

Chapter Two

Literature Review

In Chapter two, previous studies concerning referring expressions will be presented in great detail. In the first place, Section 2.1 will present the cross-linguistic models that account for nominal form developed by Givon (1983), Brown (1983), and Ariel (1990). The respective advantages of each research will be discussed.

After presenting the models, Section 2.2 will introduce a series of studies, *Pearl Stories*, conducted by Clancy (1980). She concentrates on the correlation between referential forms and their discourse contexts. In Section 2.3 we will present the study of referential forms with hierarchical approach which is taken by Fox (1987a), who investigates different English genres. In the last part of this Chapter, we will make an attempt to review previous investigations over EFL learners' referential behavior in both English and Chinese.

2.1 Topic Continuity Scale and Accessibility Theory

Many discourse analysts, including Chafe (1976, 1987), Gundel (1978), Keenan & Comrie (1977), and so forth, have provided cognitive explanation to describe the nominal behavior across languages. Henceforth, some intriguing issues concerning the nature of nominal representation in mental processing and their interaction with that of linguistic forms arise, bearing the premise that different mental statuses are

determined by means of different nominal forms (Gundel et al, 1993).

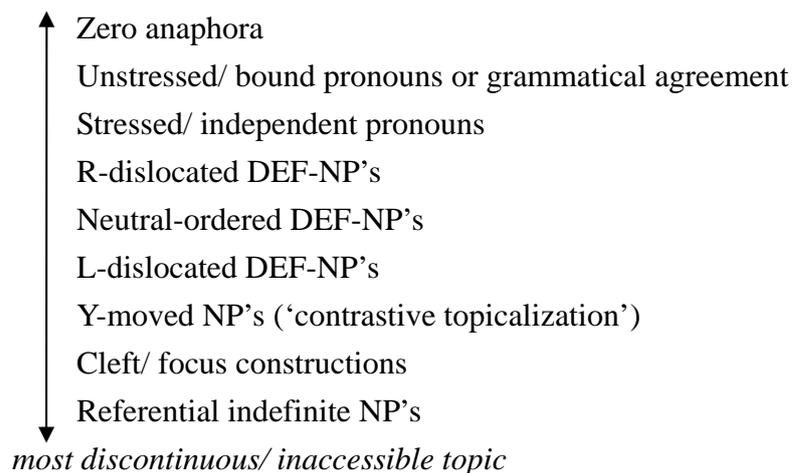
One of the prominent researchers, Chafe (1976, 1979, 1987, 1994, 1996), probes the nominal forms from various cognitive perspectives. In his opinion, factors such as given-new information, degree of accessibility, activation status, and identifiability need to be considered before an appropriate nominal form can be determined in the situational context. Furthermore, it is linguistically universal that an NP has to be reactivated or updated in regular intervals, like a number of clauses, in order to guarantee the receiver's retrieval of it. And the act of renewal is reflected on the information load of the NP itself. The general tendency of renewal is well summarized by Clancy (1992): it is conventional that at the time of utterance, the information which is thought to be 'activated' in the receiver's consciousness is prone to be realized in a weaker, more attenuated form; while more explicit forms will be used when the information is about to be activated.

Not only paving the way for cognitive approach for the research of referential form, in his studies Chafe also adopts the terms "accessibility," which is defined as "degree of activation in consciousness (Chafe, 1996: 40)," and "identifiability," which signal a representation that exists in the addressee's mind (Lambrecht, 1994: 77). Namely, the easier the retrieval of an antecedent, the more accessible the antecedent should be (Lai, 1997). This is compatible with Lyon's (1979) proposition, which

claims that accessibility reflects the “degree of salience (p. 99),” since a more salient antecedent should be easier to reclaim in the course of discourse.

Following Chafe’s approach and principle, Givon and his associates (1983) and Ariel (1988, 1990, 1996) respectively develop more systematic frameworks to explain the nominal behaviors by language users. We shall begin with the Topic Continuity Scale. In delineating the cross-linguistic topic continuity, Givon (1987) proposes the “Scale in the coding of topic accessibility,” which can be illustrated as follows.

most continuous/ accessible topic



(Givon 1983: 17)

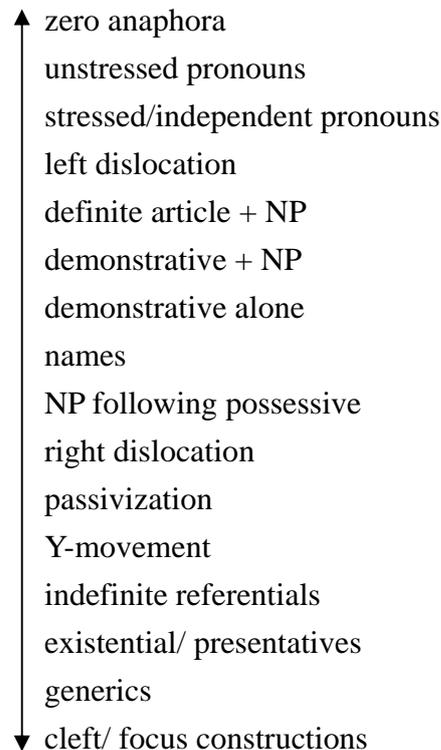
With its proposal, this model has been constantly verified across many languages, such as Japanese, Amharic, Biblical Hebrew, Hausa, Chamorro, et al in the compiled work edited by Givon (1983). According to Givon (1983) and his associates listed above, the use of referential forms should take 1) topic continuity and 2) ambiguity avoidance into consideration. The overall principle governing the use of anaphoric expressions in human languages, under Givon’s framework, consists in topic

continuity: a more continuous topic requires the use of a more reduced and accessible form such as a zero anaphora. On the other hand, a less continuous topic requires the use of more explicit and less accessible form such as full NPs. Besides, a more continuous topic is also more capable of retrieving its antecedent several clauses away than a less continuous one is. Compared with less continuous topic, a more continuous topic should persist longer and is less ambiguous. On top of that, distance also affects nominal and pronominal alternation. As the distance between a topic and its last mention increases, the topic becomes less continuous and the topic-marking mechanism needs to be adapted.

In deciding the arrangement order of his topic accessibility scale, Givon (1983) adds the insightful *Iconicity* principle, which asserts that “the more disruptive, surprising, discontinuous or hard to process a topic is, the more *coding material* must be assigned to it (p. 18).” That is to say, the information load of a referring expression should live up to the requirement of intelligibility in order for receiver’s successful interpretation, no more or less. For example, if a pronominal form is sufficient for receiver’s identification of the topical role, then there is no need for addresser to produce a definite NP.

What is of particular interest, as far as the present study is concerned, consists in the investigation on the referential forms of written English, which is provided by

Brown (1987)¹ in the same series of Givon's compiled studies. With slight modification of Givon's Scale, Brown (1983) adds the following more detailed version of topic continuity scale²:

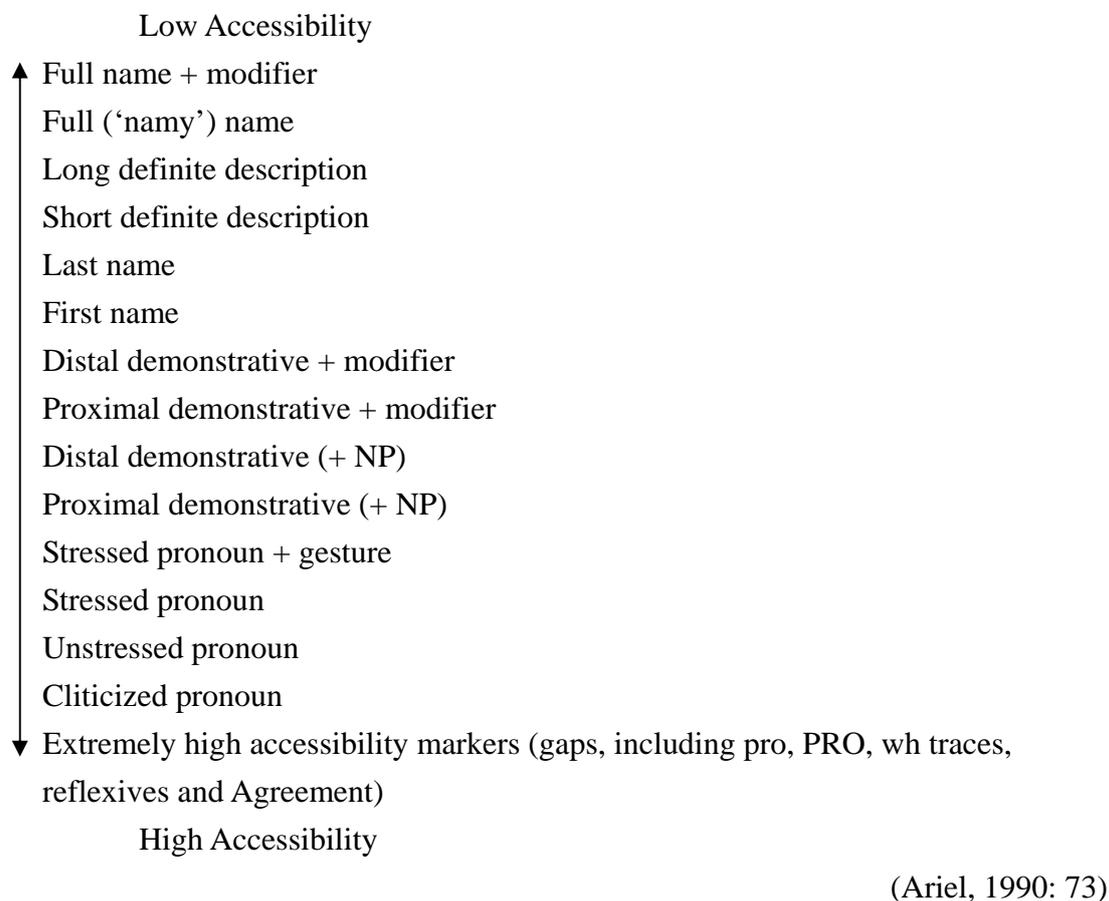


What Brown (1983) and Givon (1983) have suggested in common in their research on topic continuity is the determining factor of distance. Both suggest that more referring efforts and coding material are needed to refer to a certain entity as the distance to the last mention of it increases (Toole, 1996). Moreover, similar to Givon's device, some of the items that Brown has proposed on the scale are also topic-marking constructions such as right dislocation and passivization.

¹ The data base of Brown's research is a James Bond thriller, *Doctor No* (1958). Brown chooses this novel as his source of data because it is a narrative, and the author's purpose is to tell a story.

² In Brown's investigation, the author does not present the items on the scale as the present study does. The scale presented here is a compilation of the structures according to Brown.

Akin to the scales mentioned above, Ariel (1990, 1996) devises another framework: the Accessibility Scale, which can be illustrated as follows.



This Accessibility Scale provides a detailed depiction of conditions based on the degree of accessibility for different types of nominal forms. On the upper end are those with lower accessibility, and vice versa. That is to say, referential forms on the upper end can retrieve entities which are hard to be identified. On the lower end are forms that retrieve entities which are highly accessible.

On the top of the scale, undoubtedly, a referential form that is represented by a full name plus some additional information (i.e. modification) retrieves the least

accessible referent for the addressee. The next one goes to full name. A full name is also capable of recovering low-accessibility entity. The next one following full name on the scale is definite description. A definite description precedes last name and first name on the scale, because a definite description contains more contextual clues³ than a first name or last name does. A first name is more accessible than a last name, for “First names are shortest, they are therefore more equivocal (Ariel, 1990: 44).”

Demonstratives on the Scale are divided into four subtypes. Grammatically, English demonstratives are of two types: proximal and distal. And on the Scale they are realized as: *this*, *that*, *these* + NP, *those* + NP, promptly drawing addressee’s attention to the Physical Context. Traditionally, demonstratives have been viewed as deictic in nature (Lyons, 1977; Levinson, 1983). The interpretation of deictic form relies heavily on physical clues, such as bodily gestures, gaze, or pointing in face-to-face communication. Both Ariel (1990) and Brown (1983) make fine distinction of demonstratives because they inspect all types of NP categories, including [+human] and [-human] entities, in spoken and written form. But the primary concern in this study is the participants of the movie—the [+human] nominal phrases—in the written narratives. With this disparity, whether demonstrative will be

³ According to Ariel (1990), Gundel et al (1993), and Lambrecht (1994), during the discourse processing, for a referent to be successfully activated in addressee’s consciousness, three different ways are available: 1) Encyclopedic Knowledge Context: knowledge from the world or in one’s long-term memory; 2) Physical Context: the perception derived from the spatio-temporal environment; 3) Linguistic Context: what is being introduced into the linguistic context of communication.

included or not will be reconsidered when we devise our analytical framework in Chapter Three.

As complicated as demonstratives, pronouns in Ariel's model are divided into four subtypes. All pronominal forms, including first-, second-, and third-person pronoun, are examined by Ariel. She further suggests that whereas first- and second-person pronouns belong to the category of personal pronoun, third-person pronoun should be an independent category⁴.

To pin down the forms on the Accessibility Scale, Ariel embraces three important criteria: 1) *Informativity*, 2) *Rigidity* (i.e. Unequivocality), and 3) *Attenuation* (Phonological Size). Among them, Informativity is the most crucial factor determining where a certain referential form should be allocated on the scale. When the intended referent is thought to be hard to retrieve, the addresser will use a form of low accessibility, which is rich in information load. When the referent can be easily retrieved, the addresser can use a form of high accessibility, which is semantically empty in content. That is to say, the low-accessibility referential form contains more information load; while the high-accessibility referential form contains less information load. For example, a definite description (e.g. the young man) should be more informative than a pronoun (e. g. he) and can retrieve a less accessible referent.

⁴In real language use, there is a good reason to distinguish first- and second-person pronoun from third-person pronoun. Well said Ariel in her analysis: "first- and second-person pronouns correspond to assigned roles in conversations, while third-person pronouns refer to any person, excluding the above two (p. 47)."

The other two factors of Rigidity and Attenuation affecting the arrangement order of the scale correlate with Informativity. Rigidity refers to the power with which a referential form can disambiguate among several candidates under an equivocal circumstance. It follows that the more informative referential form is usually the more rigid one. Therefore, we can say that a first name (e.g. Sam) is more rigid than an indefinite NP (e.g. a ghost). The last criterion, Attenuation, refers to the condensing degree of a referring expression. A less attenuated form is more informative in its semantic content. For example, a full nominal phrase is less attenuated than a pronominal form and is therefore more informative.

The principles of 1) *Informativity*, 2) *Rigidity*, and 3) *Attenuation* proposed by Ariel (1990, 1996), as a matter of fact, agree with Givon's (1983) *Iconicity* tenet, as has just been mentioned above. A more informative referring expression contains more *coding material*. All these criteria will help the researcher settle down the arrangement order in designing the framework for the present study in Chapter 3.

Ariel's approach displays a unified merit (Toole, 1996), because in her framework, the effects of idiosyncratic stylistic factor in different genres can be completely ignored. Furthermore, Ariel's four dimensions of 1) distance, 2) competition, 3) saliency, and 4) unity are of particular utility in accounting for language users' choices of nominal forms. See below.

- 1) Distance: The distance between the antecedent and the anaphor (relevant to subsequent mentions only).
- 2) Competition: The number of competitors on the role of antecedent.
- 3) Saliency: The antecedent being a salient referent, mainly whether it is a topic or a non-topic.
- 4) Unity: The antecedent being within vs. without the same frame/world/point of view/segment or paragraph as the anaphor.

(Ariel, 1990: 28-29)

In general, an antecedent becomes more accessible when 1) the distance between the anaphor and its antecedent shortens, 2) there are fewer competitors between the anaphor and its antecedent, 3) the antecedent is salient, and 4) the unity between the anaphor and its antecedent is high.

In her study, not only devising a good number of referring forms, Ariel further divides all the referring expressions into three categories based on degree of accessibility, which is shown in the following Table 2.1.

Table 2.1 Ariel's accessibility degree and anaphoric forms

Degree of Accessibility	Anaphoric forms
High	deep anaphora ⁵ , third-person pronouns,
Intermediate	demonstrative pronouns, personal pronouns ⁶ ,
Low	proper names, definite descriptions,

As is illustrated in Table 2.1 above and clarified by Ariel herself, "High Accessibility

⁵ In Ariel's classification, she proposes that anaphoric expressions are of two different types: syntactically controlled vs. pragmatically controlled. For further discussion, please refer to Sag & Hankamer (1984). As far as the present study is concerned, when referring to (zero) anaphora, the author emphasizes ellipsis of a known subject.

⁶ Personal pronoun in Ariel's model refers to first- and second-person pronoun, neither of which will be inspected in the present study. As we will show in Chapter 3, no first- or second-person pronouns are produced by our subjects.

Markers imply minimal effort, while Low Accessibility Markers imply greater efforts in recovering the antecedent memory (1990: 53).” For example, a definite description can help addressee retrieve topics more easily than a pronoun does.

The framework proposed by Givon, Brown, and Ariel has been cross-linguistically testified and proved of significant value. But their models will need modification for the purpose of the present study. In Givon’s research, the primary goal is topic continuity in discourse. Some categories on Givon’s Scale are relatively marked grammatical constructions for topic identification (e.g. cleft/ focus constructions, L-dislocation DEF-NPs). They may or may not appear in our data. As for Ariel’s Accessibility Theory, her scale is devised for mixed-style data⁷. As a result, a revised framework will be needed.

Another reason for designing a revised framework is due to the fact that frameworks in previous studies are devised to account for the behaviors of all NPs, including [+human] and [-human] ones. In the present study, the author focuses on [+human] NPs, i.e. first-order entities⁸ (Lyons, 1979). Since the written data collected for this study is a synopsis of a movie, the referred NPs should be the participants in film. As has been observed in Brown’s (1983) data: “Humans are more generally the subjects of narrative (p. 323).”

⁷ The data in Ariel’s (1990) research are of various sources, including academic texts, editorial articles, interviews, short stories, oral/colloquial speech, spontaneous conversation, and so on.

⁸ In contrast with first-order entities, there are second- and third-order entities. According to Lyons (1979), first-order entities are the most basic kind of reference among the three.

Bearing in mind the need and goal of the present study, the author will propose a model that is capable of depicting advanced EFL learners' referential forms in written narratives.

2.2 Referential forms and discourse contexts

In a series of investigations of anaphoric performance based on the film *The Pear Stories*, Clancy (1980, 1992) provides penetrating cognitive explanation, and her findings are comparable to those of Givon (1983). The referring tendency, according to Clancy, is linguistically universal. Namely, the central characters, given information, and activated referents tend to be presented with more inexplicit, attenuated forms, e.g. pronoun or zero anaphor. On the other hand, the peripheral characters, new information, and inactivated referents tend to be presented with more explicit forms, e.g. full NP or complex NP. On top of that, the anaphoric forms are determined by two cognitive constrains: 1) time and 2) interference, which echo Givon's 'distance' and 'interference.' "Time" refers to the past interval between two consecutive indices of the same referent, whereas "interference" signifies the number of other interfering referents between an anaphor and the antecedent which is closest to it.

Different from Ariel's (1990, 1996) discussion, which inspects the effect of 1)

distance, 2) competition, 3) saliency, and 4) unity, Clancy (1992) declares that when inspecting referring expressions, another four predicting indices need to be considered: 1) age, 2) discourse context (Introductions, Same Subjects, and Switch Subjects), 3) plot centrality (hero vs. subordinate characters), and 4) type of narratives (picture-based vs. video-based). All these predictors are well supported through the referring behaviors shown by the subjects in Clancy's study. The indices of discourse context and plot centrality will be delineated with more details in the following paragraph. On the other hand, the predictors of age and narrative type will not be elaborated, since we only investigate English written narrative by EFL learners.

In her study on the investigation of Japanese young and adult subjects' spoken narratives, Clancy (1992) reports several critical findings. Her observation on various discourse contexts reveals that referents which are introduced for the first time are differently treated from those that are subsequently mentioned, and among the subsequent mentions in subject position, the same referent is differently treated from that of switch referent (Chafe, 1980; Dubois, 1980; Hinds, 1983). When introducing a new referent, which is often inactivated and less accessible as far as the addressee is concerned, the speaker tends to use more complex and explicit form to guarantee the addressee's retrieval of it. With the same subject being constantly mentioned, the speaker will then turn to more attenuated and implicit form to refer to it.

In addition, Clancy (1992) confirms that the effect of plot centrality derives naturally from the sociocognitive consideration. As the storyline is unveiled, the hero or the central character is the most prominent referent, and therefore does not require too much effort to be recognized by the addressee. On the other hand, the peripheral character is less likely to be constantly activated in the addressee's mind, hence demands untiring reminding from the speaker in narrating process.

In sum, after investigating Japanese EFL learners' referential behavior in their spoken narratives, Clancy (1980, 1992) has provided some intriguing findings and brilliant explanation. Additionally, owing to her defining remarks, the three discourse contexts become well-established and widespread. The contexts of Introductions, Switch Subjects, and Same Subjects designed by Clancy has been commonly adopted by subsequent researchers such as Chen (2000, 2002) and Sung (2004), whose studies will be illustrated in the forthcoming paragraphs.

The author in this study will continue using the three discourse contexts of Introductions, Same Subjects, and Switch Subjects, but with slight adjustment. The revised contexts for referential forms will be discussed in the next chapter.

2.3 Hierarchical structure in English written narratives

A further research over the distributional pattern of English anaphoric expressions is provided by Fox (1987a), whose approach is usually referred to as

hierarchical perspective or episode model (cf. e.g., Huang, 2000b; Tomlin, 1987).

After qualitatively examining the anaphora in four popular narratives⁹, Fox (1987a) proposes some interesting findings as well. Specifically, she assumes Story Grammar (Rumelhart, 1975) to be the basis in explaining the English anaphoric patterns.

Detracting the “distance” hypothesis as Givon (1983) has proposed, Fox (1987a) maintains that the long gap between two mentions of the same referent does not guarantee the use of Full NPs. Even under the context when a different character is introduced, the author can still use pronominal form to refer to the original character.

When a referential gap arises, the use of pronominal form to refer to the first character remains possible. In other words, the referential gap caused by long distance does not necessarily bring about the use of full NPs. It is episode boundary which induces full NPs.

In addition to presenting the counterexamples against Givon’s distance hypothesis, Fox (1987a) also proposes hierarchical “demarcation” in determining basic narrative units and uses it to predict the occurrence of full NPs. Rather than Givon’s results that derive from “the *overemphasis* it produces on the linear nature of texts (Fox, 1987a: 159),” Fox (1987a) claims that the author uses full NPs in order to indicate new narrative units, i.e. the beginning of developing structure. Once again,

⁹ The four popular written English narratives are: 1. *Alien* (1979), 2. *The Girl of the Sea of Cortez* (1982), 3. *Coma* (1977), and 4. *Return of the Jedi* (1983). According to Fox, fast-paced confrontations abound in these narrative fictions, and reveal significant correlation between of pronouns and full NPs.

the distance between mentions of the same referent does not seem to play any significant role in the use of full NPs. The essential determinant associated with the choice of anaphoric forms is the text structure of stories. While the Full NPs are used to indicate the slot-initial demarcation, pronominal forms are mainly used within the boundaries, regardless of the distance issue. In brief, there is no need to take distance as an influential issue in deciding what anaphoric forms to use.

Despite its explicit illustration of the data taken from English written narratives, Fox's qualitative research does not seem adequate or convincing in some ways. For instance, she denies distance as a primary consideration for the use of full NPs. Nevertheless, by inspecting a great number of data collected from several languages, Givon and his associates (1983) make significant efforts in the investigation of topic continuity, as we have mentioned. In addition, Givon does consider the influence of episode boundaries elsewhere (cf. e.g., Givon, 1992). On the other hand, unlike the various data source in Givon's research, Fox concentrates on written English narratives only. Last but not least, Fox's defining remarks over "demarcation" does not seem very perspicuous. Given a certain paragraph to be identified of its developmental structure, how exactly can the unit initial of episode be established? Bearing these flaws in mind, the author in the present study will consider the episode hypothesis as an assistant tool in our data analysis.

2.4 Referring expressions in Mandarin Chinese

In response to the tremendous interest on English nominal forms, Li and Thompson (1979) first propose the Conjoinability Hypothesis to predict the distribution of nominal forms—primarily third-person pronoun, and zero anaphora—in Chinese written discourse. After conducting an experiment of pronominal judgments and exploring two Chinese literary classics *Shui-Hu Zhuan* and *Ru Lin Wai Shi*, they reach the conclusion: “The degree of preference for the occurrence of a pronoun in a clause inversely corresponds to the degree of its conjoinability with the preceding clause (p. 330).” This general tendency suggests that full NPs usually occur in the beginning of a paragraph, pronouns are used to mark the initial position of segmentation, and zero anaphora refers to the elliptical NPs which are self-evident in the situational context (Li, 1985; Chen, 1986).

In their research, Li and Thompson (1979) also notice the prevalent use of zero anaphora in Chinese, which is quite different from the English use of zero anaphora. We just pointed out that the zero anaphora in Mandarin Chinese is a widespread device adopted by native speakers when the topic becomes familiar or highly accessible, i.e. activated in addressee’s consciousness, and understood by both speaker and hearer. In fact, in Mandarin a zero-marked subject is such an unmarked and

productive mechanism that its structural properties appear unpredictable¹⁰. On the other hand, for zero anaphora to take place in English, the syntactic context has to be a definite construction (Williams, 1988), or a coherent antecedent is required (Sag & Hankamer, 1984). Compared with Chinese zero anaphora, there are more syntactic constraints imposed on the use of zero anaphora in English. Namely, English treats zero anaphora as more of markedness than Chinese does. In view of the dissimilarities of zero anaphora between Chinese and English, it is predicted that Chinese native speakers (i.e. the advanced EFL learners) will be affected by the distribution of zero anaphora in Chinese when producing zero marker in English written narratives. And they may produce inappropriate zero marker in their written narratives.

Following Li & Thompson's research, Li (1985) and Chen (1986) dedicate further efforts in the investigation of Chinese anaphoric expressions. In their studies, the major finding is the influence of time and interference on the anaphoric choices of written narratives. Zero anaphora takes place when there is minimal time and little interference between two mentions of the same entity. Full NP occurs when the temporal span expands and more interference is involved. And pronoun is between Zero and full NP (Lin, 1992). In other words, their observation is consistent with what Clancy (1980) and Givon (1983) have claimed in previous studies. As we can see,

¹⁰ According to Li and Thompson (1979), "the nonoccurrence of anaphoric arguments in discourse must be regarded as the normal, unmarked situation (p. 322)." But despite its unpredictability, Chen (1986) identifies two cognitive constraints imposed on the use of zero anaphora in Mandarin: 1) distance to its previous mention, and 2) the interference of other referents.

there are cross-linguistic principles governing the reference management.

From a broader perspective, prior research over referential forms in Mandarin Chinese primarily aims at three ways: full NP, pronoun, and zero anaphora (cf. e.g., Chen, 1986; Christensen, 2000; Li & Thompson, 1979; Li, 1985). Such categorization for referential forms is relatively simplified when compared with that of the English language, in which a variety of referring expressions are inspected. To investigate the reference management in Mandarin, more various markers can be taken into consideration. In view of this, Lin (1992) borrows Ariel's Accessibility Theory to reexamine the Chinese referential behavior in both written and spoken narratives.

Lin's (1992) primary proposition is that Ariel's (1990) Accessibility Theory is transferable and applicable in Chinese anaphoric expression, written and spoken narrative alike. Lin (1992) states that the Chinese anaphoric forms can also be categorized into three groups based on degree of accessibility, as can be illustrated in the following Table.

Table 2.2 Lin's accessibility degree and anaphoric forms

Degree of Accessibility	Anaphoric form(s)
High	Zero anaphora, Pronouns
Mid	Determinate NPs
Low	Complex NPs, proper names, possessive NPs, bare NPs

In Lin's (1992) study, according to Table 2.2 presented above, it is reported that in Chinese narrative, high-accessibility forms are most frequent in both written and spoken narratives, low-accessibility the second, and mid-accessibility the least.

Apparently, the Chinese anaphoric expressions in Lin's model are much more various than those of previous studies (zero, full NP, and pronominal). In Lin's study, the researcher focuses on the Chinese referential forms rather than the English ones. And she has indeed made some additional contribution to the investigation of Chinese anaphoric expressions.

In this section, we have introduced studies on Chinese reference management. Li and Thompson's (1979) leading research has inspired several follow-up studies such as Li (1985) and Chen (1986) in describing Chinese referential forms. They mainly investigate the referring expressions of full NP, pronoun, and zero anaphora. Later, Lin (1992) tests the transferability of Accessibility Theory in Mandarin. Their primary contribution is that the use of referential forms needs to be based on addressee's mental state such as conjoinability or accessibility. And the contextual factors like time and interference should be considered as well. But so far, no attempt has been made to investigate Chinese EFL learners' referring expressions, let alone the comparison between Chinese and English reference management. In the next section, we will turn to this issue.

2.5 EFL learners' referential behavior

With all the previous studies on English and Chinese referring forms in either spoken or written data, more recent researchers have transferred their attention to language users and EFL learners' referring behavior in both spoken and written data.

Following Lin's (1992) investigation, Lai (1997) turns to children's use of anaphoric forms in Chinese spoken narratives. In her study, Lai (1997) devises her own analytic framework to describe the behavior of premature language users, including preschoolers, third-graders, and six-graders. The six anaphoric forms in her model are: (1) zero anaphora, (2) pronouns, (3) numeral NPs, (4) determinate NPs, (5) bare NPs, and (6) complex NPs. In her findings, Lai (1997) contends that children's choice of anaphoric forms is consistent with those of adults. In other words, more accessible antecedents are often realized by zeros and pronouns, while less accessible antecedents are retrieved by complex NPs. Furthermore, as is expected, younger learners are less competent in the manipulation of anaphoric forms. But again, the researcher has made no intention to explore Chinese speakers' reference management in English form.

Some more investigations on referential forms are conducted by Chen (2000, 2002) and Sung (2004), both aiming at EFL young learners' referring performance. While Lin (1992) and Lai (1997) emphasize the performance of Chinese anaphoric

forms, Chen (2000, 2002) and Sung (2004) make a comparison of EFL learners' referential strategies in Chinese and English. In Chen's (2000, 2004) studies, the nominal forms are divided into five types: (1) indefinite nominal, (2) definite nominal, (3) pronominal, (4) bare nominal, and (5) ad hoc. Taking a cognitive perspective of the importance of a referent in the discourse, as Clancy (1992) has done, Chen (2000) proposes that mastery of L2 referent introduction strategies developed until very late. The order of the referential forms was from maintenance skills to introduction skills. Learners have not yet fully acquired referent introduction strategies after six years of English learning. But it is worth noticing that learners have learned to use pronominals in Same Subjects Context after six years of English learning. On top of that, a U-shaped development¹¹ can be readily observed during the learning process of referent maintenance strategies in Switch Subjects Context. In sum, In Chen's (2000, 2002) studies, she highlights the developmental issue of the referential skills by primary-school children. Little has been mentioned with respect to the issue of adult learner's referential skills.

In Sung's (2004) research, similar to Chen (2000, 2002), referential forms in both languages are also coded by the same analytic model: (1) bare NPs, (2) indefinite

¹¹ By U-shaped development, Chen (2000) means: Stage 1: Learners demonstrate native-like ability of linguistic competence. => Stage 2: There is a clear drop in performance and deviant forms appear. => Stage 3: Forms with questionable grammaticality gradually disappear, and learners return to the performance which matches the norm as seen in the initial stage.

nominals, (3) determinate nominals, (4) personal nominals, (5) possessive nominals, and (6) zero anaphors (ellipses). Furthermore, like Clancy (1992) and Chen's (2002) classification, Sung encodes the discourse contexts into three: (1) introduction, (2) maintenance, and (3) switching. In her findings, Sung suggests that a) primary participants are prone to be introduced by indefinite nominals in both Chinese and English, b) children reintroduced the character by zero anaphors in Chinese and personal pronominals in English, and c) referent switching is manifested mainly by bare nominals in Chinese and determinate nominals in English.

With regard to the cause of inappropriateness on referential strategies in English narratives, Sung (2004) is in favor of the EA (Error Analysis) explanation. In the tradition of error analysis (Corder, 1981; Ellis, 1994; Robinett & Jacquelyn, 1983), at least five causes of learners' developmental error can be identified:

- 1) L1 (native language) interference: the negative effect from rules of L1,
- 2) overgeneralization: the use of a previously acquired rule in a new situation,
- 3) ignorance of rule restriction: the inability to recognize the rule reservation,
- 4) incomplete application of rules, and
- 5) false concept hypothesized.

Among the five causes, Sung (2004) declares 1) L1 interference and 2) overgeneralization of rules to be the potential grounds for referential ambiguities in children's spoken narratives. Take the lack of determiners and the different use of

ellipsis (i.e. zero anaphora) in Chinese¹² as examples. Sung (2004) thinks that it is the difference that results in the incorrect use of English bare nominals in EFL children's narrative.

Undeniably, researchers in studying EFL learners' referring expression have proposed some discerning findings with respect to preschoolers' and elementary school students' referential strategies. But the same as those analytic frameworks that account for English referential expressions, referring expressions that describe Chinese speakers or EFL learners' referring behavior in prior studies cannot be accommodated in the present study. Lin (1992) aims at the transferability of Ariel's (1990) Accessibility Scale into Mandarin, but she has failed to design a model of her own to describe the Chinese referential forms. Lai (1997) has suggested six nominal phrases which were meant for the description of young speakers' Chinese anaphoric forms, which are unlikely to be the same as those of our target subjects, namely, advanced EFL learners, and therefore cannot be taken for the present study.

Chen (2000, 2002) and Sung (2004) have devised frameworks to account for the nominal behavior of young EFL learners. Both Chen and Sung have used their own models to account for referential forms in both English and Chinese. Nevertheless, it would be unfit for the same framework to be applied in both Mandarin and English,

¹² The lack of equivalent determiners and the well-known convention of zero anaphora in Mandarin Chinese mentioned in Sung's research are based on Li and Thompson's research in 1981.

which seemed to have been taken for granted by the two researchers. As has been shown in section 2.4, the functions of zero anaphora in Mandarin and English should not be equated, or even classified under a single category. In Chinese and English, at least some of the anaphoric expressions, if not all, display distinct features, and therefore cannot be paralleled by being described with a single framework.

2.6 Summary

In the beginning of this chapter, we introduced some well-designed frameworks that are devised to describe nominal behavior in English and other languages. For example, from Givon's (1983) study on topic continuity, we learn that distance is the critical factor. When the distance between two mentions of the same referent increases, more coding material is required. The need for different amount of information load gives rise to the principle of Iconicity.

Based on Givon et al's (1983) research, Ariel (1990) then develops the Accessibility Scale, describing the use of noun phrase antecedents. Different from Givon, Ariel has turned to referring expressions rather than topic-marking constructions. Besides, in refining the degree of accessibility, Ariel suggests three groups of referring expressions: High, Mid, and Low. This classification has later been adopted by Lin (1992) in studying anaphora in written and spoken Chinese narratives.

Furthermore, we have learned the general conventions of anaphoric expressions followed by language users in the serial studies by inspecting *Pear Stories* in Clancy's (1980, 1992) investigation. The use of anaphoric forms is determined by time and interference, which are identical with Givon's distance and interference: "when referential distance is short and there is no interfering referent, hence topicality is high, a reduced anaphoric expression is typically used (Huang, 2000a: 153)."

Also, to encode the function of referential forms, Clancy (1992) proposes three different contexts of Introductions, Same Subjects, and Switch Subjects. The framework is later accepted by Chen (2000, 2002) and Sung (2004) in studying referential forms in young EFL learners' spoken narratives.

Aside from Clancy's research, we have also reviewed studies concerning anaphoric expressions in different English genres, which are contributed by Fox (1987a, 1987b, 1996). Established as hierarchical approach, Fox's episode hypothesis dismisses linear distance as a crucial determinant in anaphoric expressions. Rather, the alternation of full NP and pronominal form can be successfully predicted by means of writers' manipulation of episodic boundaries in their texts. But as we have pointed out, there seems to be some flaws in Fox's defining remarks on demarcation, and it should be refined in the methodological approach of the present study.

In the comparison and contrast of English vs. Mandarin Chinese, we have

introduced Mandarin language users' tendency and the developmental issues in referring realization, including Chinese native speakers and Chinese preschoolers' nominal forms in their output of both Mandarin and English. From Chen (2000, 2002) and Sung's (2004) studies, we find that after several years of English learning, preliminary EFL learners partially abide by the referential conventions in their spoken narrative. For example, they already learn how to use pronominals in the context of Same Subjects Context, but they have not fully acquired the Introduction strategy (Chen, 2000). In English narrative, children prefer indefinite nominals in Introductions, personal nominals in Same Subjects Context, and determinate nominals in Switch Subjects Context (Sung, 2004). Additionally, the immaturity of children's referential strategies may be attributed to L 1 interference or overgeneralization of rules.

Notwithstanding the abundant studies on referential forms as we have reviewed, little has been done in investigating the referential performance in English written narratives produced by L2 learners. Specifically, there does not seem any systematic research looking into advanced EFL learners' referential forms in written narratives. In view of such deficiency, the present study will concentrate on the investigation of the referring expressions in advanced EFL learners' written narratives.

In the next chapter, we will introduce the methodological procedure with which

EFL learners' written data are collected. With an attempt to explain learners' referring expressions, a framework will then be devised. And we will also present the steps of our data analysis.