one status. An empirical

investigation on the distribution of different forms of reference was conducted to test this prediction in five languages, namely Chinese, English, Japanese, Russian and Spanish.

The data comes from a variety of spoken and written sources, which differ in formality and degree of planning. In addition, for all the languages except Russian, they also analysed narrative film descriptions. "The speakers viewed a silent film called *The Golden Fish* and, immediately after viewing the film, described it to another native speaker of their language." (GHZ, 1993:290)

The results of the Chinese part is listed Table 1 below:

Table 1. Distribution of Chinese forms according to highest status

	In Focus	Activated	Familiar	Uni.Id[6]	Referential	Type.Id	Totals
Ø	25	1					26
ta	40						40
zhè		2					2
nèi							
zhè N	12	26	1				39
nèi N	1	7	2				10
yi N					17	2	19
N	12	17	14	49	2	10	104
Totals	90	53	17	49	19	12	240

As the table illustrates, the occurrences with *nei* have at the lowest Familiar. In other words, according to the statistics in Table 1, the necessary condition for using *nei N* requires Familiarity. GHZ (1993) comment in a footnote that it can be just an accident, or it is probably because the use of *nei* for referents that are not Familiar is relatively rare in Chinese. In the article Word order and Cognitive Status in Mandarin, Nancy Hedberg offers a more detailed discussion with examples for the argument that previous Familiarity is not a necessity for the appropriate use of *nei N*. (please refer to the literature review section: Hedberg)

---

[6] Due to space limitation, abbreviations are used. *Uni.Id* is short for the cognitive status Uniquely Identifable, and *Type. Id* is short for Type Identifiable.

The correlation between linguistic forms and highest required status is summarized in Table 2.

Table 2. Correlation between linguistic forms and highest required status

In Focus	Activated	Familiar	Uni.Id	Referential	Type.Id
Ø *ta* ´s/he; it	TA *zhe* PROX *nei* DISTAL *zhe* N		*nei* N		*yi* N ´a N´ Ø N

The correlation table clearly claims that nei requires activation when it is used alone, and nei N requires only uniquely identifiablility.

3.1.3 Summary

The Givenness Hierarchy outlined in this section is one of the main theories that my investigation based on. I will mostly stick to the definitions of each cognitive status that GHZ have discussed in my investigation. GHZ have also proposed the Coding Manual to assist the coders with more concrete syntactic constraints when they analyse examples in the data set. In order to make my investigation result optimally comparable to similar projects that previous linguists (for example, GHZ, 1993, Borthen 2010) have conducted, I will refer to the Coding Manual as one of the most principle guidelines while annotating the CS for each example. Samples of how the data is handled and analysed will be given in the methodology section, together with explanations of the approaches I use in some ambiguous situations.

Nevertheless, there are definitions of the CS Uniquely Identifable that I was not completely convinced and criteria in the Coding Manual that I found challenging to keep align with. Several corresponding problems have occurred during the analysis and discussion arised around these examples motivates me to reconsider my understanding of Uniquely Identifable. Based on the observations of the difficult examples and accounts proposed by other linguists, I will later argue to slightly revise the definition of Uniquely Identifable, so that it will capture systematic constraints on

the use of *nei* and *nei*-related nominal structures better. The discussion will take place along with the specific examples in the context.

GHZ has also conducted an empirical investigation concerning the distribution of Chinese forms according to highest status and summarised the Correlation between linguistic forms and highest required status accordingly. My investigation will follow a similar procedure with data collection, example annotation and distribution analysis. I will also compare my results and conclusion with the ones that GHZ proposed for Chinese in Table (1) and Table (2).

3.2 Relevance Theory

Relevance theory (henthforth, RT) is a pragmatic approach proposed by Deirdre Wilson and Dan Sperber (henthforth, Sperber& Wilson, 1986) to account for the mechanism and cognitive principles behind communication. Reference assignment has been a topic of study by linguists from various linguistic fields, and Relevance Theory accounts for this issue from a cognitive pragmatic perspective. Several linguists have proposed approaches to capture the essence of referent assignment with special focus on the role cognitive factors plays during the process. According to Wilson (1992), who has compared and summarized the theoretical approaches proposed by David Lewis (1979) and Herb Clark (1977), " […], the proposals of Clark and Lewis have a common structure, and clearly this structure is essentially correct. Reference assignment is affected, on the one hand, by the accessibility of potential referents, and on the other hand, by the pragmatic acceptability of the resulting overall interpretation. " (Wilson, 1992: 173). RT shares much of the structure of the accounts proposed by Lewis and Clark, yet having sharper insights on the notion of pragmatic acceptability, which, Wilson (1992) believes, has only been defined rather vaguely in the two previous proposals. That is to say, the main advantage of RT, compared to the previous accounts, would be on the criteria on which they define pragmatic acceptability.

According to Wilson (2004), RT is based on the basic assumption of human cognition that people pay attention to information that seems relevant. When an utterance is

made, the speaker expects attention from the addressee, who shall be attracted by the information being offered. In order to achieve this, the utterance is supposed to be relevant.

Relevance is defined in two terms, positive cognitive effects and processing effort. This is reflected in the definition of relevance in (18) below:

(18). Relevance of an input to an individual

18a.   Other things being equal, the greater the positive cognitive effects achieved by processing an input, the greater the relevance of the input to the individual at that time.

18b.   Other things being equal, the greater the processing effort expended, the lower the relevance of the input to the individual at that time.

(Wilson & Sperber, 2004: 609)

Contextual effects are achieved when the newly presented information strengthens an existing assumption, or contradicts and eliminates it, or yields a contextual implication by combing with it. By processing efforts they mean the mental efforts needed in order to access the interpretation. Three factors influence the amount of the processing efforts needed, the linguistic complexity of the utterance, the accessibility of the context, and the inferential effort needed to compute the contextual effects of the utterance in the chosen context. A positive cognitive effect is a worthwhile difference to the individual´s representation of the world: a true conclusion, for example.

When the speaker is processing the utterance, he/she is automatically looking for *optimal relevance.* This is reflected in the following principle:

(19)   Communicative Principle of Relevance

Every ostensive stimulus conveys a presumption of its own optimal relevance.

(Wilson&Spereber2004: 612)

Optimal relevance is defined as below:

(20)   *Optimal relevance*

An ostensive stimulus is optimally relevant to an audience iff:

20a.   It is relevant enough to be worth the audience's processing effort;

20b.   It is the most relevant one compatible with communicator's abilities and preferences.

From the definition of relevance in (18) above, we see that on the effects side, the more positive cognitive effects are yielded from an utterance, the greater the relevance. On the efforts side, each increment in processing effort detracts from overall relevance. Therefore, according to (19) and (20), when a speaker interprets an utterance, he will be seeking for an interpretation that requires the least processing efforts, yet yields the most positive cognitive effects.

Interpreting an utterance involves several processes. Wilson&Sperber (2004) describe the comprehension process as below:

(21)   Subtasks in the overall comprehension process

21a.   Constructing an appropriate hypothesis about explicit content (EXPLICATURES) via decoding, disambiguation, reference resolution, and other pragmatic enrichment process.

21b.   Constructing an appropriate hypothesis about the intended contextual assumption (IMPLICATED PREMISES).

21c.   Constructing an appropriate hypothesis about the intended contextual implications (IMPLICATED CONCLUSIONS)

(Wilson&Sperber, 2004: 615)

Wilson&Sperber pointed out that these subtasks do not happened in sequential order, they say "In particular, the hearer may bring to the comprehension process not only a general presumption of relevance, but more specific expectation about how the utterance is intended to be relevant (what cognitive effects it is intended to achieve), and these may contribute, via backwards inference, to the identification of explicatures and implicated premises." (Wilson&Sperber, 2004: 615)

From the speaker´s viewpoint, the criteria on relevance guides the speaker to formulate his utterance so that it not only gives the addressee easy access with relatively little processing effort to the intended context and interpretation, but also ensures that when the addressee processes the utterance normally, the first acceptable interpretation he/she encounters will be exactly the one intended by the speaker. This is captured by the Relevance-theoretic comprehension procedure presented in (22).

(22)   **Relevance-theoretic comprehension procedure**
22a.   Follow a path of least effort in computing cognitive effects: Test interpretive hypotheses (disambiguations, reference solutions, implicatures, etc.) in order of accessibility
22b.   Stop when your expectations of relevance are satisfied (or abandoned).

From the addressee´s perspective, he starts the interpretation process by accessing the information encoded in the utterance per se. For example, he will interpret a referring expression by considering the cognitive status being signalled and the concept being denoted by the noun – which in turn gives rise to a certain in-built encyclopaedic knowledge about the given concept. Based on this and the expectation of relevance, he will then try to find the most prominent referent that fits the given content with comparatively little processing efforts. If this interpretation is proved to have enough positive cognitive effects to the addressee (i.e. yields enough cognitive effects that are likely to be true to his sense of the common world), it can then be labelled as ´the first acceptable one with sufficient positive cognitive effects’. It is this interpretation that shall be chosen as the one intended by the speaker, and it is also the only one that the addressee is supposed to choose. According to relevance theory, once such an interpretation is derived, the addressee is definitely not supposed to continue testing other possible interpretations, because they are not in consistency with the communicate principle of relevance. On cases where the most accessible referent candidate is not a relevant one, i.e. is not the one that yields positive cognitive effects, the addressee then shifts to the second most accessible referent to test it for positive cognitive effects and thus relevance.

Wilson (1992) further argues that in order to be acceptable and comprehensible, an utterance does not actually have to be optimally relevant as long as the addressee

understands how the speaker might reasonably have expected it to be relevant. A criterion is thus built:

(23). Criterion of consistency with the principle of relevance

> An utterance, on a given interpretation, is consistent with the principle of relevance if and only if the speaker might rationally have expected it to be optimally relevant to the hearer on that interpretation.

<div align="right">Reference and relevance (Wilson, 1992:176)</div>

What is indicated by this criterion is, according to Wilson, that "[…], all the hearer is entitled to impute as part of the intended interpretation is the *minimal* context and set of contextual effects that would be enough to make the utterance worth his attention." (Wilson, 1992:177).

The argument that the first interpretation tested and found consistent with the principle of relevance is the only interpretation consistent with the principle of relevance is of utterly importance, because it reflects the least-effort strategy pursued by the addressee. As I have mentioned previously, what distinguishes relevance theory from previous theoretic accounts of reference assignment is that RT offers a more precise and predictable criterion that enables the speaker to recognise the only acceptable interpretation as soon as he sees it, and stops at that point. This cannot be achieved by the previous two approaches as they both require the addressee to test each possible interpretation to pick out the most prominent one, which is not a likely pattern of human communication.

Under circumstances where more than one referent is available, and where all are equally accessible, the addressee then seeks for a context which can yield an interpretation consistent with the principle of relevance. The criterion of consistency with the principle of relevance applies especially well for this type of situation. It claims that the first accessible context to yield an acceptable overall interpretation is the only acceptable context, and is the one the hearer should choose. When the referent of a referring expression is not overtly mentioned, and thus can only be retrieved by virtue of bridging assumptions, this criterion helps picking out the most accessible bridging assumption that meets the principles of relevance. To conclude,

Wilson (1992) claims that "inferred referents works just the same way as explicitly mentioned referents, and are affected by the same two factors: accessibility of referents and accessibility of potential context with which these might interact to yield an overall interpretation consistent with the principle of relevance." (Wilson, 1992: 180).

3.3 Bridging Inference

In many of the cases that involve reference, the referents of the referring expressions have been overtly mentioned in the previous discourse. A simple example of this is (24):

(24).   I had a chat with my classmate during the lunch break, she recommended me a nice Thai restaurant.

The third person personal pronoun she refers to my classmate in the first segment of the sentence. We assume that the two elements are in a co-Referential relation with the pronoun being the referring expression and my classmate being the referent.

However, we shall not ignore an important type of examples whereby the referents are not directly mentioned in the previous discourse, yet still possible to be retrieved or uniquely identified by associating with an entity that the addressee already has knowledge about. This phenomenon has been observed and discussed by previous linguists. Diane Blakemore (1992) stated: "[…] in many cases the interpretation of a referring expression has to be *bridged* by assumptions which are not directly mentioned in the preceding utterance, but which are constructed by a series of inferences on the basis of what the hearer knows or believes." (Blakemore, 1992: 75). Building on Clark (1997), Blakemore refers to this type of bridging assumptions as "implicated assumption". An example of a bridging assumption is (25´)[7], which is inferred from (25):

(25).   Nigel bought a fridge. The door fell off three weeks later.

---

[7] The example is quoted from Diane Blakemore´s article

(25´).   The fridge had a door.

The referent of the referring expression *the door* has not been previously mentioned in the first segment where two other nominals are mentioned, namely *Nigel* and *a fridge.* In order to assign an antecedent to the anaphor *the door,* a bridging assumption such as the one in (25´) is needed.

According to Blakemore (1997), what the bridging assumption in (25´) does, is "giving the hearer access to an antecedent for *the door.* In terms of the approach being adopted here, this means that the hearer incorporates the mental representation made accessible by (25´)[8] into the propositional content of the utterance." (Blakemore, 1997: 76) According to my understanding, this can be illustrated in the map below:

Figure 1. Choosing the bridging assumption

---
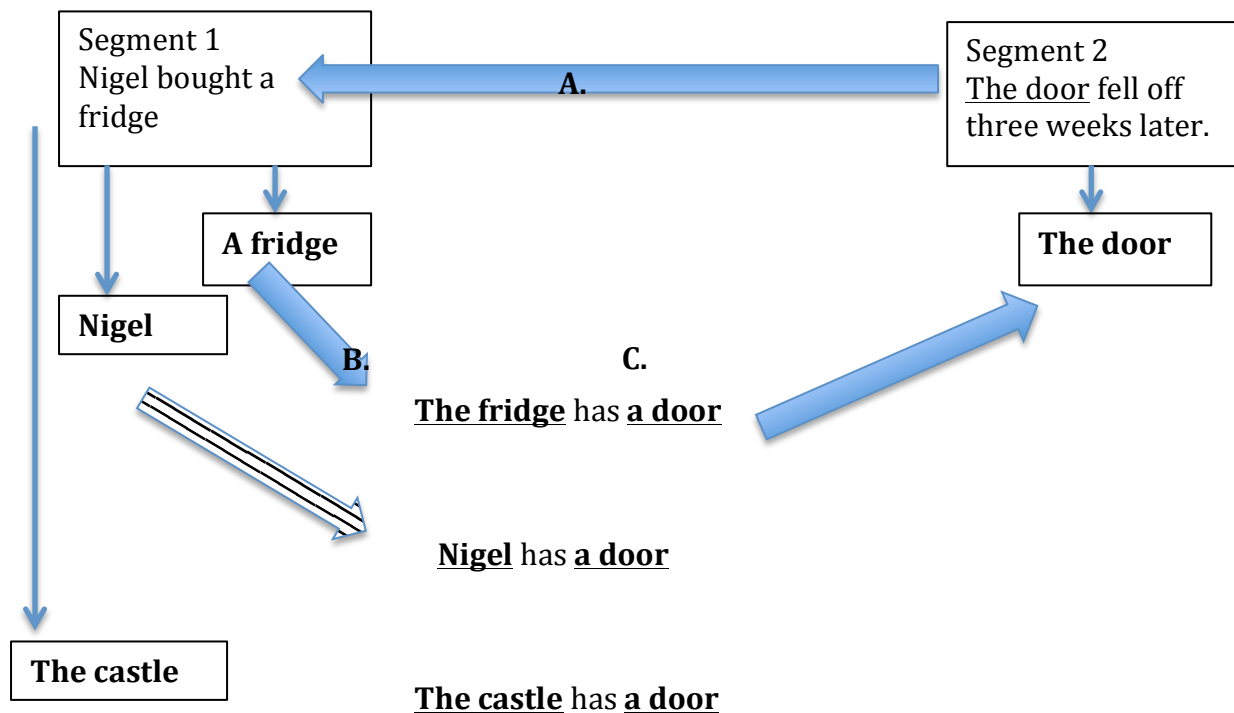
[8] The index is directly quoted from the original chapter, in this section, (2´) responses to (2´).

Example (25) is divided into two segments. It is only possible to interpret the referent of the door by virtue of a bridging assumption. The referring expression in segment (25) is *the door*, the definite article indicates a Uniquely Identifable referent. What (A) illustrates is that the addressee searches back in the immediately previous discourse for pragmatic inferences that can contribute to identify the referent. What (B) presents is the process of selecting the ideal bridging assumption. In this stage, the addressee tends to pick out a nominal that is relatively prominent for him, according to the relevance-theoretic comprehension procedure. In segment (1), there are two nominals, *a fridge* and *Nigel*. The processing efforts required for interpreting both nominals seem to be rather equal at first glance, with *a fridge* being slightly easier to be accessed due to its two-word-closer position to segment (2). However, the distance influenced by syntactic position is not the determining reason why ´a fridge´ requires less processing effort than ´Nigel´. Admittedly, when being presented with the concept ´a fridge´, the addressee gets easy access to the assumption that a fridge has a door/doors based on his knowledge of the world. On the other hand, the assumption allegedly inferred from ´Nigel´ is that Nigel has a door – which is not an assumption the hearer has already stored in mind. The hearer could have established such an assumption on the spot when a proper context is achieved, but doing so would require more processing effort than just retrieving it. Plus, such an unusual assumption as the person has a door would not be able to yield positive contextual effects for the addressee, since it is very unlikely to be a true description of the world. Comparing to the in-built assumption of ´fridge – a door´, accessing ´Nigel has a door´ is thus both more processing-effort demanding and yields less positive cognitive effects. As it has been pointed out previously, it is the degree of processing effort being the major measurement in this stage, therefore, the addressee will choose ´a fridge´ to link with. Now that the addressee has been offered an antecedent for the anaphor in the bridging assumption, he can build a mental representation of the door by far. What follows naturally is that he is able to incorporate the mental representation of the door into the proposition of the utterance in segment (2), as what is marked by (C).

The bridging assumption "The castle has a door" is, in fact, not a likely bridging assumption, because the nominal ´castle´ is not connected to any concept that is previously mentioned. It therefore requires so much processing effort to be accessed that it violates the assumption of not putting the addressee to unjustifiable processing

efforts. According to my understanding of relevance theory, the addressee is highly likely **not** to continue testing other bridging assumptions once he/she has achieved the most ideal one, that is to say, the addressee stops once he is satisfied with the comparatively most prominent bridging assumption that is also consistent with the principles of relevance, i.e. one that yields satisfactory amounts of positive cognitive effects. He would continue examining and testing only if the most accessible one is not relevant. Thus, in example (25), he/she will most probably not even consider a remote bridging assumption such as ´The castle has a door".

To summarize, ´The fridge has a door" is chosen, since it yields adequate positive contextual effects for the least amount of processing effort.

The bridging process being marked with index is presented below:

Nigel bought a fridge$_i$.
The fridge$_i$ has a door$_j$.
The door$_j$ fell off three weeks later.

Figure 2. The bridging process marked with index

According to Blakemore (1997), the addressee does not necessarily need to know the bridging assumption beforehand in order to interpret the utterance. Following this point, the addressee does not need to have previous access to the bridging assumption that the fridge has a door; it may be created on the spot. In the cases of (25) and (25´), the assumption arises mostly from what the addressee already knows about the world, i.e. a fridge must have an entry that we call a door for things to get through. Nevertheless, in other examples such as (26), Blakemore argues, the addressee does not need previous access to the bridging assumption.

(26).   I walked into the room. The chandeliers sparkled brightly.

The bridging assumption needed in (26) is that the room has chandeliers. This is less certain according to our knowledge of the world, since chandeliers are not something a room is obligatorily equipped with. The addressee will, however, still interpret the

chandeliers as those in the room that has been mentioned in the first segment of (26). The philosophy behind this is that the speaker automatically promises that an utterance is consistent with the principle of relevance once it is uttered. That is, the addressee is also aware of the fact that he/she does not need to spend unjustifiable processing effort to interpret the utterance. If there is an already-Activated room in the immediately previous discourse, a room that is possible to have chandeliers, based on his trust to the ´promise´, he would naturally interpret the room as the unique one that has the chandeliers the speaker saw.

In the Coding Manual (GHZ, 1993), Gundel et al proposed the criteria for judging each CS. For the CS Uniquely Identifable, one of the criteria they proposed is, " A unique referent can be created via a ´bridging inference´ by association with an already Activated referent. (e.g. A house….the front door)" (GHZ, 1993:4)). An example that they use to illustrate this, is (27).

(27).   She got into bed, laid her head on the pillow, and in two minutes was sleeping like a child. (From *Murder after Hours*, Agatha Christie)

The addressee will interpret the referent of the referring expression ´the pillow´ as the one that is put on the exact bed that just has been Activated. The bridging assumption is that the bed has a pillow. The pillow is in other words not just any random pillow, but the one being placed on the bed that is Activated in the addressee´s short-term memory. This is the interpretation that yields a unique pillow for the least processing effort and meanwhile results in a plausible description of the world.

3.4 Referential use Vs. Attributive use

As for definite descriptions, Keith S. Donnellan (1971) has proposed the argument that definite descriptions have two uses. The same definite description can be used attributively or referentially[9], depending on the speaker´s intention.

---

[9] The term *Referential* in Donnellan´s sense is not the same as the cognitive status Referential in the Givenness Hierarchy theory. In Givenness Hierarchy theory, the cognitive status Uniquely identifiable is assumed to be compatible with both Referential and attributive readings in Donnellan's sense, and Donnellan's distinction of the two uses of definite description is not reflected in the Givenness Hierarchy.

By attributive use, Donnellan (1971) means that "A speaker who uses a definite description attributively in an assertion states something about whoever and whatever is the so-and-so." (Donnellan, 1971: 198). The two essential words here are *whoever* and *whatever*, they indicate that the content of the so-and-so is irreplaceable and fixed, but the identification of the entity that fits in the description is not.

About Referential use, he states the following: "A speaker who uses a definite description referentially in assertion, on the other hand, uses the description to enable his audience to pick out whom or what he is talking about and states something about that person or thing." (Donnellan, 1971: 198). The term being used is *whom* and *what,* this indicates that what the speaker intends to communicate about, is a particular, fixed entity.

The distinctions between the two uses can be illustrated from several perspectives. I will use the formula " the X is Y " as a simple example to demonstrate the distinctions, where X stands for the definite description and Y stands for a name or some aspects of X.

Considering it from the speaker´s perspective, when he uses the utterance attributively, he intends to focus on the property X and make X irreplaceable. The definite description yields a referent X′, and there is only one X′ that uniquely fits the description of X. The content in Y describes something about this unique referent X′. In other words, Y is attached to X′. If X′ cannot be found or does not exist, Y is empty.

When the speaker uses the utterance referentially, on the other hand, he intends to use X as a device to draw the addressee´s attention to single out the particular entity he has in mind, name it X". The addressee focuses on X" and intends to let the addressee know that it is X" he is talking about. The speaker chose to use X as the device, mostly probably due to the consideration that X is the most ideal choice of assisting the identification, nevertheless, X is still replaceable. Any device that enables the addressee to single out X" and get through the idea that it is X" the speaker intends to continue to talk about, is fine.

To sum up, in the attributive use, it is presupposed that something or whatever other thing is X. In the Referential use, it is presupposed that only a particular entity is X.

Seeing it from the addressee´s perspective, when the expression is used attributively, he does not necessarily need to have knowledge about the referent previously, since he is enabled to construct such a referent with the encoded information. On the other hand, when the expression is used referentially, the addressee is expected to realize the intended entity. Therefore, I think, a certain degree of Familiarity about the referent is usually needed, if not, at least some shared knowledge between the two is required. But this is not always true. For example, when the addressee is told to "bring me the cup *on* the table" by the speaker, yet nothing is placed on the table. Instead, the addressee sees the cup *under* the table and picks that cup out. In Referential use, the addressee´s reaction indicates the success of the communicative act, but in this example, no previous Familiarity is involved.

As for the truth value of X in ´the X is Y`, the two uses present distinguished patterns. In the original example proposed by Donnellan (1971), i.e. "Smith´s murder is insane", a referent that uniquely fits the description is required when *Smith´s murder* is used attributively. If not, for example, if Smith actually dies naturally, such a referent cannot be found. As a result, no insanity can be assigned to any entity and it is thus impossible to assign a truth value; we get a truth-value gap. As I have mentioned previously, in my opinion, when the utterance is used attributively, Y is attached to X. Y does not exist when X does not.

This is most commonly not the case in Referential use, whereby the identification process may remain intact even when X is false. This is a result of what X functions as, namely as a device to assist the identification rather than a frame that the referent must flawlessly fit into so as to be chosen. This can be explained by example (28) below:

(28)

Speaker: – I´m thinking of the man with the mesmerizing blue eyes we met last night at Smith´s party.

Addressee: -- Yea, I remember him. He has charming eyes, but they are definitely not blue.

The addressee does not share all of the speaker´s presumptions about the referent of *the man with the mesmerizing blue eyes*, yet he is still able to pick out the intended person that the speaker assumes he would. Whether or not the addressee acknowledges the truth of the definite description does not disturb the identification process. In order to achieve the right identification, the addressee only needs to approve that the speaker chooses X as the device and follows his sense. He does not really need to accept X to be true; it only ought to be truthful for the speaker (i.e. when the speaker formulates the utterance, he reasonably believes that X is an ideal device to assist the addressee to pick out what is intended in that context).

To sum up, in the attributive use, if nothing is X, the speaker does not refer to anything, a truth-value gap exists. In the Referential use, if nothing is X, the speaker stills states something ad hoc (with a slightly misleading description) and asserts that the referent is Y. Both the speaker and the addressee are satisfied as long as Y is identified.

The Referential-attributive distinction plays a role during the data analysis when some examples only have heavy definite descriptions rather than a clear antecedent for the anaphor. The addressee needs to construct the referent through inference by interpreting the expression attributively.

In a footnote, GHZ present their opinion about the distinction between the CS Referential and the term Referential use in Donnellan (1966)´s sense.   They acknowledge that indefinites may be used referentially or nonreferentially, but definite expressions are **always** used referentially in their sense. The distinction between the two uses claimed by Donnellan is not reflected in Givenness Hierarchy theory. GHZ state: "[…] definite expressions are always used referentially in the sense that the speaker intend to refer to a particular entity in using them – either one they are acquainted with and intend to refer to irrespective of whether the description actually fits (Donnellan´s ´Referential´ use), or one which the description actually fits,

irrespective of whether the speaker is directly acquainted with it (Donnellan´s ´attributive´ use)." (GHZ, 1993: 276).

3.5 How reference is understood

So, what does the process of a successful communication concerning coding and decoding of the *form-cognitive status* correlation look like? Generalizing from GHZ´s theoretical accounts, I assume that the ideal process of this kind can be depicted as the procedure routine in Figure 3:

Figure 3. How is reference understood

I. Utterance composing stage (by the speaker)

a. The speaker intends to refer to a particular object, individual, proposition or a set of individuals. The intended referent is chosen;

b. The speaker assumes/calculates the mental location of the intended referent/interpretation in the addressee´s memory and attention state with considerations on the addressee´s knowledge about the referent, the referent´s degree of Familiarity to the addressee, the context where the referring expression is posited etc. The assumed cognitive status is decided;

c. The speaker selects a specific referring expression (i.e. a ´cipher´) whose form matches the assumed cognitive status according to his/her unconscious knowledge about the language;

d. The speaker utters the 'cipher'

II. Utterance interpretation stage (by the addressee)

a. The addressee catches the 'cipher' from the speaker's utterance. The referring expression/definite description is recognized;

b. The addressee starts decoding by first accessing the information encoded in the referring expression itself. He searches his brain database for the conditions of using the referring form (which was already introduced and memorized during language acquisition). The set of possible referents is restricted according to the encoded cognitive status. Meanwhile, he also

recalls the cultural/encyclopedic knowledge or shared personal experience with the speaker to evaluate the degree of Familiarity;

c.   The addressee decides the referent based on the understanding of step b, contextual assumptions and other pragmatic inferences. The referent is identified.

Chapter 4. Previous Claims about *na*

4.1 Hedberg (1996)

The purpose of this article, as the author states, is to show that the Givenness Hierarchy framework can shed new light on the claims concerning the encoding of definiteness in Mandarin proposed by Li and Thompson (1976, 1981). The article discuss the behaviours of both the distal demonstrative *neì*[10] and the proximal demonstrative *zhè*, I will only focus on the arguments about neì in this review. Li and Thompson´s (1989) well-known claim about the Chinese distal demonstrative *na* is that it is beginning to function like the English definite article *the*, and this is what Hedberg (1996) argued for.

GHZ (1993) constructed the Givenness Hierarchy for Chinese as in (3) below:

(3). THE GIVENNESS HIERARCHY: MANDARIN

In Focus	> Activated >	Familiar	> Uniquely Identifable >	Referential	>	type
identifiabl {	Ø {TA		{***nei* N**}			{*yi*
N ´a N´						
ta ´s/he; it}	*zhe* PROX					Ø N }
	***nei* DISTAL**					
	zhe N}					

As (3) illustrates, GHZ (1993) claim that Mandarin does not have linguistic forms to signal that a referent is Familiar or Referential. Regarding on the correlation between *nei* and a specific CS, they claim that the distal demonstrative *nei* requires only activation when it is used alone, and it encodes Uniquely Identifiable when in the form *nei N*.

According to Hedberg (1996), the claim that *nei* is beginning to function equivalent to a definite article in English, (i.e. *the*) is not easy to be explained in the Givenness Hierarchy due to the logical relation between the CSs. The reason is that *nei* is initially a distal demonstrative whose English counterpart is *that*, which signals

---

[10] The pronunciation being recorded in the article is nei, I chose to follow the author´s intention in this chapter of literature review.

Familiar in the hierarchy for English. According to GHZ (1993), the six CSs on the hierarchy are implicationally related, that is, if the referent is Familiar to the addressee, it is automatically also Uniquely Identifable to him. Thus, it is unclear if *nei* is supposed to be more precisely translated as *the* in a case where it is actually translated as *that*, since using the latter acknowledges using of the former.

A more promising example given by Li and Thompson (1975) is quoted in (4), whereby *nei* occurs in a relative clause.

(4). Wo diu   le   de nei ben shu
      I   lose ASP ATT that CLS book
     "the book I lost"

Hedberg (1996) argues that the addressee is still required to know the fact that the speaker lost a pen previously to succeed in retrieving the referent. " […] the information encoded in the nominal alone is still insufficient to enable the addressee to uniquely identify the referent without previous Familiarity." (Hedberg, 1996: 185). Once the condition of Familiarity is met, *that* can be used as well. Therefore, it is still unclear if *the* is the only counterpart of *nei* in this example.

The ideal example for *nei* signalling Uniquely Identifable prosed by GHZ (1993) is (5), whereby the referent is only Uniquely Identifable to the addressee.

(5) 我 昨晚   睡 不 着。   隔壁 那 只 狗 让 我 一直 醒 着。
    wo zuowan shui bu zhao   gebi nei zhi gou rang wo yizhi xing zhe
     I last night sleep NEG PART next door that CL dog make me always awake PART
    " I cound´t sleep last night. The dog next door kept me awake."

The addressee does not necessarily need to know that the speaker´s neighbour has a dog previously for the appropriate use of *nei*. The information encoded in the nominal is adequate enough for the addressee to uniquely construct the representation, therefore the referent does not need to be Familiar to the addressee.

Hence, Hedberg(1996) agrees with the observation proposed by Li and Thompson (1989) that *nei* has one of the functions of the definite article property.

4.2 Brøseth & Jin (2008)

While GHZ (1993) proposed the correlation between each of the cognitive status and their corresponding linguistic forms in five languages, Heidi Brøseth and Fufen Jin (Heidi & Jin, 2008) focus on one particular CS Referential and its relevant referring forms in Chinese. The study was conducted on a comparative base with analysis of examples from Chinese, Norwegian and English.

Chinese does not have an article system such as the one of English, that is, indefinite article and definite article do not exist in Chinese. However, as Brøseth & Jin state: "[…] several authors (Li and Thompson 1981; Gundel et al. 1993; and Robertson 2000) maintain that the distal demonstrative determiner in Chinese "*nei* CL N" is *beginning* to take on some functions associated with the definite article." (Brøseth & Jin, 2008:115). A similar phenomenon has been observed in Norwegian where the pre-positioned definite articles *den* (´the´) is homophonous with the distal demonstrative determiner *den* ´that´. Based on the assumption proposed by previous linguists and the pattern of Norwegian homophonous articles, Brøseth & Jin proposed the hypothesis that Chinese at the current stage of development has two homophonous linguistic forms, *nei* the emerging definite article and *nei* the distal demonstrative determiner, and that they signal two different cognitive status depending on the context. Nevertheless, although the tendency that *na* starts to functions equivalently to a definite article is observed, it is not tested and exemplified further in their article.

The novelty of the proposal in Brøseth & Jin´s article is that both Chinese and Norwegian associate the cognitive status Referential with a particular linguistic form. The Chinese part of this argument contradicts the claim made by GHZ (1993), as the latter acknowledge no Chinese linguistic signalling Referential. Conversely, as Brøseth & Jin (2008) claim, some examples of their Chinese data indicate that the referent of the form ´nei CL N´ is only Referential. However, this does not lead them to conclude that the emerging definite article *nei* has the cognitive status Referential as its necessary and sufficient condition. Rather, they propose a new approach: In addition to the distal demonstrative determiner *nei* and the emerging definite article

nei that already have been observed, Chinese has an indefinite article *nei* as well. It is this third *nei* that can signal Referential.

Cconcerning the emerging definite article *nei*, Brøseth & Jin (2008) hypothesize that *nei* CL N in Chinese will continue to have a demonstrative function while developing into a definite article, and that "the demonstrative determiner nei N ´that N´ could have a higher cognitive status than the emerging definite article nei N ´the N´, which has Uniquely identifiable as its necessary and sufficient condition." (Brøseth & Jin, 2008:121). This hypothesis has not been investigated in the article, but they suggest that the necessary condition for the demonstrative *na* should be at least Familiar, following the assumptions in the Givenness Hierarchy theory.

To summarize, Brøseth& Jin (2008) propose that Chinese has three homophonous linguistic forms of *nei* that signal different cognitive statuses. The emerging definite article *nei* N ´the N´ signals Uniquely Identifiable. The demonstrative determiner *nei* N ´that N´ is suggested and supposed to have at least Familiar as its necessary condition. Last but not least, the indefinite article *nei* of *nei* CL N signals Referential. Therefore, Brøseth& Jin (2008) claim that Chinese does have a linguistic form whose necessary and sufficient condition is Referential.

In Chinese, a nominal is commonly required to be proceeded by a specific classifier. For example, the classifier 棵(ke) only precedes nominals denoting trees. Therefore, it is possible to use *na CL* (without the nominal) solely as a referring expression, since the classifier encodes information about the referent´s ontological category. It is likely that classifiers play a role in the reference assigning process by narrowing down the referent set.

Both *nei N* and *nei CL N* are possible referring forms in Chinese, but the latter is more frequently and formally used. At the beginning of the conclusion part in Brøseth & Jin (2008)´s article, they state:" In this paper, we questioned the assertion of Gundel et al (1993) that nei CL N has the cognitive status Uniquely identifiable as its necessary and sufficient condition". A closer examination of the GHZ (1993) article shows that such an assertion has not been made. More precisely, what GHZ (1993) concludes is

that nei N (**without CL**) has Uniquely identifiable as its necessary and sufficient condition. Although the examples presented in GHZ (1993)´s paper seem to indicate that by ´nei N´, they actually mean ´nei CL N´ or even more, but this is not clearly underspecified. Considering the special role of classifiers, I choose to distinguish *nei CL N* and *nei N* as two referring expression forms in my data analysis.

Lastly, the pronunciation of the form 那 should be clarified, since the notion of ´homophonous forms´ has been mentioned in the Brøseth & Jin (2008) article. The citation pronunciation of 那 in the Chinese dictionary is nà, it is also explicitly stated that 那 is read as its citation pronunciation in structures of *na* and *na* N; on the other hand, it should be read as neì when it is followed by whichever classifier or the cluster of a number and a classifier. In other words, *na* N, *nei* CL N are read differently.

Chapter 5. Methodology & Data collection

5.0 Introduction

This section is meant to outline the framework of the investigation being conducted in this thesis. I will first explain the motivation behind such an investigation and the basic premises on which the investigation based on, and then describe the data collection process with considerations of the motivation and intended goals. I will also illustrate the way in which the data were prepared with necessary screening and how the data were sorted. On the practical level, the process through which each datum was analysed and reviewed will be explicitly demonstrated by a sample.

5.1 Preliminaries

The main premise of the Givenness Hierarchy theory is that " different determiners and pronominal forms conventionally signal different cognitive statuses (information about location in memory and attention state), thereby enabling the addressee to restrict the set of possible referents. " (GHZ, 1993: 275). Similar to the choice of using other linguistic elements such as intonation, it is believed that the form of referring expression is determined by the cognitive status of the referred entity in the addressee´s memory.

The logical relation between the six universal cognitive statuses on the hierarchy is implicational, as GHZ claimed. That is, when a referent has the CS Familiar, it is automatically also Uniquely Identifable, Referential and Type Identifiable to the addressee. Accordingly, the definite article ´the´ with Uniquely Identifable as its sufficient and necessary CS can also be used for all the CSs lower than Uniquely Identifable. On the other hand, however, when an entity is referred by an expression with ´the´, it does not necessarily have the CS Uniquely Identifable. Using ´the´ only means that the lowest possible CS of the referent is Uniquely Identifable, it is possible for the referent to be Familiar, Activated or In Focus to the addressee. Thus, it is merely ideal that there exists an exclusive mapping between a certain CS and its corresponding referring form in the distribution.

5.2 Motivation

One of the major goals of this thesis is to study the Mandarin word *na*, on what it encodes. Naturally, the main motivation of the investigation is to get empirical evidences and insights regarding the distribution and meaning of *na* and noun phrases embedding *na*.

Following this sense, the investigation also attempts to capture the tendencies for the CSs indicated by the various noun phrase structures embedding *na* (if such tendencies exist), as complements to the major issue.

Moreover, this investigation is expected to serve as a test for the argument that Mandarin does not have a linguistic form encoding the CS Referential, proposed by GHZ (1993) as part of the results of an investigation testing predications about the relation between linguistic forms and various CSs. An investigation of the same purpose as mine has been conducted by Brøseth & Jin (2008), whose results contradict those of GHZ (1993)´s. My investigation will offer new perspective to this issue: Focusing only on the understanding of *na*, it is thus more specific comparing to the one conducted by GHZ (1993), where a variety of Mandarin pronominals and determiners were studied; Meanwhile, I take some other pragmatic features into account, as these features will inevitably influence the reference assignment.

Last but not the least, it is also an important motivation of the investigation to deepen my understanding about the Givenness Hierarchy theory by applying it to my data analysis, with special regards on the understanding of the CSs Uniquely Identifable and Referential. Analysing the data often involves other necessary approaches in addition to that of the Givenness Hierarchy theory, Relevance Theory is one of them. The investigation is therefore a study of how the Givenness Hierarchy theory interacts with other theories in the communication level.

5.3 Data collection and preparation

5.3.0 Considerations

My investigation involves data from two source books, namely the Chinese translated version of the novel *Sophie´s World* and a novel written by a native *guó yǔ* speaker.

Novels are the first genre I consider for the data collection due to several reasons. Firstly, they are equipped with continuous contexts, and this makes the annotation more reliable, especially in examples where the context plays a heavy role in the interpretation.

Secondly, published books have gone through several revises, thus they contain relatively fewer technical problems such as unnecessary grammatical mistakes. Since unsatisfying data will be reviewed with particularly cautious discussion and may serve as opponent evidences for the concluded tendency, it is important to ensure that they are not irregular due to misuse of the grammar.

New vocabulary, temporary usage and random occurrences emerge constantly in the daily use of colloquial Chinese, yet not all of them are recognised and will eventually be written into the grammar. Since the use of written language is relatively more conservative and formal, it has advantages in reflecting the authentic grammar in this sense.

5.3.1 The two source books

5.3.1.1 Source book 1: The Chinese translated version of Sofies verden
Sofies verden (Sophie´s World) is a novel introducing the history of philosophy written by the Norwegian writer Jostein Gaarder. It follows the story between the heroine Sophie Amundsen, a teenager Norwegian girl and a mid-aged philosopher Alberto Knox, who introduced the philosophical thinking and the history of philosophy to Sophie. Originally written in Norwegian, the book has been translated to around 53 languages, including English and Mandarin Chinese.

The Chinese version I got was purchased in a local bookstore in main land China where *pǔ tōng huà* is the language spoken by the readers. The book was translated by

Baosen Xiao, a native *guó yǔ* speaker from Taiwan. I, as a native *pǔ tōng huà* speaker did not notice the fact that the book was written in *guó yǔ* while reading, this is mostly due to the Familiarity between the written languages of *pǔ tōng huà* and *guó yǔ* except for some certain vocabulary which did not appear during my reading. The book did not come with preface, nor did it have translator´s notes explaining basic information regarding the translation[11]. Therefore, there lacks confirmed answers to some important questions such as ´Which language was the book translated from, directly from Norwegian or via English?" "If the book was translated from the English version, for example, which edition?"

Nonetheless, it is at least clear that the first source book is a piece of translated work from a certain original copy. As translated works are done under constraints from the original language, the use of the language is usually rather conservative. It is also inevitable to be influenced by the original language in aspects such as word order, sentence punctuation, etc.

5.3.1.2 Notes of A Desolated Man

The second source book is a postmodern, first-person tale of a contemporary Taiwanese gay man reflecting on his life and love on the edge of the mainstream Taiwanese society. The narrator/ speaker[12], Xiao Shao, recollects a series of friends and lovers, as he watches his childhood friend, A´Yao, succumbing to complications from AIDS.

The book was written by a female native *guó yǔ* speaker in a poetic, complicated yet succinct language style, its language almost approaches the point between vernacular Chinese and literary Chinese[13], which is quite concise. Compare to the straight and

---

[11] I have also checked the various editions of the book, as well as searching for information on the Internet.

[12] In spite of the actual gender of the writer (a female writer), I will refer to the speaker by ´he´ since it is this male figure that is the narrator of the book.

[13] Literary Chinese: (文言文, wényánwén) is the form of written Chinese used from the end of the Han Dynasty (206 BC – 220 AD) to the early 20th century when it was replaced by vernacular written Chinese. The language has different grammar, vocaburary etc than those of modern Chinese. The language is no longer in use.

narrow language in the first source book, language of the second one is so considerably more unrestrained and literary that almost all the examples are consisted of complex sentences with more than commonly frequent punctuation. The latter will directly influence coding for the CSs, since the criteria listed in the Coding Manual (my principle guidelines for the coding) are partly defined by the punctuations. Nevertheless, I believe that such a free language style is positive for the data variety in the sense that it unveils some usages that are unusual yet acceptable for the grammar.

5.3.2 Collecting and preparing the data

The procedures of data collection and preparation of the two source books are similar, I will use Notes of A Desolated Man as an example to illustrate the process.

Initially, all the sentences containing *na* were collected from the electronic version of the source book by using the searching function of Microsoft Word, this guarantees that all the occurrences of *na* are trapped without omissions resulting from human errors. A total number of 316 matches were found as a result of this step.

The word *na* is cited as being corresponding to a distal demonstrative and a modal particle in most of the Mandarin dictionaries, where our understanding of *na* departs from. The second step is then to manually eliminate the data that are irrelevant for the study of referent assignment, such data include sentences where *na* functions as a modal particle or part of a discourse marker, and sentences where *na* occurs as component of some certain nouns, e.g 刹那[14] chà *nà* (a short moment).

The remaining data are assumed to be relevant for referent assignment. *Na* functions similarly to a pronominal when it is used alone. On the other hand, it is widely recognised that the distal demonstrative *na* has started to function like a definite article in Mandarin, as the language lacks an article system. Examples where *na´*s

---

[14] Na is usually not used as component of nouns in Mandarin. Some exceptions: It appear in proper names as a family name; It is used as component in a transliterated noun such as 刹那(chà *nà*, ´a brief moment`), a transliteration from Sanskrit, being introduced to Chinese via Buddhist scriptures.

usage approaches a pronominal are categorised separately, and the rest of the data are grouped according to the noun phrase structure where *na* occurs, for example, Group.*na* CL N, Group.*na* N, etc.

5.4 Coding

5.4.1 The Coding Manual (2006: GHZ)

The Coding Manual (Coding Protocol for Statuses on the Givenness Hierarchy) offers guidelines to the annotators who code cognitive status within the Givenness Hierarchy theoretical framework. It is a four-page instruction where criteria for each CS are explicitly listed with regard to more concrete linguistic properties such as syntactic and semantic performances of the referring expression. For instance, one of the criteria for the CS In Focus is: The referent is In Focus if it is the referent of a DP earlier in the same sentence, which is more strictly and explicitly defined comparing to the abstract description of its nature. The Coding Manual (2006) makes it possible for consequent coding throughout the investgation, it is especially practical for comparing among the results of several similar projects. One of the motivations of my investigation is to test the arguments about some aspects of Mandarin claimed by previous linguists, therefore, I will strictly refer to the Coding Manual during the annotation to ensure the results comparable to those of the previous investigations.

5.4.2 The coding process

At the beginning of the Coding Manual (GHZ, 2006), GHZ illustrate the introductions about how to use the protocol. The instruction is stated as such: "When determining cognitive status using the protocol, imagine you are the speaker/writer and ask yourself what you can assume about the cognitive status of the intended referent for the addressee at the point just before the form is encountered. Check the criteria for each status in the order[15] they are listed below. That is, start with the cognitive status IN FOCUS. If none of the criteria apply, try ACTIVATED. If none of the criteria apply, try FAMILIAR, and so on. Stop when you find a criterion that applies. This is

---

[15] The six cognitive statuses are listed from the highest (i.e. In Focus) to the lowest (i.e. Type Identifiable)

the highest cognitive status for the referent/interpretation you are checking." (Coding Manual, GHZ: 2006). I will adopt this procedure during my data analysis.

5.4.3 A coding sample

The following example demonstrates how the coding was conducted for each datum. Text in the example is provided with word-by-word gloss together with pin yin in the second layer and free translation in English in the end of the example. The relevant noun phrase whose referent is coded for CS is marked in bold in both the pin yin layer and the gloss layer.

(29).

诡辩学家　　与　自然派　哲学家　　有　　一　个　　共　通　点 (sg1)

gui bian xue jia yu   zi ran pai zhe xue jia   you   yi ge gong tong dian
Sophists   and   natural philosopher have   one CL common point

那　　就　是：他们　都　批评　　传统的　　　神话
na jiu shi   tamen dou piping   chuantongde   shenhua
that just be   they both criticize traditional   mythology

"The Sophists and the natural philosophers have one characteristics in common, that is, they were critical of the traditional mythology."

In (1), the referring expression *na* refers to *yi ge gongtongdian* ('one characteristic in common'), which occurs in the direct object position of segment 1. The referent meets one of the subcase for the status In Focus according to the Coding Manual: it is the syntactic focus of the immediately preceding main clause, which makes it qualified for the criterion: it is the referent of a DP in a syntactically prominent position (incl. non-overt subjects) in the main clause of the immediately preceding sentence. Therefore, the referent has the CS In Focus.

51

Chapter 6. The Investigation

6.1 Group Na.

The referring expression in this group is *na*. It is presumed that *na* functions similar to a pronoun rather than a determiner in this group. The data are grouped into two subcategories according to their sources.

6.1.1 Examples from Source book 1

The group includes 14 examples with the relevant referring expression *na*. 7 of these 14 referring expressions have In Focus referents, occupying 50% of the data set. Expressions whose referents are Activated and Familiar occur twice each. There are also 3 data whose referents are Uniquely Identifiable.

6.1.1.1 Data analysis

6.1.1.1.1 The relation between non-reflexive predicative NP and its subject in copular sentence

While analysing the data of this group, one of the most frequently encountered issues involves the interpretation of copular sentence, and I found it utterly necessary to clarify the relation between (non-reflexive) predicative NPs and subjects in copular sentences before we go about the analysis itself. I will use Example (30) to demonstrate my arguments.

(30)

另外　也　有　某　些　东西　随　　　着　烟雾　往上　　升, (SG1)
lingwai ye you mou xie dongxi sui   zhe yanwu wang shang sheng
besides also have certain PL thing   along with PART fog   upwards   rise

那　是"气"(SG2)
na shi qi
that be   air

"Besides, there is something rising upwards along with the fog, that is ´air´. "

The referent of *na* is *mou xie dong xi ("something")*, which is in the syntactic focus of an existential sentence, namely segment 1. Segment 2 is a positive copular sentence with the following structure: the pre-verbal *na* precedes the copula *shi* that has a

strong identifying reading, which is followed by the post-copular nominal *water*. Concerning the relation between *na* and the post-copular nominal *water*, it is not uncommon to argue that *na* is coreferential with *water* [16] among linguists. Nonetheless, this is not always the case that subjects are coreferential with the post-copular nominals in positive sentences. If we mark the referent of *na* as X, and the referent of the post-copular nominal water as Y (i.e. water), what the copula *shi* does is to signal that *X* and *Y* coincide by virtue of its identifying reading. That is to say, the entity referred by *na* coincides with water, but *na* per se does not refers to water (otherwise the interpretation of the copular sentence will become ´water is water´ when substituting *na* with its referent, which does not yield positive cognitive effects to the addressee.) To conclude, *na* does not refer cataphorically in examples of this kind, and the relation between non-reflexive predicative NP and its subject in copular sentences is NOT coreferential.

6.1.1.1.2 A Uniquely Identifable referent retrieved by attributively interpreting the definite description

With regard to the interpretation of definite descriptions, Donnellan (1971) claims that they can be used referentially and attributively. [17]In the GH framework, referents with both readings are possible to have the CS Uniquely Identifable. Example (32) demonstrates the scenario where a Uniquely Identifable referent is retrieved by attributively reading the definite description.

(31):

<div align="center">

木材燃烧时发出´噼啪！噼啪｀的声音，　那是水。
segment 1　　　　　　　　　　　　　　　　segment 2

mucai ranshao shi fachu pipa pipa　　de　shenyin *na* shi shu
Wood　burn make out Onomatopoeia　ASSOC noise that be water

</div>

---

[16] In fact, it is not unusual to mark predicative NPs in a positive phrase like example (2) as being coreferential with their subjects. (e.g. Mitkov, 2002).

[17] For presentation of the distinction between attributive use and Referential use of definite description porposed by Donnellan (1971), please refer to the relevant section in the litterature review part.

"The wood makes crackling noise while it is being burned, that is water."

The referring expression *na* occurs in segment 2 of (31), and segment 1 narrates the fact that the woods make crackling noise while they are being burned. Segment 2 is copular phrase with the copular *shi* (´is´). As has been discussed, the copula only signals that the referent of *na* and the referent of the predicative NP, i.e. *water* coincide. The post-copular NP is only one part of the predication that is attached to the subject referent. Identifying the referent of *na* will not be disturbed when *water* is substituted because the real referent of *na* is connected with the preceding segment backwards in the context.

The thing is, *na* does not have an overtly-mentioned referent in the previous discourse, instead, the addressee gets a definite description. According to Donnellan (1966), a definite expression can have an attributive reading or a Referential reading, depending on the speaker´s intention. He states: "A speaker who uses a definite description attributively in an assertion states something about whoever and whatever is the so-and-so. A speaker who uses a definite description referentially in an assertion, on the other hand, uses the description to enable his audience to pick out whom or what he is talking about and states something about that person or thing." (Donnellan, 1966:198). In the first case, each detail of the descriptions counts during the identifying process, since the speaker states something about a whatsoever thing that fits the description. Thus, the attributives are essential. In the latter use, i.e. when the definite expression is used referentially (not in the same sense as what it means by the cognitive status ´Referential´ in the Givenness Hierarchy Theory), the fact of drawing the addressee´s attention to single out a referent is more essential. The picked-out referent does not necessarily need to completely fit the description.

Back to (31), following the principle of optimal relevance, the referent of *na* shall be interpreted as ´whatever entity that causes the crackling noise´ so that the interpretation can yield enough positive cognitive effects. Such a referent is not overtly stated, it needs to be inferred from segment 1 based on the attributive reading of the definite description.

The referent of *na* has the CS Uniquely Identifable, it is partly retrieved through contextual inference with the already Activated proposition of segment 1, whereas Uniquely Identifable referents can be either Referential or attributive in Donellan´s sense, it is identified by attributively interpreting the definite description in this example.

6.1.1.1.3 A higher-level topic referent

(32).

哲学家　　走 到 那 两 人　　旁边，　取 下 他的 扁　帽，
zhexuejia　　zou dao *na* liang ren　　pangbian　qu xia tade bian mao
philosopher　go to that two person　aside　　take off　his　beret

说 了 一些 苏菲 听 不　懂　　的 话。 (sentence 1)
shuo le yixie sufei ting bu dong de hua
say ASP some Sophie hear NEG understand ASSOC word

苏菲　想，那 一定 是 希腊文。(sentence 2)
Sufei xiang *na* yiding shi xilawen
Sophie think that must be Greek
Sophie thinks, that must be Greek"

" the philosopher goes to those two guys, taking off his beret and said some words that Sophie does not understand.

Na in the copular phrase (in sentence 2) refers to the language that is associated with "Some words that Sophie does not understand" from the preceding context, but this does not mean that *na* is directly coreferential with the post-copular *Greek* in sentence 2. Although it seems unlikely, the post-copular DP *Greek* is merely part of a predication that holds of the subject referent. To figure out the referent of *na*, we need to investigate the immediate preceding context.

No explicit antecedent of *na* has been mentioned in the previous discourse. We do have a suspected candidate in sentence 1, i.e. 一些苏菲听不懂的话 ("some words that Sophie does not understand"), but this expression is not the genuine antecedent of *na,* considering the function of the copula. Now that we have reached the agreement that referents of referring expressions by the two sides of the copular coincide, in this context, it means that the referent of *na* coincidences with the referent of *Greek*, which is a singular entity (i.e. the Greek language). Following this sense, the entity it

coincides with is expected to be singular as well, therefore, the plural candidate "some words that Sophie does not understand" does not fit in this equation. A more satisfying solution would be: *na* refers to the name of whichever language that is related with the expression "some words that Sophie does not understand". Its antecedent can be inferred from the definite description, but is not the expression per se.

The referent of *na* is Uniquely Identifable as it can be exclusively retrieved through contextual inferences and by interpreting the definite description attributively. Nevertheless, one last problem must be tackled before we can come to finally conclude that the referent is Uniquely Identifable is: Is the language name Greek a higher-level topic of any random bunch of Greek words. If so, the referent will meet one of the criteria proposed for In Focus and thus have a much higher cognitive status. The relevant criterion is listed in (34):

(33).   A referent is In Focus if it is a higher level topic that is part of the interpretation of the preceding clause (whether it is overtly mentioned there or not). For example:

(34).   The kitchen has a new countertops and a beautiful tile floor. There's also a big walk-through closet. Would you like to take a look at <u>it</u>? Both the kitchen and the closet are In Focus. (the kitchen is the higher-level topic).

As (34) illustrates, the kitchen has been overtly mentioned in the preceding clause, which also introduces things that are placed in the kitchen. The walk-through closet is clearly In Focus (by the pronoun *it*), so is the kitchen, as the in-focus closet is located inside the kitchen, which is regarded as a higher-level topic. This reminds me of the ´bridging inference´ that is proposed as a criterion for the CS Uniquely Identifable. In the latter case, a referent can be uniquely identified via ´bridging inference´ by associating with an already-Activated entity, e.g. the room – the window. One can argue that a bridging inference exists between the language name Greek and Greek words, and the referent is then Uniquely Identifable. One can also argue that the language name is a higher-level topic of any random bunch of Greek words, the referent is then In Focus. One can even argue that the language name can be inferred

from the words under this name, and does not need to be overtly mentioned. In this case, the referent is uniquely mentioned.

According to (33), it does not play any role whether the higher-level topic has been overtly mentioned before or not, this can be used to account for the fact that Greek, as the name of the language, has not been mentioned.

I am personally uncertain about this example, but I will regard the referent as a higher level topic for the moment and regard it as In Focus in the end.

6.1.1.1.4 An Activated datum

(35) ”我　没　有　男　朋友.”　苏菲　啜泣　说，(SG1)
　　wo　mei　you　nan　pengyou　sufei　chuoqi　shuo
　　 I　NEG　have　male　friend　　Sophie　sob　　say
　 ” I don´t have a boyfriend.” Sophie said while sobbing, ”

　 ”那 是 我 编　 的，(SG2)
　 na shi wo bian　　de
　 that be　I　fabricate PART
　 ” that is what I made up with ”

　 因为　　那时候 我 说 白　 兔　 的 事情 让　你 不 高兴.”(SG3)
　 yinwei　*na* shihou wo shuo bai　tu　de　shiqing rang　ni　bu gaoxing
　 because that time　I　say white rabbit PART thing　make you not happy
　 ”because I´ve made you unhappy that time when I talked about the white rabbit”

Example (35) is the utterances of Sophie (the speaker), as the answer to her mother (the addressee)´s question of whether she has a boyfriend or not. The referring expression *na* occurs in segment 2. *Na* is controlled by the verb *bian* (´fabricate´) in the same segment, considering the context and the principle of optimal relevance, the entity that *na* stands for should be a fact (not a statement, as one can make up a fact that is connected with a certain statement, but not the statement itself).

A statement can be directly derived from the speaker´s utterence, the statement is ´I have no boyfriend´. The fact associating with this statement is the fact that the speaker does not have a boyfriend, which seems to be the referent of *na*. As a result, the overall interpretation of the sentence is: The speaker fabricates the fact that she does not have a boyfriend. Unfortunately, this interpretation is wrong. If we wait until the

sentence has been fully processed, it turns out that the referent of *na* is the opposite, it is the underlying fact being denied by the verb *bian* (´make up, fake´). That is, the real referent is the fact that the speaker has a boyfriend. This judgment is done due to the information being encoded in segment 3 in the subsequent context, which explains Sophie´s motivation of lying about having a boyfriend (Because she needs the excuse of having a boyfriend to cover the philosopher who makes her talking oddly, since talking about philosophy annoys her mother.)

In this example, the referent of *na* is retrieved mostly through pragmatic reasoning and contextual inferences only when the sentence that contains it has been fully processed. The first judgment we made without accessing information in segment 3 has been proved to be wrong. The information encoded in the referring expression does not really contribute much to the referent assignment as it is really breif. The referent meets one of the criteria proposed for Activated. The criterion is: The referent is Activated if it is a proposition, fact, or speech act associated with the eventuality (event or state) denoted by the immediately preceding sentence(s). The referent is thus Activated, only to note that correctly retrieving it requires the sentence to be fully accessed.

6.1.1.1.5. A Uniquely Identifiable datum

(36).

所谓　　哲学，　我们　指　　的　　是 耶稣 基督　　降生　　　前　 六　 百　　　 年　左右，
suowei zhexue　　women zhi　de　　shi yesu jidu jiangsheng qian　liu bai　　　nian zuoyou
so-called philosophy we　　 refer PART be Jusus Christ born　　　 before six hundred year around

在 希腊　演进　 的　 一　 种　　　崭新 的　思考 方式。(sentence 1)
zai xila　yanjin　 de　 yi zhong zhanxin de　sikao fangshi
in Geece develope ASSOC one type　 new　 ADJ think　way

在　　那　以前，人们　在　　各种　 宗教　　 中
zai　 *na*　yiqian renmen zai　 gezhong zongjiao zhong
PREP **that** before people PREP various　 religion　inside

找　　到　 了 他们　心　 中　 问题　　的　　答案。(sentence 2)
find reach ASP they　heart inside question ASSOC answer
zhao dao　le tamen xin　 zhong wenti　 de　 da´an

" By philosophy we mean a brand new way of thinking that emerged around 600 B.C in Greece. Before that, people had sought for the answers of the questions of their hearts in various religions."

Example (36) contains two sentences, the referring expression *na* occurs in an

inserted position in a circumpositional phrase in sentence (1). It is preceded by a preposition and followed by a constituent that is likely to be a postposition. Judging by both sentence (1) and sentence (2), the constituent phrase of *na* seems to be the time when philosophy emerged or the fact that philosophy emerges, if we were about to make a phrase to replace *na*. However, *na* comes in this particular circumpositional phrase with prepositions that require the referent of *na* to be a time concept, thus the possibility of *na* referring to the fact of philosophy´s emergence is excluded.

Notice here that the time point 600 B.C. has been mentioned in sentence 1, yet the time described by *when philosophy emerged around 600 B.C* differentiates from what is described by *600 B.C*, as the latter holds the possibility for time slots in which events other than the emergence of philosophy can also possibly take place. The referent of *na* is not 600 B.C, because the content encoded in the subsequent phrase in the same sentence requires the time of philosophy´s emergence to be emphasized in order to make the existence of *na* relevant.

There is no NP antecedent referring to the time when philosophy emerged in the previous discourse (sentence 1). The addressee is somehow able to retrieve the referent of *na* despite the fact that sentence 1 merely explains what is philosophy by describing how and when it emerges. The interpretation is conducted through inferences of sentence 1.

The referent has the CS Uniquely Identifiable, and it is uniquely identified through contextual inferences when sentence 2 that contains it has been fully processed. The constraint on the referent´s ontological category made by the prepositions also contributes to the retrieval.

6.1.2 Data from source book 2

The group has 17 data, among which the expression with Uniquely Identifable referent occurs twice, those with In Focus and Activated referents occur 5 and 9 times respectively, and one expression referring Familiar referent has also been spotted.

6.1.3. Summary of Group *na.*

The analysis above involves several issues. I first clarify that the relation between non- reflective predicative NP and its subject in copular sentence is not coreferential, as being a premise for the analysis in this group, where *na* in the majority of the data occur in positive copular sentences.

The most common use of *na* is to refer to entity/entities with clear antecedent in the immediate previous discourse. The referents typically have the CS In Focus, with several exceptions of Activated, Familiar and Uniquely Identifable as well. Although it seems to be a pronominal only for singular referents (without the plural morpheme *men*), one datum has been found in which *na* refers to plural referents, i.e. *na − the two films that I am unable to direct.*

Another common situation is that the anaphor *na* does not have a clear corresponding antecedent in the previous context, yet the referent has been mentioned. In order to retrieve such referents, the addressee needs to interpret the definite description attributively (in Donnellan´s sense). I coded these examples as being Uniquely Identifable.

The reference distribution concerning data from the two source books is integrated and summarized in the table below.

Table 3. Reference Distribution of referring expression *na*

CS	**Inf**	Act	Fam	Uni	Ref	Type	Total
sum	**12**	11	3	5			31
Percentage (%)	**38.71**	35.48	9.68	16.13			100

As being illustrated, the referring expression *na* is commonly used for referents who are In Focus, which is in accordance with the briefness of the expression. The frequency tendency descended gradually from In Focus to Familiar, and rises in Uniquely Identifiable. 16.13% of the data have referents that are Uniquely Identifiable, it is the lowest CS that systematically occurs.

6.2 Group. *na* N

Analysis in this chapter is the major part of the data investigation. It contains the analysis of data from two source groups.

6.2.1 Data from Source book 1

Source book 1 includes 12 data, the referring expressions in 3 of them refer to normal entities, and those in 9 of them refer to time concepts. Among the latter 9 examples, 3 of them have the canonical form *na N*. Unexpectedly, the remaining 6 data share a slightly different structure, namely *na tian (´day´) N*. Strictly speaking, these 6 expressions do not fit the form of this group since they actually have *na N1 N2* structure rather than *na N*. I categorize them into this group because the fixed N1 functions more like a grammatical component than a noun, and its lexical meaning is rather weak. Within such a structure, N1 is fixed as tian (´day´), while N2 varies within the range of nouns indicating time concepts. For the convenience of discussion, I thus mark this subcategory as *na tian TN (a nominal indicating time concept)*.

6.2.1.1 Coding of a typical datum of na tian TN

Data in the *na tian TN* subcategory share a common property in terms of the ontological category of their referents, i.e time period. N1 is obligatory and fixed as *tian (´day´)*, and N2 is the head noun indicating a time concept whose duration is shorter than a day. To illustrate, (37b) is an example from this subcategory and (37a) its English counterpart.

(37a) that　　Ø　　　　afternoon

(37b)　那　　天　　　　下午
　　　na　　tian　　　xia wu
　　　that　　day　　　afternoon

The Mandarin expression in (37b) means *that afternoon*. As it can be seen clearly from above, apart from the two accurately mutually mapping forms, one element is added in the Mandarin expression, i.e. 天 *tian* (´day´), the *N1*. The same applies for all the time concepts shorter than a day in Mandarin, including *morning*, *noon*, *afternoon*, *evening*, *night*, etc.

Now let us have a closer examination of the additional element. The mandarin form 天 (tian) corresponds to:

a).   A common noun meaning sky, e.g. 蓝天, *lan tian* ("blue sky");

b).   A common noun representing the time concept day. e.g. 前天, *qian tian* ("the day before yesterday"); 那天, *na tian* ("that day")

c).   A classifier indicating the duration or the frequency of an event. e.g. 睡了 三 天 , *shui le san tian,* ("to have slept for three days");

In my opinion, the *tian* of N1 corresponds to (b), i.e. it is a common noun meaning day, although being incapable of showing its own duration if followed by TNs such as *afternoon*. In *na tian TN*, the TN is the head noun, while *tian* is semantically weak. It functions as the entity that conveys the referent of TN, which is also the head referent of the whole expression. For instance, the duration indicated by *na tian xiawu* ("that day afternoon") is the time length of an afternoon, not a day.

N1 is not a classifier, because it can be preceded by only one numeral, namely *yi* ("one"), and this is against the most essential character of a classifier, considering that classifiers have to be capable of being headed by various numerals so that the quantities are marked.

An ideal example to demonstrate the link between tian (´day´) and the TN is (38), in which the referring expression *na tian xia wu* ("that day afternoon") occurs in the last sentence, and the whole previous context is used to activate the day that conveys this particular afternoon. The referent is Uniquely Identifable as the day that conveys it can be uniquely identified.

(38).

如今, 在 学校 上 课 时, 她 变得 很 难 专心 听 课。
Rujin zai xuexiao shang ke shi ta biande hen nan zhuanxin ting ke
Now at school have class when she become very hard focus listen class
"Sophie felt it difficult to concentrate while she was having class nowadays."

最后 一 堂 课 的 下 课 铃 响起 时 她 飞快 走 出 学校,
zuihou yi tang ke de xia ke lin xiangqi shi ta feikuai zou chu xuexiao
last one CL class POSS over class bell ring when she quickly go out school
"When the last class was over, she rushed out of school quickly"
(… …)

苏菲 打开 信箱 时, 感觉 自己 心跳 加 快。
Sufei dakai xinxiang shi ganjue ziji xintiao jia kuai
Sophie open mailbox when feel herself heart jump more quickly
"Sophie felt that her heart was jumping faster when she opened the mailbox"
(……)
当 她 关上 园 门 时, 发现 有 一 个 大 信封 上 写着 她的 名字
dang ta guanshang yuan men shi faxian you yi ge da xinfeng shang xiezhe tade mingzi
when she close garden door when discover have one CL big envelope on write her name
"When she closed the gate of the gardern, she saw a big envelope with her name on the cover"
(……)
苏菲 看 了 看 手表, 时间 是 两 点 四 十 五 分。
Sufei kan le kan shoubiao shijian shi liang dian si shi wu fen
Sophie read PART read watch time be two hour fourty five minute
"Sophpie read the watch, it was a quarter to three."
(……)
那 天 下午, 苏菲 的 妈妈 回家 时 苏菲 仍 处于 震惊 状态 中。
Na tian xiawu sufei de mama huijia shi sufei ren chuyu zhengjing zhuangtai zhong
That day afternoon Sophie POSS mum get home when Sophie still located shock situation in
"That afternoon, when Sophie´s mum got home, Sophie was still in shock."

The datum in (38) is consisted of sentences selected throughout the source chapter whereby all the events take place during the same day, and the selected sentences offer the addressee information about time shifting within this particular day. The referring expression, *na tian xiawu* (´that day afternoon´) occurs in the last sentence of (38). It is not difficult to recognize a schedule-like clue running through the day in the narration, meanwhile, no words signalling date changing is used ever since the first expression marking a new day occurs. Therefore, the day is unique for the addressee despite the lack of a date, since he is offered enough details that only exist on that particular day.

na tian xiawu (´that day afternoon´) occurs after the day has been Activated in the addressee´s memory, and since there is no word indicating that the date has been changed, the addressee would naturally assume that it is the afternoon of the very

same day that *na tian* (´that day´) refers to. The afternoon is, therefore, also Uniquely Identifable.

The syntactic-structural distinction between the Mandarin expression and its English counterpart does not result in crucial distinction in reference retrieval. In both languages, the *afternoon* can be uniquely identified by association with the day, as long as the day is Uniquely Identifable per se. The only difference is that Mandarin has overtly marked the link in the form by using *tian* (´day´). English, on the other hand, hides the link and leaves it to pragmatic inferences alone, i.e. a bridging inference as outlined in the relevant section of the theoretical review part earlier in this thesis.

In sum, referring expressions with time concept referents in *na tian TN* form occur 6 times in A1. The referents are Uniquely Identifable by associating with the days that include them. In order to locate the TN, the addressee needs to locate the day first, and this specific day is normally located via particular events that take place during the day. The referent is Uniquely Identifable.

6.2.1.2 A counter example of na tian N form

Most occurrences of *na tian N* (´that day´) in this group have a Uniquely Identifable referent that is Uniquely Identifable. However, datum (39) may be a counter example.

(39)
"现在 你 要 说 实话。 你 是 不 是 整 晚 都 在 外面?
Xianzai ni yao shuo shihua ni shi bu shi zheng wan zou zai waimian
Now you must say truth you yes NEG yes whole night totally PREP outside
" Now you must tell the truth? Have you been outside for the whole night (tonight)?

那 天 晚上 你 为什么 没 换 衣服 就 睡 了?
Na tian wanshang ni weishenme mei huan yifu jiu shui le
That day night you why NEG change cloth directly sleep ASPT
Why did you go to bed directly with out changing into pajamas that night?"

Datum (39) was selected from a conversation between Sophie and her mother where the latter was questioning Sophie about her recent weird activities. Notice here that the night mentioned in the first sentence (denoted by *zhengwan* ´the whole night`) is not the same one as the referent of our target referring expression in the second

sentence. I ignore the reader´s perspective for the moment and take Sophie as the addressee in this case, because the assumption of GHZ´s theory is that the speaker will choose a form that fits the CS he has presupposed for the addressee, and in the conversation segment of example (39), the speaker (Sophie´s mother) was certainly not talking to the readers.

The referring expression *na tian wanshang (´that day night´)* is followed by an event that takes place during the time period referred to. It would be easier for the addressee to retrieve the referent if she had talked about the night or even some particular events from that night with the speaker, because then the referent would be Familiar to the addressee in her long-term memory. Nevertheless, the context has not shown any sign that the addressee and the speaker have talked about that night. There is no evidence of any already-Activated events that the referent can build a bridging inference with. In fact, I, as a reader, am surprised that the addressee is even able to retrieve the referent without requiring more information. The only reason I can think of is that the event of Sophie sleeping without changing into pajamas is so rare and unusual for both the addressee and the speaker that it discriminates that particular night from all other nights – thus making it Uniquely Identifable. It may also be that Sophie actually has a representation of that night in her long-time memory, in which case it is Familiar to her. There is, however, no evidence for this in the text.[18]

Notably, the discussion above is based on the condition that the referent (a time concept) has not been overtly mentioned or discussed and is most likely not represented even in the addressee´s long-term memory. Thus in order to locate it, the addressee needs to access additional information from the events taking place during the period referred to, if not directly asking for the axis such as the date. In other words, more processing efforts are required and bridging inference between the head referent and the day that includes it is involved. The existence of bridging inference is crucial because it basically means that regardless of whether the event is Activated (in the addressee´s short-term memory) or Familiar (in his/her long-term memory) or new but unique, the referent will always be at least Uniquely Identifable.

---

[18] This is possible, because the writer has not necessarily written everything in the book.

6.2.2 Examples from source book 2

The group has 32 examples, among which referring expressions with Familiar referents occur 18 times (56.25%), those with Uniquely Identifiable referents occur 6 times, and those with Activated and In Focus referents occur twice and 4 times respectively. There is also one example whose referent is Referential and one whose referent is Type Identifiable.

6.2.2.1 Uniquely Identifable or Familiar

Following GHZ´s (1993) definition for the CS Familiar, the addressee can identify the intended referent because he already has a representation of it in memory. This includes two subcategories: In short-term memory if it has recently been mentioned; in long-term memory if not. The latter case can be that the intended referent is an entity that the addressee is Familiar with due to his knowledge about the world, some specific cultures and some personal experience that he shares with the speaker. The intended referent in (40) demonstrates how a referent can be retrieved via cultural knowledge.

(40)

此 间， 我 每每   看 它 一 鱼, 好 寂寞的鱼 啊 我 发   出 像   耶和华神 的   喟叹 SG1
ci jian wo meimei   kan ta yi yu hao jimode yu a wo fa   chu xiang yehehuashen de   weitan
this period I every time see it one fish very lonely fish INTJ I make ASPT as   Jesus Christ ASSOC sigh

「那人   独 居 不   好, 我 要 为 他 造 一 个 配偶 帮助 他 segment2」SG2,
na ren   du ju bu hao wo yao wei ta zao yi ge peiou bangzhu ta
that person alone live NEG good I must for him build one CL spouse help   him

我 亦 认真   考虑 过 是否 要   去 後山溪   捞 一 尾 同类   来 相伴 SG3
wo yi renzhen kaolv guo shifou yao   qu houshanxi   lao yi wei tonglei   lai xiangban
I also seriously consider ASPT whether going to to back mountain creek catch one CL same kind for accompany

"During this period, I have always looked at this lonely fish and sighed like Jesus Christ "it is not good to let that person live alone, I need to create a spouse to help him." I have also considered seriously about the idea of catching a wild fish from the Back Mountain Creek to accompany him."

The intended referent of the relevant expression *na ren* has not been mentioned previously, yet can be identified when the sentence has been fully processed. Segment1 mentions a lonely fish that just has lost his friend, and segment2 is marked

as a quotation uttered by Jesus Christ, as being pointed out in segment 1. Segment3 continues the narration by telling that the speaker once has thought of finding a new partner for the fish from the river nearby. If the addressee has the story of Adam and Eve in paradise lost in his mental storage of cultural knowledge, it cannot be more obvious that the referent of *na ren* is Adam, as the words *du ju* (´live alone´), *zao* (´create´) and most of all *ye he hua* (´Jesus Christ`) all point to the Adam in that famous story. The speaker compares the fish as the lonely Adam who is in need of a partner.

The referent is Uniquely Identifable if we regard it as being retrieved through bridging inference by association with the story and Jesus Christ, in other words a context and an entity that can be assumed to be known by the hearer through cultural/ encyclopaedic knowledge. However, it is Familiar if we regard it as an entity that the addressee already has in his long-term memory, as part of the story (if he knows the story, he is highly likely to know Adam as well). I am personally in favour of the first judgement as the referent was not referred directly by its name or similar (such as *1960s*), but by an expression (i.e. *na ge ren, that CL person*) that can refer to any person. Bridging inference with the story plays a critical role in the retrieval.

Moreover, for the addressees who have not stored such a story in their long-term memory, or those who have no idea about Jesus Christ, they can at most accept that the speaker refers to an unique entity, the referent thus has no higher CS than Uniquely Identifable in this most conservative situation.

6.2.2.2 Type Identifiable or not?

In the case of example (41), the speaker´s intention plays a role when I was judging the CS for the referring expression *na xuan zhuan mu ma yin yue* (´the music of merry-go-round´) which occurs in sentence 2. Sentence 1 and sentence 3 are included in the example but not glossed, considering that they only serve as background context for judging the speaker´s intention, as well as space limitation.

(41)

後来我看到隐遁的麦可杰克逊终於让欧普拉去他的梦幻谷采访，晚上凉风里他走到外面，奇怪他的庄园和游乐场修整得那样人工一丝不苟，像一所优良的公共设施，一座模型陪葬物。S1

" Later, I saw it on the TV that Michael Jackson finally let Oprah Winfrey interview him in his Neverland Valley Ranch. He went outside of his house in the evening with cool breeze. It was strange that his ranch and amusement park were so artificial and meticulous that the whole estate looked like a well-equipped public construction (rather than a home) as well as a model to be buried with the dead."

游乐场　　　永远　　令 我 伤感，　想　 到 马戏，小丑，假日，童年，曲终人散，
you le chang　yongyuan ling　wo shanggan xiang dao　maxi　xiaochou jiari tongnian quzhongrensan
amusement park　always　make me sad　　think ASPT circus　clown　holiday childhood music end people gone

而 那 旋转木马　　　音乐　　　真　　是 太 荒凉，
er *na* **xuan zhuan mu ma yinyue** zhen　shi tai huangliang
and **that merry-og-round　music**　really be too desolated

像　一缕 亡 魂 依　　绕　　　不 去　还 在　凭吊　　　往日　繁华。S2[19]
xiang yi lv wang hun yi　　rao　　bu qu　hai zai　pingdiao　wangri fanhua
like　a ray dead spirit snuggle surround NEG leave still ASPT pay homage to old day glory

" Amusement parks always sadden me, circuses, clowns, holidays, childhoods, feelings of being desolated, and the music of merry-go-round sounds so lonely that it feels like a wondering ghost paying homage to the past glory."

麦可对摄影机介绍他的旋转木马跟摩天轮，灿晶晶开亮著似两盘钻石座落於绒黑夜幕中。他说他有时会半夜一人去开旋转木马骑，天啊这是我所见过最最寂寞的人。S 3

" Michael Jackson presented his merry-go-round and Ferris wheel in front of the camera, they shone like two plates of sparkling diamonds in the velvet dark night. He said that he sometimes came and took a ride on the merry-go-round alone at midnight. My goodness! This is the most lonely person I have ever known."

The referring expression *na xuan zhuan mu ma yin yue* (´music of merry-go-round´) occurs in sentence 2. Sentence 1 narrates the fact that Michael Jackson finally allowed Oprah Winfrey to interview him in his private estate, which has a manor and an amusement park. In sentence 2, the speaker expresses her feelings about amusement parks and things involved with amusement parks in general. Sentence 3 continues the narration about what Michel Jackson did during that interview.

Now let us examine the nominal closer. One solution is to consider yin yue ("music") as the head noun, and ´merry-go-round´ as the element that modifies it, accordingly, the interpretation will be the music that comes from a specific merry-go-round. In this sense, all kinds of music are possible as long as it comes from that specific merry-go-round, it can be a birthday song, a national anthem; and the determiner *na* is used to single out the specific ´merry-go-round´. Nevertheless, if this solution is to be the intended interpretation, the speaker would want to adopt an associative phrase structure by inserting the corresponding classifier for amusement park after the

---

[19] S2: sentence 2

determiner, and then connect the two nouns with an associative phrase marker *de*. (i.e. Before*: na xuan zhuan mu ma yin yue*; After: *na **ge** xuan zhuan mu ma **de** yin yue*) The other solution is to consider xuan zhuan mu ma yin yue ("music of merry-go-round") as a compound noun that consists of two nouns, namely merry-go-round and music. The interpretation is thus: the type of music that typically comes from any merry-go-round. The addressee only needs to have knowledge about what kind of thing "music of merry-go-round" is to achieve the intended interpretation. In this case, the determiner is used to control the whole compound noun, and the referent is Type Identifiable.

Secondly, let us review the context where the referring expression occurs. The referring expression occurs after five parallel nominals. All the nominals (*amusement parks, circuits, clowns, holidays, childhood*[20], etc.) in sentence 2 are only Type Identifiable to the addressee, because the speaker has no intention to discuss a particular amusement park or a particular clown. Two *amusement parks* have been mentioned in the context. One is Michael Jackson´s amusement park that is already Activated in sentence 1, but it is irrelevant in the reference retrieval here. The other one is the Type Identifiable *amusement park* being mentioned at the beginning of sentence2. Linking with this one does not yield more positive cognitive effects than what the addressee needs to have, because the speaker was not expressing his feelings about the music of merry-go-round from that specific amusement park. Thus the second amusement park is also irrelevant. Judging by the context, the speaker´s intention is now quite clear: he only wants to discuss about some types of things in general. It is not logical that he suddenly talks about a particular entity right after mentioning five parallel Type Identifiable referents in the same sentence. In this sense, whether *na* exists or not has no influence on the interpretation, it might be here due to other reasons. The intended referent is the type of music that typically comes from any merry-go-round, and it is Type Identifiable in this example.

---

[20] Mandarin does not mark common nouns with the plural morpheme 们 （*men)* even when the referents are plural, the plural morpheme only occur after some certain words such as 我们(wo-men, we); 姐妹们(jie mei-men, sisters); 们 men will not be used to form a word such as 很多游乐场们(hen duo you le chang- men, many amusement parks), which is impossible.

This is a bit controversial. One possible opponent argument is that as long as the addressee can exclusively recognize this particular type of music from any music, this whole group is Uniquely Identifable. If this is the case, then ´bed line´ is a Uniquely Identifable group as they can be distinguished from any other kinds of lines, yet this is not true in an example such as " Bed lines must be changed frequently", where being able to understand what the compound word denotes is enough for the required interpretation. In the example, the referring expression is listed together with nominals whose denoted entities are very different from each other, rather than among a row of nominals denoting various music types. Consider the two contexts below:

(42).

42a.   amusement parks, circuits, clowns, holidays, childhoods, music of merry-go-round

42b.   birthday songs, national anthems, music of merry-go-round

In (42a), the Type Identifiable interpretation is satisfying enough for the interpretation, while in (b), it is true that the addressee has to exclusively recognize this particular type of music from other music types, and thus this whole group is Uniquely Identifable. The actual context of our example is that of (42a).

The next potential opponent argument is that the referent is Referential as it is mentioned in the subsequent discourse, which is true. However, the definition of ´Referential´ is: " The speaker intends to refer to a particular object or objects. To understand such an expression, the addressee not only needs to access an appropriate type- representation, he must either retrieve an existing representation of the speaker´s intended referent or construct a new representation by the time the sentence has been processed." In this example, if it is true that the speaker intends to refer to some particular objects, he intends to refer to the music of merry-go-round as a unique group, the referent group is then Uniquely Identifable; if he does not intend so, then the referents are Type Identifiable. There does not exist a third possible interpretation to fit in between those two interpretations. It would be odd to regard the referent as Referential, even though it does meet one of the criteria for Referential in the Coding Manual (GHZ, 2006).

Hence, I would consider the referent to be only Type Identifiable.

6.2.2.3 Referential or not?

The data in the second source book includes one particular example whose referring expression is actually used for Referential. Intuitively, (6) is a controversial example from several perspectives, yet I decided to include it considering that unsatisfying data are also inspiring data for deeper insights and discussion.

(43)

我 仍 记得 那 人 姓 施, 我们 每 星期 周末 会面,
wo ren jide ***na ren*** xing shi women mei xingqi zhoumo huimian
I still remember **that person** family name shi we every week weekend meet

延续 一 个 月, 他 突然 在 不是 应该 连络 的 时间 打电话 找 我,
yanxu yi ge yue ta turan zai bushi yinggai lianluo de shijian dadianhua zhao wo
last one CL month he suddenly at NEG should contact ASSOC time call find me

向 我 借 两 万 元。
xiang wo jie liang wan yuan
towards I borrow two ten thousand yuan

"I still remember that this person has the family name of Shi, we used to meet each other every weekend, (this) has lasted for one month. (One day), he suddenly called me at the time when he is not supposed to do so, wanting me to lend him 20, 000 Yuan."

The sentence in which the relevant referring expression occurs is 我仍记得那人姓施 (wo ren ji de *na* ren xing Shi, *I still remember that person family name shi),* it comes in SVO (subject verb object) order with *wo* (´I´) being the subject, *ji de* (´remember´) being the verb, and the rest of it being the object. As I said, it is a controversial example, which is due to the way in which the speaker composed the phrase. In fact, I found it difficult to even retrieve the referent. The referring expression is 那人 *na ren* (´na ren, that person´),* being followed by 姓施 (´xing Shi, family name Shi´), the sentence could mean that the author remembers the particular fact that the person´s name is not any other family names but Shi. This solution is grammatically right, however, pragmatically unacceptable. Judging by the subsequent context, the speaker has no intention of emphasizing that he remembers the person´s family name.

The sentence might predict that the speaker remembers this person, who happens to have the family name Shi. This interpretation becomes pragmatically obvious and turns out to be favoured by the context, but the way in which it is put is grammatically wrong. In this interpretation, 那人姓施(" na ren xing Shi, *that person whose family name is Shi*) is the direct object of the verb (ji de, "remember"), therefore, it has to be a noun phrase, at least. However, the structure the speaker composes the word is not qualified for being a grammatically-right noun phrase. Ideally, the speaker will want to adopt a relative phrase that carries modifying information about the head noun, and the relative clause is supposed to be connected to the head noun by the particle / associative phrase marker *de*, resulting in 那个姓施的人 (*na* ge shi xing **de** ren, the person whose family name is Shi). In this sense, the referent is Uniquely Identifable.

I propose a third solution here. The referring expression is only *na* ren ("that person"), and the sentence needs to be read as such:

我	仍	记得	那	人	**,**	**(他)**	姓	施,	我们	每	星期	周末	会面,
wo	ren	jide	***na***	**ren**		ta	xing	shi	women	mei	xingqi	zhoumo	huimian
I	still	remember	**that**	**person**		he	family name	shi	we	every week		weekend	meet

The comma and the third person singular pronoun (underlined) are inserted. In this solution, the referent is Referential, as it is evidential from the context that the speaker intends to refer to a particular referent (that is newly introduced), and it is mentioned in the subsequent discourse. The referent is not only retrieved by virtue of his family name, but the whole subsequent context after the comma, which includes information about the fact that the referent borrowed money from the speaker in a shabby way. The addressee is able to construct a representation of the referent by the time the sentence has been fully processed. This interpretation is compatible with the rest of the sentence, and it makes the sentence grammatically right. The referent of the referring expression na ren (´that person ´) meets both the criteria for Referential, and is Referential according to the Givenness Hierarchy theory.

To be honest, following the sense of my proposal, I believe the referent is too unique to be only Referential, since the addressee has a representation of this particular referent due to those very unique events in the context. This involves my questioning

about the definition of the CS Uniquely Identifable, I will discuss it with deeper insight in Chapter 7. Residual Issues about the Givenness Hierarchy theory.

6.2.2.4 The referring expression na ren ("that person")

One single referring expression repeats 9 times among the 33 data of the group. The expression is *na ren* ("that person").

Among these 9 occurrences of *na ren*, one of them has an in-focus referent, one has a Referential referent (as has been discussed above), one has an Activated referent, and the remaining 6 examples have Familiar referents. Some of these 9 examples seem to shown a tendency that their referents are particular persons or well-established names, for instance, Adam from the Bible, Sakyamuni Buddha of the Buddhism religion. Those names s are well-established concepts in their own cultures.

In addition, the referring expression is also used for referents that yield strong emotional reaction for the speaker. An example of thi latter case is (44):

(44)

我 灼 苦 等 着 他应该 给我一个 交待，
wo zhuo ku deng zhe ta yinggai gei wo yi ge jiaodai
I anxiously painfully wait ASPT he ought to give me one CL explanation

他 跟 那 人， 他 跟 我，我们，到底 是 要 怎样？
ta gen **na ren** ta gen wo women daodi shi yao zenyang
he and **that person** he and I we on earth be going to how

他 却 不 提。
ta que bu ti
he however NEG mention

" I was waiting anxiously and painfully for an explanation from him.
How is he going to handle the relation between that person and him, him and me, us.
But he does not mention it at all. "

The expression *na ren* ('that person') refers to the affair lover of the speaker´s boyfriend. Admittedly, the addressee identifies the referent by virtue of previous Familiarity about this person in his long-term memory. However, I have this intuition that *na ren* has developed from a normal referring expression indicating a person with Familiarity to a symbol of a particular person. The referent is a person that evolves

strong emotion for the speaker and thus strong cognitive effects for the addressee (as it is a monologue book, the readers, as the addressee, have similar feelings as the speaker), it has particularly prominent identifying effects in comparison with other referents. The reference assignment can be dated back to so further away that in other cases the speaker would need to use a more complicated linguistic form. Therefore, I believe that *na ren* not only indicates that the referent is Familiar, but also reveals some subtle feelings of the speaker and his attitude towards the referent.

In this example, the speaker may not want to assign a name to such a person, he thus uses *na ren* to represent him every time this referent occurs. It is in principle possible for the speaker to use a proper name such as the third person personal pronoun 他 ta (´he´), but he chooses not to. Based on expectations of optimal relevance - the interpretation will be the optimally relevant one, given the speakers preferences and abilities. Using the less informative and thus more processing demanding form *that person* ('that CL person') raises expectations of extra cognitive effects that would not have been achieved by the other means. I, as the addressee, do have a more profound impression of this particular person.

Another example concerning the speaker´s preferences is (45), in which the expression *na shi* refers to sexual intercourse.

(45)

夜	风	潮糊糊	刮	涂	我	脸	我	心	臆	阿尧	大约	是	去	干	了	那	事。

ye feng chao hu hu gua tu wo lian wo xin yi A´Yao dayue shi qu gan le *na* shi

evening wind damp blow paint/apply I face I heart guess A´Yao probably be go do ASPT that thing

" The damp evening breeze blowed my face, I guessed that A´Yao probably had sex (when he was absent)"

The relevant expression is *na* shi (´that thing´), judging by the context, it denotes the activity of sexual intercourse. Considering the background that Chinese culture is relatively conservative regarding mentioning sexual intercourse in natural discourse, speakers usually adopt some specific vocabulary to avoid uttering the word directly. The speaker´s preferences of utterance composing play a role in this example. According to the Relevance Theory, the speaker automatically seeks *optimal relevance* when he composes the utterance. *Optimal relevance* is defined in (47):

(46)   *Optimal relevance*

An ostensive stimulus is optimally relevant to an audience iff:

a.   It is relevant enough to be worth the audience's processing effort;

b.   It is the most relevant one compatible with communicator's abilities and preferences.

As b) points out, the speaker will not go against his own willing and preferences in producing the utterances. There might be ostensive stimulus that is more economical, yet if it is against the speaker´s preferences, he is likely to use a less economical one that he is able to utter. *Na shi* is one of this kind. Instead of using an explicit term encoding sex intercourse, the speaker uses *na* shi (´ that thing´), which is vague yet mild. It is a choice made due to preferences influenced by the speaker´s culture.

Similarly to *na ren* in (45), *na shi* also conveys some aspects of the speaker´s attitude towards the referent, but *na ren* has not reached the point as *na shi*, which almost conventionally indicates sexual intercourse. However, the 9 data in this group seem to indicate that it has started to share the same sense of referring to referents that the speaker would like to avoid addressing directly. They are not rare, another Chinese idiom 那厮 *na si* (´that guy´) is also widely recognized as indicating the speaker´s teasing altitude towards the person referred to. The fact that all these expressions involve ´*na*´ rather than other demonstrative determiner (e.g. the proximal demonstrative *zhe* (´this´)) can be an inspiring clue for further investigation.

6.2.3 Summary of Group *na* N

The group includes examples with referring expressions in the form *na N*. It consists of data from two source books. The subgroup from source book 1 contains 12 data, and the one from source book 2 contains 32 data.

In the discussion of the first subgroup, I have focused on the 9 referring expressions whose referents are time concepts, and have divided them into two subcategory, considering that 6 of them have a rather particular structure, namely *na tian* TN. The table below summarizes the distribution of *na N* in the first subgroup.

Table 4. Distribution of *na N* according to highest status (from subgroup1)

	Inf	act	Familiar	uni.Id	ref	Type.id
Na N (A1)		2	1	**9**		

Concerning expressions with time concept referents in general, I argue that they can be uniquely identified by association with the events taking place during the period referred to. According to my observations, under circumstances where no previous Familiarity is involved in the reference assignment of time concepts, the referents can usually be rightly interpreted based on not only the information encoded in the nominal, but also the predication of the rest of the sentence and pragmatic inference.

Moreover, I explored the distinction between the English referring expression ´that night´ and its Chinese counterpart, i.e. *na tian wanshang* ("that day night") regarding reference assignment. By a closer examination of the additional element *tian* in the Chinese expression, I concluded that they come in a rather frozen structure, *na tian TN,* where *tian* (´day´) is the additional element functioning as a fixed component. In both English and Chinese, an *afternoon* can be uniquely identified once the day that includes it is uniquely identified, and this link is overtly marked in the Chinese version by *tian* (´day´), whereas it is hidden in its English counterpart. In spite of that, the reference assignment is more or less similar, and the referents in both cases are Uniquely Identifable if no previous Familiarity is involved.

In the second subgroup, I end up having a rather complex pattern whereby reference of *na N* distributes across each of the cognitive status on the hierarchy.

Table 5. Distribution of *na N* according to highest status （group2）

	Inf	act	Familiar	uni.Id	ref	Type.id
Na N (A2)	4	2	**18**	6	1	1

Notably, referents that are Referential and Uniquely Identifable only occur once each, I have discussed these two data in depth in the content above, yet there are still unsettled issues in judging the CS for the referent that is contemporarily considered as Referential, and the datum itself seems to have a problem. As for the datum with the Type Identifiable referent, it is also quite controversial. Therefore, these are not

convincing enough to be considered for the conclusion, plus, they do not occur systematically.

The second subgroup also includes examples where the expressions refer to time concepts, but they come in the normal structure of *na N* (rather than *na tian TN*). Comparing to the examples in *na tian TN*, these examples are free from the distraction of the link between the day and the smaller time portion (for instance *night*), therefore, they demonstrate the idea that a time concept referent can be uniquely identified by what is stated in the rest of the sentence where the referring expression occurs, the significance of context inference is thus emphasized. These examples call for deeper insights on the nature of the cognitive status Uniquely Identifable. I will discuss it as residual issues later in the thesis.

To conclude, reference distribution of the two subgroups of data are integrated in the table below.

Table 6. Reference distribution of Group. Na N

cs	inf	act	fam	uni	ref	type	total
sum	4	4	19	15	1	1	44
Percentage (%)	9.1	9.1	43.18	34.09	2.27	2.27	100

According to the summarized table, Familiar is the most frequent CS, covering 43.18% of the data. To be conservative, I will only consider Uniquely Identifiable as the lowest CS that occurs systematically, and conclude that na N encodes Uniquely Identifiable accordingly.

6. 3 Group *na* CL N

6.3.1 Referential or not?

The group contains two pieces of data whose referents seem to be Referential to the speaker. As far as the data analysis has processed, the usage of *na* being involved in these two examples is the one that approaches the nature of the cognitive status

Referential the most. Considering that the two referring expressions are used quite similarly, I will only include one of them as example (47) to demonstrate how they are analysed.

(47).

我 继续 写, seg1 此 刻 我 的 心情 seg2 (Sentence1)
Wo jixu xie ci ke wo de xinqing
I continue write this moment my emotion
"I continue to write about my feelings now"

你 还 记得 那首 词吗, seg1 水远山长愁煞人[21], seg2 就是这样 seg3 (Sentence2)
ni hai ji de *na* **shou ci** ma shui yuan shan chang chou sha ren jiu shi zhe yang
you still remember **that CL poem** water far hill high sadness kill people just be this look

"I continue to write about my feelings at this moment, do you still remember that poem? The distance has saddened me so much. (My feeling) is just like that."

The referring expression *na shou ci* (´that CL poem´) occurs in segment 1 of sentence 2. Sentence 1 explains the background that the speaker continues to write about his feeling on the postcard after a short break. The addressee of S1 is the reader of the book, as it is a part of the narration; the addressee of S2 is the receiver of the postcard, as it is directly quoted from the postcard. Since the target referring expression occurs in sentence 2, the receiver of the postcard is the actual addressee for the referring expression. Sentence 2 contains three segments, the speaker asks the addressee if he still remember *that poem* in segment 1, segment 2 is the poem line being written on the postcard. In segment 3, the speaker confirms that his feeling is just like what is described by segment 2. Judging by this context, the referent of *that poem* is the line being explicitly stated in segment 2.[22]

---

[21] From the poem 黄陵庙 *(The Huangling Temple) by* 李群玉 *(Qunyu Li, Tang Dynasty).* The poem is: 黄陵庙前春草生，黄陵女儿茜裙新。轻舟小棹唱歌去，水远山长愁煞人。

[22] Theoretically, it is also possible to argue that the referent of that poem is the poem from which the line in segment 2 is selected. I am uncertain about the speaker´s considerations when he uses shou, the classifier for a piece of poem as a whole, because he could have used the more specific classifier 句(jù) that is only for lines and sentences. However, considering that the writer already had written several unnatural sentences in the book, this might be another one of them, but I am not 100% sure.

It seems plausible to claim that the referent is Familiar to the addressee based on shared personal experiences between the speaker and the addressee, because the speaker asks the question ´Do you still remember…´ immediately before the referring expression, which is the object of the question. The word hai (´still`) and the verb ji de (´remember´) suggest that the speaker assumes that the addressee has heard the poem previously. The referent thus has the CS Familiar, despite that the referring expression alone is not enough to ensure that the addressee can uniquely identify the referent, I will discuss about this fully later.

6.3.2 na precedes a proper noun

The group includes a rare datum whose referring expression does have the structure *na N*, yet the *N* is a proper noun. Similar to English, proper nouns in Chinese are not supposed to be preceded by any determiner, the referring expression in (48) is a counter example to this rule.

(51)

他 师父　　的　　师父 跳 到 七十六　　岁，跳　　那 位 特洛伊　　皇后，
ta shifu　　de　　shifu tiao dao qishiliu　　sui tiao ***na wei teluoyi huanghou***
he master POSS master dance til　seventy-six year dance **that CL Troy　　queen**

年老的　　海克芭 看　　著　　她 所爱 之　　人　一 个 个 於　眼 前 死 去 的　　意象，
nianlaode haikeba kan　　zhe　ta suoai zhi　　ren　yi ge ge yu　yan qian si qu de　　yixiang
aged　　Hecuba witness APST she love ASSOC person one CL CL PREP eye front die go ASSOC imagery

如　此 告别　　了　　舞台。
Ru　ci gaobie　　le　　wutai
like this bid farewell to ASPT　stage

"The master of his master had danced until the age of seventy-six, playing the Queen of Troy and the imagery of the aged Hecuba witnessing those who she loved passing away in front of her eyes one by one. This was how (he) bided farewell to the stage."

The referring expression is *na wei teluoyi huanghou* (" that CL Queen of Troy), and its referent is the Queen of Troy that is being portrayed in a specific piece of ballet

Nonethless, I will not take this possibility into consideration, because the original poem that the two lines are selected from has four lines, and not every line in the poem describes the speaker´s feeling. As he confirmed in segment 3, it is the line in segment 2 that precisely describes his feelings at that moment. Therefore, he should have no intention of reminding the addressee of the whole poem.

dance. In the referring expression, the determiner *na* precedes a proper noun, i.e. the Queen of Troy. More specifically, it is a full name in Mulkern´s sense. Mulkern divided proper names to two categories, full names and single names. She states: " A full name is frequently used to refer to an entity for which the addressee is not yet expected to a representation (…)" and single names generally include family names, given names and nicknames. A full name such as *Queen of Troy* does not require the addressee to have a representation of the entity in his long-term memory, therefore, the referents of full names have CSs that are no higher than Uniquely Identifable. The addressee accepts that there is a unique queen of Troy, it is not just a queen of an unspecified dynasty or country. Therefore, the referent of *teluoyi huanghou (*"Queen of Troy") is Uniquely Identifable to the addressee even without the information encoded in the determiner and the classifier. Given the grammatical rule that no determiner is required before proper nouns in Chinese, it is thus challenging to explain the speaker´s motivation of using *na.*

Intuitively, one possibility is that the speaker intends to hint the addressee that he refers to a specific entity, and that he will continue to talk about this entity. It is true that the speaker talks about this specific entity in the subsequently discourse, and in this sense, the referent is Referential. However, if this is the case, using *na* is unnecessary, because the referent is Uniquely Identifable to the addressee by the proper noun per se. According to the Givenness Hierarchy theory, a Uniquely Identifable referent is also Referential to the addressee by definition. The speaker does not need to use another component to signal that the referent is Referential once more. The second possibility is that the speaker assumes the referent to be Familiar to the addressee. In this example, the addressees are the readers of the book, and the speaker assumes the readers all have a previous representation of the Queen of Troy in their memories. This can hardly be true in the normal sense of writing. *Na* is thus not used to signal previous Familiarity of the referent.

If the speaker does not use *na* wrongly, the only reason to explain his choice of using *na* is that he does not regard the Queen of Troy as Uniquely Identifable, or, it is not precisely enough. A normal nominal usually undergoes a lexical pragmatic process during the reference assignment. For instance, in the sentence ´I forgot to visit the bank, therefore I can´t return the money that I bought from you." In order to achieve

optimal relevance, the bank needs to be interpreted narrowly, as ´the bank office that deals with private individuals and the one that I have a bank account with deposit money." Similarly, the referent should be interpreted as ´the queen of Troy inside the ballet world´, not just the historical figure. It is possible that the historical figure ´the queen of Troy` is not exactly the same as the one in the ballet world, as the latter might be an artistic figure created by the script writer based on the historical figure. In this sense, *na* is used to mark the narrowed interpretation. This interpretation is reflected by the subsequent context, where the discussion focuses on the performance of the artistic figure.

6.3.3 Summary of Group. *na* CL N

Group na CL N. includes examples whose referring expressions come in the form *na* CL N, which is probably the closest form to what is meant by *na* N by GHZ (1993), as they did not explicitly distinct *na* N and *na* CL N in the summary table of their empirical investigation.

Table 7. Reference Distribution of Group. Na CL N.

CS	Inf	Act	Fam	Uni	Ref	Type	Total
sum	3	5	26	10			44
Percentage (%)	6.82	11.36	59.09	22.73			100

Table 9 illustrates the reference distribution of group. *na* CL N. Referring exressions in *na CL N* form are largely used to refer referents whose CS is Familiar, the tendency is pretty clear comparing to the percentage of the second ranked CS. The data amount of Familiar referents almost triple that of Uniquely Identifable data. Reference distribution of *na* CL N ranges from the highest In Focus to Uniquely Identifiable, but no lower than that. Hence, *na* CL N encodes Uniquely Identifiable.

6.4. Group Na CL Adj N

The group consists of a total number of 15 examples, all of the 15 referents have the cognitive status Familiar.

6.4.1 A typical example

(49).

苏菲，当 你 看 到 一 个 影子 时，一定 会 假定　　有 一 样 东西 投射 出 这个 影
子 s1
Su fei dang ni kan dao yi ge yingzi shi yiding hui jiadign　you yi yang dongxi toushe chu zhe ge yingzi
Sophie when you see ASPT one CL shadow when surely will presuppose exist one CL thing reflect out this CL shadow
Sophie, when you see a (random) shadow, you will normally suppose that it is the shadow of a certain object.

你看到一只动物的影子，心想那可能是一匹马，但你也不太确定。S2
Ni kan dao **yi zhi dongwu de yingzi** xin xiang *na* keneng shi yi pi ma dan ni ye bu tai queding
You see ASPT **one CL animal POSS shadow** heart think that maybe be one CL horse but you neither NEG very sure
When you see **the shadow of a kind of animal**, you might wonder that the animal is a horse, but you are not sure.

于是你就转过身来，瞧瞧这匹马。S3
Yushi ni jiu zhuan guo shen lai qiaoqiao zhe pi ma
Thus you then turn around look at this CL horse
Therefore you turn around to look at the horse directly.

而比起**那模糊的影子**，这匹马当然显得更俊秀，轮廓也更清晰. S4
Er bi qi ***na* mohu de yingzi** zhe pi ma dangran xiande geng junxiu lunkuo ye geng qingxi
And compare PART **that vague shadow** this CL horse surely seems more handsome figure also more clear
Comparing to **that vague shadow**, this horse is certainly more handsome, and the figure is clearer.

The referrign expression *na* mohu de yingzi (that vague shadow) appears in S4, it refers to the vague shadow of a certain type of animal, which has been previously mentioned in sentence 2. The interpretation meets one criterion for the status ´Familiar´ provided in the Coding Manual: It was mentioned at any time previously in the discourse. It is retrieved from the addresse´s long-term memory.

The other 14 referents are in more or less the same situation. They have been overtly mentioned in the previous discourse, but rather far away, and they are retrieved from long-term memory.

6.5 Group *na* CL RC N Vs Group. RC *na* CL N

The following two groups contain data whose nominals contain a relative clause that describes the referent of the expression. Mandarin allows for two possibilities concerning the position of relative clause (henceforth, RC) in a nominal. A RC can

either precedes the *na* CL cluster or follows it, both the *na* CL RC N and RC *na* CL N combinations are grammatical.

I will examine and discuss data of the two groups together to explorer whether the referents´ cognitive status is influenced by word order.

6.5.1 Group. *na* CL RC N

6.5.1.1 Data from Source book 1.

6.5.1.1.1 A typical example for expressions indicating Familiar referents

Example (50) is a typical datum of this group.

(50).

苏菲正坐在秋千上，想着哲学课程
sufei zheng zuo zai qiuqian shang xiang zhe zhe xue ke cheng
Sophie ASPT sit PREP swing POSTP think ASPT philosophy cuorse

与席德（那 位 收 不 到 她 父亲 寄 来 的 生 日 卡 的 女孩）
Yu xide *na* wei shou bu dao ta fuqin ji lai de sheng ri ka de nv hai
And Hilde that CL receive NEG ASPT she father send come PART birthday card ASSOC girl

之间究竟有什么关系？
zhi jian jiu jing you shen me guan xi
PART between actually have what relation

Sophie is now sitting on the swing, trying to figure out the relation between the philosophy courses and Hilde (the girl who would not be able to get her birthday card sent by her father).

The relevant expression is *na wei shou bu dao ta fuqin ji lai de sheng ri ka de nv hai* , *("*the girl who would not be able to get her birthday card sent by her father"*)*. It starts with *na* together with the obligatory classifier for human beings, being immediately followed by the RC, and ends with the head noun nv hai (´girl´).

The relevant expression occurs within the parenthesis of the source text to explain its preceding proper noun *Hilde*. Previously in the context, Sophie received a birthday card that was supposed to be sent to a girl called Hilde. That is, this ´Hilde´ girl has been previously mentioned to the addressees (readers of the source book) and thus meets one of the criteria for Familiar, i.e. it has been mentioned anywhere in the

previous discourse. In fact, for readers who remember that Hilde is the girl who did not get her birthday card, the referent is even In Focus.

Example (50) serves as a sample of the majority (9 of 11, 81.82%) of Group. DM+RC, and the remaining two data have Uniquely Identifable referents. Such a clear pattern seems to indicate a tendency of mutual preferences between expressions in DM+RC order and Familiar referents. Nevertheless, data in the second source book present a contradictory trend, where Familiar referents counts for 42% of the whole group, being slightly exceeded by the percentage (57%) of Uniquely Identifable referents.

6.5.1.2 Examples from source book 2.

6.5.1.2.1 Examples for Uniquely Identifable

(51)

我只把视线留在那杯冰冻冒珠浮堆鲜奶泡沫红樱桃的咖 啡上

wo zhi ba shi xian liu zai *na* bei bing dong mao zhu fu dui xian nai pao mo hong ying tao de ka fei shang

I only PART eyesight focus at *na* CL(cup) frozen emit bead float stack fresh milk froth red cherry ASSOC coffee POSTP

"I only focuses on that cup of iced coffee with fresh-milk-made-froth and red cherry (as its topping)"

In (51), the cup of coffee referred by *na bei bing dong mao zhu fu dui xian nai pao mo hong ying tao de ka fei (`that cup of iced coffee with fresh-milk-made-froth and red cherry`)* can be exclusively identified due to the detailed description of its topping, which is encoded in the RC. Since the referent has not been mentioned previously and can be identified through the encoded information in the nominal per se, it is judged as being Uniquely Identifable. Example (51) represents 57% of the data of group. na CL RC N in source book 2.

6.5.1.3 Summary of Group. na CL RC N

Data from the two source groups have presented inconsistent patterns. Source book 1 has 10 data in total, among which 8 of them (i.e.80%) have Familiar referents, and the remaining 2 data have Uniquely Identifable referents. The mutual preferences between this data group and the CS Familiar are quite clear, judging by the table. However, the data from source book 2 argues differently. The group has a total

number of 23 examples, and referents in 21 of are judged as being Uniquely Identifable. The tendency in both groups are overwhelmingly clear, it is thus risky to conclude whether CS is the most preferred one. The lowest CS in both subgroups is Uniquely Identifiable, it is thus properiate to say that na CL RC N encodes Uniquely Identifiable.

6.5.2 Discussion about RCs

"Fox and Thompson (1990, p.301) identify two major types of relative clauses according to their functional roles: characterization and identification. In the first type, the relative clause provides a characterizing assertion or description of a new head NP referent in a particular discourse situation to supply additional descriptive information regarding the head noun. In the second type the relative clause makes the referent of a head NP relevant at a point in an particular discourse situation when it is first introduced." (Ming, 2010:332)[23] The two examples in (57) are used to illustrate the two discourse functions.

(52).   a. This man (who I have for linguistics) is really too much.

    b. There´s a woman in my class who´s a nurse.

(Ming , 2010:332)

The RC in (52a) is claimed to have the discourse function of identification, because the RC contains the given referent ´I´, which is used to ground the newly introduced referent ´this man´. On the other hand, the RC in (52b) only makes a characterizing assertion about the new referent ´a woman´, it is not used to anchor it. The two examples in (52) both have ´who´ as the relative clause particle, and none of them uses a comma to pause the sentence, the distinction between identification and characterization is not similar to that between restrictive and non-restrictive relative clause.

The referring expression in (53) contains a RC of characterisation, as it does not contain any given referent to anchor the new referent. On the other hand, the head

---

[23] The original Fox & Thompson (1990) article is not reachable, therefore the MING (2010) citation is cited.

referent A´Yao can actually be retrieved by the proper noun, there is no need for the RC to function as identification neither.

(53).

妈妈 每次上　楼 送茶 食铺床，添被褥向 我 传　述 主的道理，
ma ma mei ci shang lou song cha shi pu chuang tian bei ru xiang wo chuan shu zhu de dao li
mother every time upstairs deliver tea food make the bed supply blanket towards me sermonize Lord POSS ideas
是籍我 讲 给那个 根本 　不　 听的 　阿尧。
Shi jie wo jiang gei *na ge gen ben bu   ting de   A´Yao*
Be via I tell to   **that CL completely NEG hear PART A´Yao**

"(A´Yao´s) mother had been sermonizing to me when she came upstairs to bring us food, make the bed, and to bring us blankets. She was actually preaching via me to A´Yao, who strongly refused to be sermonized."

The relevant nominal is *na ge gen ben bu ting de A´Yao (A´Yao, who strongly refused to be sermonized)*, it refers to A´Yao. As the main character of the book, he is frequently mentioned and thus has constantly renewed Familiarity to the addressee. The proper name alone is adequate enough for the referent retrieval. The information encoded in the RC is not intended for identifying the referent, but for offering additional information to yield more positive cognitive effects.

6.5.3 Group. RC na CL N

Now let us progress to the next group, where the relative order of RC and na CL is reversed.

6.5.3.1 Examples from Source book 1.

6.5.3.1.1 A basic example

A basic example for the data is (54).

(54).

如今，除了苏菲以外，
ru jin chu le su fei yi wai
nowadays except for Sophie PART out

大家都认为这行老树篱就像园子另一边那个兔笼子一般，没有什么用处。S2

Da jia dou ren wei zhe hang lao shu li jiu xiang **yuanzi ling yi bian** *na* **ge tu long zi** yi ban mei you shen me yong chu
Everybody all think this CL old bushes just like *garden other one side that CL rabbit hutch* the samee NEG have any usage

´To everyone but Sophie, the old hedge was just as useless as the rabbit hutch at the other side of the garden. ´

The relevant expression is *yuanzi ling yi bian na ge tu long zi (the rabbit hutch at the other side of the garden.),* it can be exclusively identified due to the RC. The referred rabbit hutch has not been mentioned previously, therefore, it does not have a CS higher than Uniquely Identifable.

The RC contains an already grounded entity, namely yuanzi (´garden´). It is not any random garden but the one that has been Activated in the previous context. Since the RC functions to identify the referent, it belongs to RC of identification in Fox & Thompson (1990)´s sense. Such grounding progress has been observed and listed as ´bridging inference´ in GHZ´s Coding Manual, being the second subtype of the CS Uniquely Identifable.

As has been mentioned, there exist two exceptional Uniquely Identifable examples in Group *na* CL RC N, where all the rest are Familiar. In Group RC *na* CL N, the majority examples have Uniquely Identifable referents. Nevertheless, concerning the discourse function of the RC, those two exceptional Uniquely Identifable examples in the previous group do not behave exactly the same as those Uniquely Identifable ones in this group, even though they have the same CS. In the former two cases, their RCs contain information that are independent from previous mentioning, yet the RCs in the latter examples always involve information that requires previous Familiarity. See (55) and (56):

(55).
在此之前，大家已经商量好要想的东西是
zai ci zhi qian da jia yi jing shang liang hao yao xiang de dong xi shi
 PRED this PART before people already discuss well will want PART thing be

那只正在隔壁花园里玩耍的猫咪"毛毛"
na **zhi zheng zai ge bi hua yuan li wan shua de mao mi mao mao**
that CL ASPT next door garden in play PART cat maomao

"People had decided the thing they wanted was the cat that was playing in the garden next-door being named as Maomao. "

The referring expression is *na zhi zheng zai ge bi hua yuan li wan shua de mao mi Maomao ("the cat that was playing in the garden next-door being named as Maomao").* The referent is identified by virtue of the RC that contains the word *garden*. The garden has not been Activated in the previous discourse, rather than being a grounded entity to ground other new entities, it is mentioned for the first time per se. Still, it can be accepted that there is a garden where the cat is playing when the sentence is uttered, and the referred cat is Uniquely Identifable because of this.

The interesting part is, when *the garden* that is used to ground a new referent has been mentioned previously, the word order changes, this can be demonstrated by (56) from the previous group.

(56)

这栋红房子坐落在一个很大的园子中.S1
zhe dong hong fangzi zuo luo za yi ge hen da de yuanzi zhong
this CL red house locate PREP one CL very big garden POSTP
´The red house was surrounded by a large garden´

(…… context about the garden)

如今，除了苏菲以外，
ru jin chu le su fei yi wai
nowadays except for Sophie PART out

大家都认为这行老树篱就像园子另一边那个兔笼子一般，没有什么用处。S2
Da jia dou ren wei zhe hang lao shu li jiu xiang **yuanzi ling yi bian** *na* **ge tu long zi** yi ban mei you shen me yong chu
Everybody all think this CL old bushes just like *garden other one side that CL rabbit hutch* the samee NEG have any usage

´To everyone but Sophie, the old hedge was just as useless as the rabbit hutch at the other side of the garden. ´

Similar to (54), the referent of (55) is also Uniquely Identifiable through the information encoded in the RC. The only difference is that the garden in (54) has been mentioned, which means that the RC in (54) provides an already Activated entity for the new referent to anchor, while that in (55) provides completely new information. With regards to their internal orders, the referring expression in (54) comes in na CL RC N order, while the one in (55) comes in RC na CL N order.

6.5.3.2 Examples from the second source book

Example (63) has a possessive phrase in the position of RC.

(56).

至此我惊悚 发觉除 了   他那 个家，我们 的   窝，
zhi ci wo jing song fa jue chu le   ta *na* ge jia wo men de   wo
til this I horrifying realize except PART his that CL home we   POSS nest

我竟 然   再 无   可 与 他连 系 的   点，线。
Wo jign ran   zai wu   ke yu ta lian xi de   dian xian
I   unexpectedly more NEG can with he connect ASSOC dot line

The modifier is a possessive marker, namely ´his´, it is adequate enough to uniquely identify the referent once the identification of the person is clear. Normally speaking, *na CL* is not necessarily required if the speaker only wants to communicate ´his home´. Somehow, the speaker has chosen to insert the demonstrative *na* (attached by a classifier). Rather than considering the possessive phrase as a variation of the RC, it seems that the demonstrative is later inserted. As a result, we now have the difference between the original *ta jia (´his home´)* and *ta na ge jia (´that his home´)*.

The subsequent parallel nominal, *wo men de wo ("our ghetto")* explains that it is a place that the speaker used to live in together with his boyfriend. Intuitively, the difference between ta jia (´his home´) and ta *na* ge jia (´that his home´) in this context is that the referent of the former can be both newly introduced or Familiar as long as the personal pronoun has been explained, but the referent of the latter one (with *na CL*) should be Familiar to the addressee. The speaker has seemingly used *na* to emphasize the addressee that it is a referent that the he has particular Familiarity with.

A similar case is (57) where the person being referred to was present in the immediate spatio-temporal context.

(57)

他那一 身   家当，帅奇表，金项链，红绳络一块绿玉挂在颈下，
ta *na* yi shen   jia dang shuai qi biao jin xiang lian hong shen luo yi   kuai lü yu gua zai jing   xia
he that one CL/body property Swatch watch golden necklace red   line   one   CL green jade hang PREP neck POSTP

大胆 小 妖精多半有   人   养他 罢。
Da dan xiao yao jing duo ban you   ren   yang ta ba
Daring little demon   probably have person support he MOOD

´His set of accessories, Swatch, golden necklace, a green jade hanging by a red line, little daring demon, he must be supported by someone (rich).`

The referring expression is *ta na yi shen jia dang (´ that his suite/set of accessories´)*, it refers to the person´s outfit as a whole. The speaker describes a stranger who came uninvited and sat opposite to him by the other side of the table in a café. The person is within the speaker´s eyesight, so is his outfit. Since he has been describing the stranger´s movements in the preceding sentences, the person is Activated by means of the speaker´s eye gaze. The speaker uses a third-person personal pronoun, which indicates that the person has been mentioned, thus the referent (the outfit) is uniquely identified through this already Activated entity as it is worn by the person.

Notably, the referring expression *ta na **yi** shen jia dang (´ that his one suite/set of accessories´)* contains not only the determiner *na*, but also another word yi (a, one), a numeral that is believed to have started functioning as an indefinite article/determiner in Mandarin. Nevertheless, this *yi* should not influence the referent retrieval, a closer examination reveals that it is only a numeral rather than an indefinite determiner (the expression already had one). Imagine that the speaker describes the eyes of the same stranger in exactly the same scenario, he would use *ta na liang zhi yan jing (his that two CL eyes)*. Therefore, yi in (57) does not function as a determiner here: It can be substituted, and replacing it will not influence the CS.

The referring expression serves as an outline key word for the description in the subsequent sentences, where each of the ornamets are scanned and listed. The speaker was watching the boy from a distance away, he uses the distal demonstrative character of *na* to indicate that the referent is somewhere further away, yet still within eyesight. In this case, the referent is Activated to the addressee by meeting one of the criteria for Activated: It is something in the immediate spatio-temporal context that is Activated by means of a simultaneous gesture or eye gaze.

6.5.3.3 Summary of Group. RC na CL N

To summarize, the two source groups have 21 referring expressions. A total number of 13 tokens are found whose referents are Uniquely Identifable, taking 61.94% of the whole data set. Expressions with Familiar referents account for 38.10% (i.e 8 out of

21) of the data set. Thus, Uniquely Identifable is the most frequent CS, and it is also the lowest.

6.5.4 Comparison between Group. na CL RC N and Group. RC na CL N

The Cognitive statuses of the head referents are certainly not the only reason that determines the internal word order of a nominal. Other factors influence the relative order of the relative clause and the determiner as well. Therefor, a clear mutual preference between a certain word order and a certain CS is not expected.

The two subgroups of Group RC na CL N offer two mutually incompatible pattern, that is, one group of data has Uniquely Identifable as its most frequently referred CS, while the other group prefers Familiar the most. The pattern shown by Group na CL RC N is relatively clearer. As the data show, Uniquely Identifable is the most frequently referred CS by expressions in DM + RC structure.

However, we do find a minimal pair, namely example (54) and (55). The referents in both of the examples are Uniquely Identifable through the information encoded in the RC. The difference is that the RC in (54) contains an entity that has been previously mentioned, it functions as an already-Activated entity to anchor the new head referent, whereas the RC in (55) provides completely new information. Other factors being similar, the referring expression in (54) comes in *na CL RC N* order, while the one in (9) comes in *RC na CL N*. This might be a pair of examples to demonstrate the mutual influence between internal word order of a nominal and the CS, but not strongly so.
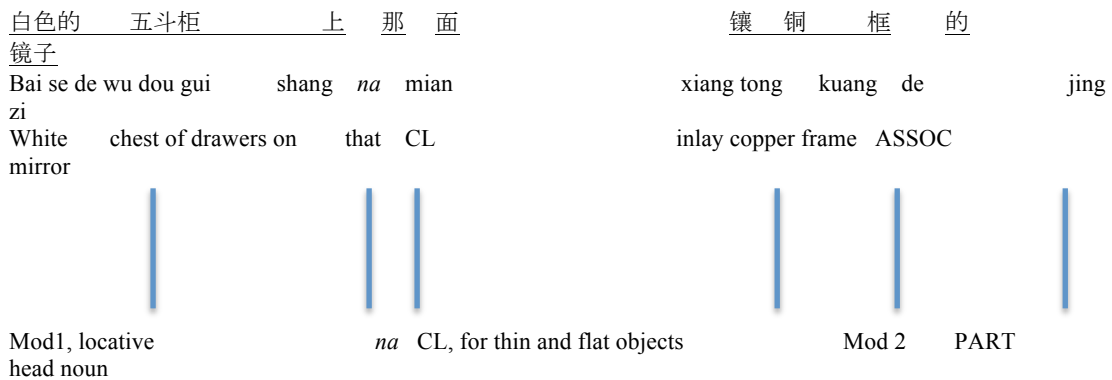
6.6 Group. Mod1 *na* CL Mod2 N

Referring expressions in this group have the most complicated surface structure among all the groups, they come with two modifiers that are posited by both sides of *na*, and a head noun.

6.6.1 A typical datum

 I will illustrate the interior of such a nominal by annotating the definite expression in (58).

白色的五斗柜上 那 面 镶铜框 的 镜子 是普通的镜子还是魔镜？

白色的	五斗柜	上	那	面	镶	铜	框	的	镜子
Bai se de	wu dou gui	shang	*na*	mian	xiang tong	kuang	de		jing zi
White	chest of drawers	on	that	CL	inlay copper frame	ASSOC			mirror

Mod1, locative head noun			*na*	CL, for thin and flat objects		Mod 2	PART	

"the copper-framed mirror on that white chest of drawers"

As has been indicated by the data of the previous groups, modifiers preceding the determiner commonly contain locative or possessive information about the referent, the modifier close to the head noun usually contain eventive information or information other than locative or possessive. In this example, Mod 1 describes the location of the head noun, i.e. mirror, as being on the white chest of drawers (this chest of drawer is highly likely to be the only one in the room, otherwise the speaker would have specified it with a determiner). The classifier is *mian*, theoretically, it can be used for thin and flat objects in general, yet in the real natural discourse, it is almost only precede ´mirror´ and ´drum´. Although unable to precisely pick out its corresponding noun, the classifier has sharply narrowed down the referent set. Mod 2 depicts the most prominent part of the referent´s appearance, as having an inlaid copper frame. Besides, the head noun explicitly points out the ontological category of the referent, namely mirror. Precise as it is, the definite description exclusively selects the only qualified referent that has passes through all the sifting. The information being encoded in the nominal per se is perfectly sufficient to uniquely identify the referent, and no previous Familiarity, or contextual inference or other pragmatic factors are required.

6.6.2 When the head noun = proper noun + common noun

One inspiring datum in this group attracts my interest. The referring expression is a nominal of a referring expression in the form Mod1 *na* CL Mod2 N like other data, but the name of the referred person (a proper noun, but not a full one) is inserted

between the Mod2 and the head noun, which generally consists of a common noun. The result is, for instance, Justin boy.

The addressee would spare the processing effort of accessing the 2 modifiers if the inserted proper name is a full name such as Justin Smith, as this kind of name is commonly prominent enough to draw the addressee´s attention and usually contains adequate enough information to uniquely locate the referent in a certain context. The thing is, in our case, the speaker inserted a first name, which is incapable of uniquely identify a referent by this proper name itself. It occurs with a modifier, which is used to identify the referent, and a common noun that points out the referent´s category. The example is (59).

(59).

跟 小 朋 友 追 逐 射 水 枪 比赛 电 动 玩 具 打 枕 头 仗 弄 得 羽絮 四飞 ,
gen xiao peng you zhui zhu she shui qiang bi sai dian dong wan ju da zhen tou zhang nong de yu xu si fei
with little friend chase shoot water gun compete video game fight pillow battle make PART feather four fly

并 跟 小 鬼 当 家 那个 窜 红 全 美 片 酬 暴 涨 的 克 金 小 鬼
Bing gen **xiao gui dang jia** *na* ge cuan hong quan mei pian chou bao zhang de ke jin xiao gui
And with **Home Alone** that CL jump popular whole America film salary explosive rise ASSOC Culkin little ghost

结 成 莫逆.
jie cheng mo ni
connect become close friends

"Michael Jackson shot water gun with kids, plays video games with them, and fight pillow battle with them, making the feathers in the pillow fly all around.

He also became close friends with the little Culkin24 who became famous across the US and got rich because of (the film) Home Alone."

This example is interesting because it shows the addressee´s preferences during referent retrieval. The referring expression we are discussing is *na ge cuan hong quan mei pian chou bao zhang de ke jin xiao gui (the little Culkin who became famous across the US and got rich because of (the film) Home Alone)*, it comes in mod 1 *na* CL mod 2 N (a tight apposition) order, and has a head noun consisting of an inserted proper noun (i.e. the name *Culkin*) and a common noun (i.e. xiao gui, an expression meaning little kid)

_____

[24] Macaulay Carson Culkin (born August 26, 1980) is an American actor. He became widely known for his portrayal of Kevin McCallister in *Home Alone*

For those addressees who have the name Culkin and the film *Home Alone* in their long-term memory, the proper name alone is enough to ensure the referent to be Familiar to those addressees. While processing the referring expression, the two already Familiar names shall be more prominent to them than other descriptive content, as they contain more distinguishable information than other words in the modifiers . If this is the case, the modifier functions as confirmation or characterisation (rather than identification) of the referent since the retrieval process has ceased once the referent has been identified.

For those addressees who have no special knowledge about Culkin or *Home Alone,* these two proper nouns do not offer more specific information than pointing out the ontological categories, i.e. the name of a person/living creature and the name of either a film, a book or something whose name needs to be in italics[25] respectively. Nevertheless, it is not necessary for them to have previous knowledge about the names neither, if they were just to uniquely identify the referent. As long as they can accept that there is a certain young actor who plays the leading role in a specific film and gets rich, the referent is Uniquely Identifable to them.

6.6.3 Summary

Source book 1 and Source book 2 have 3 and 14 data respectively. Referents in all of the data in both of the two groups are Uniquely Identifable. Therefore, it is clear from the data that referring expressions with Mod 1 + *na* CL + Mod 2 + N structure are preferably used to refer to Uniquely Identifable referents.

6.7 Summary and conclusion of the Investigation

The data of this investigation were categorised into 7 groups according to the internal structure of the nominals. The 7 groups are:

- Group 1. *na*
- Group 2. *na* N

---

[25] in the Mandarin written system, they are not marked in italic but by a pair of angle quotation marks

- Group 3. *na* CL N
- Group 4. *na* CL Adj N
- Group 5. RC *na* CL N
- Group 6. *na* CL RC N
- Group 7. RC1 *na* CL RC2 N

The reference distribution of the 7 groups of data are collected in Table 8.

Table. 8 The reference distribution of na-embedded expression in Group 1 -7

	Inf	Act	Fam	Uni	Ref	Type	Totals
na	12	11	3	5			31
na N	4	4	19	15	1	1	44
na CL N	3	5	26	10			44
na CL ADJ N			15				15
RC *na* CL N			8	13			21
na CL RC N			11	23			34
RC1 *na* CL RC2 N				17			17

In group 1, *na* functions as a pronominal on its own. The conceptual information encoded in the expression is brief, thus retrieving the referent relies largely on the context and other relevant factors apart from the encoded cognitive status. In addition to concrete entities, this pronominal *na* is used to refer to abstract concepts, which commonly do not have clear antecedent in the previous discourse. Retrieval of the referents thus involves attributive reading of the descriptive expressions. The data analysis indicates that In Focus is the most frequent CS by this referring expression, but the reference distribution ranges from In Focus to Uniquely Identifiable. The lowest CS that occurs systematically is Uniquely Identifiable, and according to the Givenness Hierarchy theory, the lowest systematically occurred CS is the one that being encoded by the referring expression, therefore, I would conclude that the pronominal *na* encodes Uniquely Identifiable. This conclusion is not in accordance with the result of GHZ (1993)´s investigation, where the data has leaded them to conclude that *na* encodes Activated when it is used alone.

Apart from Group1, *na* in the following 6 data groups occurs in nominal together with other components, and it always precedes a noun. Group 2 and 3[26] seem to be regarded as *na N* as a whole in the investigation conducted by GHZ (1993), yet they

---

[26] Possiblely this also holds for group 4-7

were treated as two separate data groups in this thesis, considering that classifiers in Mandarin can contribute to narrow down the referent set by setting constraints on the ontological categories of the referents.

In Group 2, Familiar is the most frequent CS, covering 43.18% of the data. It is followed by Uniquely Identifiable, whose referents sums up to 34.09% of the group. One suspicious datum where the referent seems to be Referential occurs, however, this alone will not lead me to conclude that Mandarin has a determiner *na* that encodes the status Referential, as the datum itself is controversial and is likely to be the result of an unnatural use. The reference of *na N* distributes from In Focus to Type Identifiable, yet the two lowest CSs only have one supporting datum each, plus the data were proved to be ambiguous and weak in arguing systematically. Hence, *na N* is considered as allowing the referent to have a CS ranging from Uniquely Identifiable to In Focus.

As for Group 3) *na* CL N, reference distribution of *na CL N* ranges from the highest In Focus to Uniquely Identifiable, but no lower than that. Uniquely Identifiable is the lowest systematically occurred CS. 59.09% of the data are used to refer to Familiar referents, almost tripling the percentage of data used for the second-ranked CS (i.e. Uniquely Identifiable). Therefore, it is good chance to say that referring expressions in *na CL N* structure commonly have the CS Familiar, but the CS that *na* encodes is Uniquely Identifiable.

All the data in Group *na* ADJ N have the CS Familiar. Data of this group was not as adequate as those in other groups, as the second source book does not contribute any example to this group. According to the 17 data from the first source book, referring expressions in *na ADJ N* are commonly used for Familiar referents. However, native speakers have confirmed that created examples in this form can be used for referents that are Uniquely Identifiable, but no lower than that. Therefore, I conclude that Uniquely Identifiable is the encoded CS of *na ADJ N*.

The components in referring expressions come in mutually reversed order In Group 5) RC *na* CL N and Group 6) *na* CL RC N. Reference distribution of Group 5 covers across Activated (with one datum), Familiar and Uniquely Identifable. The data

analysis of Group 5 indicates that referring expressions with RC *na* CL N internal structure systematically favour the CS Uniquely Identifable. The same holds for Group 6. The internal word order of the referring expressions does not seem to influence the CS, as being indicated by the data.

Referring expression in Group 7) RC1 *na* CL RC2 N are equipped with the most complicated internal structure. As a result of the richness and complexity of components, the nominal encodes the most conceptually rich information among all data groups. Referring expressions in all of the data are used for Uniquely Identifiable referents, indicating that the form *RC1 na CL RC2 N* is commonly used for referents with the CS Uniquely Identifiable.

To conclude, the investigation has shown that there are two *na*. One is a pronoun in Group 1, which encodes Uniquely Identifiable. Notably, it is a pronoun that generally occurs in the subject position of a sentence, and it rarely occurs in object position. The most frequent CS for this pronoun *na* is In Focus, it should be helpful for Mandarin learners to know that *na*, as a pronoun, is commonly used for referents that are In Focus.

Apart from Group 1) *na*, the rest 6 data groups can be seen as representing a second version of 'na', i.e. the determiner 'na'. They have shown that reference of *na*-embedded referring expressions distributes to Uniquely Identifiable as the lowest, indicating that the word encodes Uniquely Identifiable. Na is unlikely to be used for indefinite entities as the reference does not distribute to CSs lower than Uniquely Identifiable. In other word, the claim made by Brøseth & Jin (2008) of Mandarin having an indefinite determiner *na* is not supported by my investigation.

My result is in accordance with GHZ (1993) that *na N* encodes Uniquely Identifiable, if what they mean by *na N* includes the referring expression forms from Group 2 – 7. The result is also in consistency with Hedberg´s (2003) conclusion of the Mandarin distal demonstrative *na* having one of the function of the definite article property, as the data has proved that *na* shall be used for definite entities.

Chapter 7. Residual issues about the GH theory

7.1 The dilemma concerning the logical relation between CSs

Two major problems about the theory have emerged during the investigation. One is the relation between the CS Familiar and Uniquely Identifable, the other one the definition of the CS Uniquely Identifable.

If we were to draw a line to divide the six CSs on the hierarchy into two groups, the distinction line would be placed between Familiar and Uniquely Identifable, as I see it. CSs that are posited left to the line (all higher than Familiar) encode that the referents have been previously mentioned or otherwise Activated, and that the addressee does not encounter such referents for the first time or the earlier. On the contrary, referents whose CSs are right to the line can be newly introduced in the discourse by definition.

According to the claim of the GH theory, the logical relation between the six CSs is implicationally related. As GHZ (1993) stated, " IN using a particular form, the speaker thus signals that she assumes the associated cognitive status is met and, since each status entails all lower statuses, she also signals that all lower statuses (statuses to the right) have been met." (GHZ, 1993:275 – 276). For instance, a Familiar referent is by default also Uniquely Identifable. That is, if the referent has been mentioned to the addressee (and being stored in his short – term or long – term memory), this automatically means the addressee can uniquely identify the referent.

By ´Uniquely Identifable´ GHZ (1993) mean that "the addressee can identify the speaker´s intended referent on the basis of the nominal alone." (GHZ, 1993:277). Notice that they have explicitly pointed out that no other factors and information shall be involved during the retrieval, thus the addressee is expected to construct or retrieve the referent only on the basis of the referring expression.

Nevertheless, these two claims are not supported by my data. Here is one of the counterexamples:

(60).

我 继续　写，　　此　刻　我的　心情 (Sentence1)

Wo jixu　　xie　　ci　ke　　wo de xinq ing

　I　continue write t　his moment my　emotion

"I continue to write about my feelings now"

你　还记得　　那首词　吗，seg1 水 远　山　长　愁　煞　人[27]，seg2 就是这样 seg3 (Sentence2)

ni　hai ji de　***na* shou ci**　ma　　shui yuan shan chang chou　sha ren　　jiu shi zhe yang

you still remember **that CL poem**　　　water far　hill　high sadness kill people j　ust be this look

"I continue to write about my feelings at this moment, ***do you still remember that poem？ The distance has***
saddened me so much. (My feeling) is just like that."

This example has been discussed in Group *na* CL N, the relevant part of original discussion is quoted as below:

The referring expression *na shou ci* ("that CL poem") occurs in segment 1 of sentence 2. Sentence 1 narrates that the speaker continues to write about his feeling on the postcard after a short break, as to explain the context. Sentence 2 consists of three segments. In segment 1, the speaker asks the addressee if he still remembers *that poem*, and segment 2 is the quoted poem line that the speaker writes on the postcard immediately. Segment 3 explains the relevance of segment 2 by confirming that his feeling is just like what is described by the line in segment 2. Two addressees are involved. The addressee of Sentence1 is the reader of the book, as it is part of the narrated monologue; the addressee of Sentence 2 is the receiver of the postcard, as they are directly quoted from the postcard. We chose to ignore the reader´s perspective for the moment and regard the receiver of the postcard as the actual addressee, because the target referring expression occurs in sentence 2. Judging by the continuity of the context, the referent of *that poem* is the line being explicitly stated in segment 2.[28]

---

[27] From the poem 黄陵庙 （*The Huangling Temple) by* 李群玉 *(Qunyu Li, Tang Dynasty)*. The poem is: 黄陵庙前春草生，黄陵女儿茜裙新。轻舟小棹唱歌去，**水远山长愁煞人。**

[28] Theoretically and strictly, it is possible to argue that the referent of *that poem* is the poem from which the line in segment 2 is selected. I am uncertain about the speaker´s considerations when he uses *shou*, the classifier for a piece of poem as a whole. Because he could have used the more precise classifier 句(*jù*), which is only for lines and sentences. However, this seems more likely to be a casual unsatisfying usage of the classifier.

x

99

It seems plausible to claim that the referent is Familiar to the addressee based on shared personal experiences between the speaker and the addressee, because the speaker asks the question ´Do you still remember…´ immediately before the referring expression, which is the object of the question (in segment 1). The word *hai* (´still`) and the verb *ji de* (´remember´) suggest that the speaker assumes that the addressee has heard the poem before, and the question per se indicates that he is uncertain of whether the addressee can retrieve the intended referent or not. The referent is judged as being Familiar, because it meets one of the criteria for Familiar: It can be assumed to be known by the hearer through cultural/encyclopedic knowledge or shared personal experience with the speaker.

Following this sense, the referent is Familiar, and it is automatically Uniquely Identifable to the addressee among all the poems he had in his memory. Now if we recall the nature of Uniquely Identifable being described in the GH theory, the addressee should retrieve the poem line based on the nominal (i.e. *na* shou ci, that CL poem) alone. It is not likely to be true for the addressee to accept the unique identifiability of a poem line based on this nominal, and he can unlikely achieve so without other information, not to mention pragmatic reasoning and influence of the context which are inevitable.

A test can be done to confirm the dilemma by deleting segment 2 (the quoted poem) and see if the sentence will be influenced. Judging the way the speaker constructs sentence 2, segment 2 is the equivalence of the referring expression. In other words, the referent of the referring expression *na* shou ci should conincide with segment 2, they predict the same thing. The addressee shall be able to retrieve the referent since it is Familiar (regardless of being in his long or short term memeory, regardless of the amount of processing efforts needed since this is not a measurement to influence the

---

I will not take this possibility into consideration for two reasons. Firstly, the original poem as a whole contains four lines, but not each line in the poem describes the speaker´s feeling. (As he confirmed in segment 3, it is the line in segment 2 that precisely describes his feelings at that moment). Thus, he has no intention or motivation to remind the addressee of the whole poem. Secondly, the way the speaker constructs the sentence has made segment 2 an equivalence to the referring expression na shou ci. Since only one line is mentioned in segment 2, the intended referent referred by segment 1 should coincide with this one line (rather than a whole poem).

CS), and thus deleting segment 2 is supposed to have no influence on interpreting sentence 2. The new sentence becomes *´Do you still remember that poem, my current feelings are just like this.´* We have no way to confirm if the sentence is acceptable for the receiver of the postcard, as we are not the addressee, neither is the speaker. He is uncertain about whether the addressee can retrieve the referent based on this fairly brief referring expression, thus the intended referent is revealed in the immediate subsequent discourse. The new sentence is not pragmatically comprehensible to me, as the reader (like a second addressee), and it seems to be the same for the speaker. On the other hand, the sentence is perfectly understandable if segment 1 (containing the referring expression) is eliminated.

To summarize, it is rather obvious that the speaker presumes the referent to be Familiar to the addressee when the sentence is uttered, but it is also clear that such a referent will not meet the nature of Uniquely Identifable in GHZ (1993)´s sense. Hence, there is a logical dilemma for the relation between the CS Familiar and Uniquely Identifable, as the former does not entail its immediate lower CS in this datum.

7.2 Proposing the modification

Borthen (2010) encountered the example in (62) while annotating the plural form of Norwegian first-person personal pronoun, i.e. vi (´we´).

(62). "[…] – Kanskje det var de på den andre siden som sprengte den, sa hun, - eller kanskje det var oss…
- Kanskje det, sa han. – Vi var så unge da det begynte, sa hun. […]"

´[…] Maybe it was those on the other side who blew it up, she said, or maybe it was us-.
- Maybe, he said. – We were so young when it started, she said. […].´

(Borthen, 2010: 1806)

The referring expression vi ('we') has two possible interpretations. One is the set that includes the speaker (Ilin) and Ariel, who have been mentioned in the previous sentence. The other candidate is the set of people on Ilin and Ariel´s side of the river ,

which is explicitly referred to one sentence back. The right interpretation yields after the sentence has been fully processed, due to what is stated in the rest of the sentence, i.e. the set of persons referred to were so young when the problem started. As Borthen (2010) argues, the addressee knows that Ilin and Ariel were young at that time, whereas it is very unlikely that everybody on their side of the river were young. Therefore, the set that only includes the speaker (Ilin) and Ariel is the preferred interpretation. This example demonstrates that the referring expression alone is unable to uniquely identify the referent, the right interpretation is achieved through a joint power of the information encoded in the referring expression, the content of the sentence/context where the referring expression occurs, and pragmatic reasoning. Moreover, in her study, there is a very strong tendency to 'vi' to refer to sets of referents that are Uniquely Identifable or higher. These kind of cases are exceptions, if the old definition is kept. A revised definition of 'Uniquely Identifable' is, according to her, a more useful linguistic term than the old one.

I am in favor of the modification proposed by Borthen (2003), as it accounts for the significant role played by contextual inferences and other pragmatic reasoning and inevitable factors during the comprehension. As a result of the modified definition of ´Uniquely Identifiable´, the referent of *na shou ci* in (60) can be uniquely identified when the sentence has been fully processed, thus the logical relation that its actual CS Familiar entails Uniquely identifiable does not have problem anymore.

As a result of the modification, the new definition approaches the definition of the CS Referential, or in other words, makes the boundary between Uniquely Identifable even more vague, as expanding the range from the nominal itself to all the context before or after the referring expression. Concerning this, Borthen (2010) states in a footnote that: "One consequence of this definition of the status ´Uniquely Identifiable´ is that it becomes hard to differentiate it from GHZ´s definition of the status ´Referential´. My view on this issue is that the status ´Referential´ should not be part of the Givenness Hierarchy at all, but rather a feature that can be cross-classified with any of the cognitive statuses." (Borthen, 2010: 1807).

The modification justifies those referents that do not fit the original definition of Uniquely Identifable, yet is more than just Referential. In the dilemma of example (1), it solves a potential problem concerning the relation between two adjacent CSs.

7.3 Addition to ´Bridging inference´

Lastly, my data has motivated some new types of ´bridging inference´ that can be added to the criterion below:

A unique referent can be created via a 'bridging inference' by association with an already Activated referent.(e.g. A house….the front door). Previous linguists have offered new examples to this criterion, for example, ´drive – the car´ (citation). I propose two new examples concerning the retrieval of referents who fall in the ontological category of time concepts. The relevant example is listed in (61) below.

(61)
如今，在 学校　上　课　时，　她 变得 很　难 专心　听　　课。
Rujin　zai xuexiao shang ke　shi　ta　biande hen　nan zhuanxin ting　ke
Now　at school　have class when she become very hard focus　listen class
"Sophie felt it difficult to concentrate while she was having class nowadays."

最后 一 堂 课　的　下　　课 铃 响起 时　她 飞快　走 出 学校,
zuihou yi tang ke　de　xia　ke　lin xiangqi shi　ta　feikuai zou chu xuexiao
last　one CL class POSS over class bell ring　when she quickly go out　school
"When the last class was over, she rushed out of school quickly"
(… …)

苏菲　打开　信箱　时, 感觉 自己　心跳　加　　快。
Sufei　dakai xinxiang shi ganjue ziji　xintiao　jia　kuai
Sophie open　mailbox when feel　herself heart jump more quickly
"Sophie felt that her heart was jumping faster when she opened the mailbox"
(……)
当　她 关上　　园　门 时,　发现 有 一 个 大 信封　　上 写着 她的　名字
dang ta guanshang yuan　men shi　faxian　you yi ge da　xinfeng　shang xiezhe tade mingzi
when she close　garden door when discover have one CL big envelope on　write her　name
"When she closed the gate of the gardern, she saw a big envelope with her name on the cover"

(......)

苏菲　看　了　看　手表，　时间 是 两　点 四 十 五 分。

Sufei　kan　le　　kan　shoubiao shijian shi liang dian si　shi wu fen

Sophie read PART read watch　　time　be two　hour fourty five minute

"Sophpie read the watch, it was a quarter to three."

(......)

那　天　下午，　　苏菲　的　妈妈　回家　时　苏菲 仍 处于　震惊　　状态　　中。

Na　tian xiawu　　sufei　de　mama huijia　　shi　sufei　ren chuyu zhengjing zhuangtai zhong

That day afternoon Sophie POSS mum get home when Sophie still located shock　situation　　in

"That afternoon, when Sophie´s mum got home, Sophie was still in shock."

(from Group. Na N, example (2)).

The referring expression, *na tian xiawu* ("that day afternoon") occurs in the last sentence of (2). The head referent ´afternoon´ is uniquely identified through the day that includes it. And the day (that is used to ground the new referent) per se is identified through a series of chronologically-described events that take place within the same particular day. This applies to referents whose ontological category is time concept, they can be uniquely identified through the particular events that take place during the referred time peiod.

This examples motivates two new types of ´bridging inference´. One, a time period can be uniquely identified through inference by association with one or several unique events that take place during the referent time. Two, a time periode can be uniquely identified if a time concept whose duration is longer has been uniquely identified (but not necessarily vice versa). e.g day – the afternoon of the day.

104

Chapter 8. Conclusion

The Mandarin character 那 *na* is cited as corresponding to to a distal demonstrative, and a conjunction in the Mandarin dictionaries, this thesis has focus on studying what is encoded in a third *na* that can be used to assign reference.

As the result of the investigation concerning the reference distribution of the third *na*, I conclude that this third *na* should be further divided in to two *na*, a pronoun that encodes Uniquely Identifiable, and a determiner that encodes Uniquely Identifiable. My result about the determiner *na* which encodes Uniquely Identifiable agrees with GHZ´s (1993) claim that *na* in *na* N encodes Uniquely Identifiable. It has also proved Hedberg´s (2003) claim that *na* has started to have the function of a definite article in Mandarin. However, my conclusion about the pronoun *na* encoding Uniquely Identifiable does not agree with GHZ´s (1993) claim that it encodes Activated.

The result of my investigation does not support Brøseth & Jin´s (2008) claim that Mandarin has an indefinite *na*, as reference of my data has not systematically distributed to CSs that are lower than Uniquely Identifiable.

Hence, Mandarin has a distal demonstrative *na*, a conjunction *na*, a pronoun *na*, and a *na* that has started to function similar to a definite article.

Bibliography

Blakemore, Diane. "The Proposition Expressed." In *Understanding Utterances: An Introduction to Pragmatics*, by Diane Blakemore, 65-83. Wiley-Blackwell, 1992.

Borthen Kaja, Brøseth Heidi, Fretheim Thorstein. "Annotation manual for the Norwegian NP-from project." *Working Papers ISK*, May 2008: 81-125.

Borthen, Kaja. "On how we interpret plural pronouns ." *Journal of Pragmatics* 42, no. 7 (July 2010): 1799-1815.

Borthen, Kaja "Predicative NPs and the annotation of reference chains." *COLING '04 Proceedings of the 20th international conference on Computational Linguistics*. Stroudsburg: Association for Computational Linguistics Stroudsburg, PA, USA, 2004.

Brøseth Heidi, Jin Fufen. "The cognitive status Referential and its referring forms." In *Comparative Grammar and Language Acquisition in the Age of Globalization: Norwegian and Chinese*, edited by Tor A. Åfarli and Fufen Jin, 107-123. Trondheim: Tapir Akademisk Forlag, 2008.

CASS Institute of Linguistics dictionary editing team. *The Contemporary Chinese Dictionary.* The 5th Edition. Beijing: Commercial Press , 2008.

Donnellan, Keith S. "Reference and Definite Descriptions." In *Readings in the philosophy of language*, edited by Jay F. Rosenberg & Charles Travis, 195-211. Englewood Cliffs, New Jersey: Prentics-Hall, INC., 1971.

Gundel Jeanette K., Hedberg Nancy and Zacharski Ron. "Cognitive Status and the Form of Referring Expressions in Discourse." *Language* (Linguistic Society of America) 69, no. 2 (June 1993): 274-307.

Gundel, Jeanette K, Bassenea Mamadou ,Gordonb Bryan , Humnicka Linda, Khalfaouia Amel. "Testing predictions of the Givenness Hierarchy framework: A crosslinguistic investigation." *Journal of Pragmatics* 42, no. 7 (July 2010): 1770-1785.

Gundel, Jeanette K. "Relevance Theory Meets the Givenness Hierarchy. An Account of Inferrables." In *Reference and Referent Accessibility*, edited by Thorstein Fretheim & Jeanette K. Gundel, 141-153. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1996.

Gundel, Jeanette K., Hedberg, Nancy, Zacharski, Ron, Mulkern, Ann, Custis, Tonya, Swierzbin, Bonnie, Khalfoui, Amel, Humnick, Linda, Gordon, Bryan, Bassene, Mamadou, Watters, Shana,. "Coding Protocol for Statuses on the Givennes Hierarchy." May 2006.

Hedberg, Nancy. "Word Order and Cognitive Status in Mandarin." In *Reference and Referent Accessibility*, edited by Thorstein Fretheim & Jeanette K. Gundel &, 179-192. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1996.

Li Charles N, Thompson Sandra A. *Mandarin Chinese: A Functional Reference Grammar.* University of California Press, 1989.

Matthews, P.H. *The Concise Oxford Dictionary of Linguistics* . Oxford University Press, 2007.

Ming, Tao. "The Relative Position of Demonstratives and Relative Clauses in Mandarin Chinese." *the 22nd North American Conference on Chinese Linguistics (NACCL-22) & the 18th International Conference on Chinese Linguistics (IACL-18).* Cambridge: Harvard University, 2010. 323-340.

Mulkern, Ann E. "The Game of the Name." In *Reference and Referent Accessibility*, edited by Thorstein Fretheim & Jeanette K. Gundel, 235-250. Amsterdam/Philadelphia: John Benjamins Publishing Compaby, 1996.

Wilson Deirdre, Sperber Dan. "Relevance Theory." In *The Handbook of Pragmatics*, by Deirdre Wilson and Dan Sperber, edited by L.R. & Ward, G. Horn, 607-632. Oxford:, 2004.

Wilson, Deirdre. "Reference and relevance." *UCL Working Papers in Linguistics*, 1992: 165–191.

Yip, Moira. *Tone.* Cambridge University Press, 2002.

Data Source Book

贾德，乔斯坦. *苏菲的世界*. Translated by 萧宝森. 作家出版社, 2007.
Garrder, Jostein. *Sophie's World: A Novel About the History of Philosophy.* Translated by Baosen Xiao. Zuojia Publisher, 2007.

朱天文. *荒人手记* 济南: 山东画报出版社, 2009.
Zhu，Tianwen. *Notes of A Desolated Man.* Ji Nan: Shandong Magazine Publisher, 2009.