



# Referential choice in Mandarin child language: A discourse-pragmatic perspective

Chiung-chih Huang\*

National Chengchi University, Graduate Institute of Linguistics, No. 64, Sec. 2, Zhi-nan Rd., Wenshan District, Taipei City 11605, Taiwan

## ARTICLE INFO

### Article history:

Received 22 September 2009

Received in revised form 24 November 2010

Accepted 13 December 2010

Available online 1 February 2011

### Keywords:

Reference

Referential choice

Child language

Mandarin Chinese

## ABSTRACT

This study explored Mandarin-speaking children's referential choice in natural conversation from a discourse-pragmatic perspective. The data consisted of two Mandarin-speaking children's natural conversation with their mothers, collected when the children were between the ages of 2;2 and 3;1. The subject and object arguments of the children's utterances were coded for the categories of referential forms and informativeness features. The referential forms included three categories: (1) null forms, (2) pronominal forms, and (3) nominal forms. The informativeness features included eight categories: (1) absence, (2) newness, (3) query, (4) contrast, (5) differentiation in context, (6) differentiation in discourse, (7) inanimacy, and (8) third person. The results showed that the children were sensitive to the informativeness features, and that their referential choices were made in accordance with discourse-pragmatic principles. Such sensitivity was observed since the children were as young as 2;2 and throughout their development. In addition, the results showed that the referential choices for subject and object arguments differed significantly in both children's data; the differences appeared to be associated with the asymmetry in informativeness between subject and object arguments. The results of the study are consistent with those reported in Allen (2000), Clancy (1997) and Serratrice (2005), and provide an important piece of cross-linguistic evidence for the discourse-pragmatic account for children's referential choice.

© 2010 Elsevier B.V. All rights reserved.

## 1. Introduction

Research in first language acquisition has shown that child language at the early stages of language acquisition is characterized by the omission of arguments. Children may omit the subject argument, the object argument or both in their utterances. Such a phenomenon occurs cross-linguistically, regardless of whether the target language requires overt arguments as in English (Bloom, 1970, 1990; Hyams, 1986; Valian, 1991) and Danish (Hamann and Plunkett, 1998) or permits omitted arguments as in Chinese (Wang et al., 1992), Inuktitut (Allen, 2000), Japanese (Hirakawa, 1993), and Korean (Clancy, 1993, 1997).

Different types of explanation have been proposed to account for the phenomenon of argument omission in child language. From a grammatical perspective, it has been suggested that the child starts out with a grammar that is different from the adult's. That is, the child's early grammar permits argument ellipsis where the adult's grammar would not. Later, the child's grammar would change into one more appropriate to the adult language (Hyams, 1986; Hyams and Wexler, 1993; Radford, 1990). Another type of explanation is from a performance perspective (Bloom, 1993; Valian, 1991). The performance account assumes that the child has adult-like grammatical structures from the earliest stages of language learning but omits arguments as a result of immature or limited processing resources. That is, the child can only cope with

\* Tel.: +886 2 29393091x88161; fax: +886 2 29387466.

E-mail address: [cchuang@nccu.edu.tw](mailto:cchuang@nccu.edu.tw).

producing utterances of limited length. Since the processing load of a sentence is assumed to be greater at the beginning of the sentence, subjects are omitted more frequently than objects. As the child's processing capacity matures, argument omission gradually declines until it largely disappears.

### 1.1. Discourse pragmatics and referential choice

In addition to the grammatical and performance accounts, some researchers have more recently adopted a discourse-pragmatic perspective to explain the child's referential choice; in other words, the child's referential choice may be discourse-motivated (Allen, 2000; Clancy, 1993, 1997; Guerriero et al., 2006; Narasimhan et al., 2005; Serratrice, 2005).

In language acquisition research, grammar and discourse are frequently treated as separate domains that do not interact in any significant way. However, in research on adult grammar, there is a long and flourishing tradition of theoretical approaches that consider discourse pragmatics as crucial for a comprehensive understanding of how speakers use syntax in discourse (Ariel, 1990, 1996; Chafe, 1976, 1994, 1996; Givón, 1984; Halliday and Hasan, 1976; Huang, 2000; Levinson, 1987, 1991). In this use-oriented perspective the choices speakers make are the end results of the interaction of syntactic and pragmatic principles and can only be understood by resorting to an integrated mode of explanation that draws simultaneously on both levels (Serratrice, 2005). This has important implications for the study of language development, suggesting that the acquisition of grammar may be related to the referential strategies used by adults in conversations with young children (Clancy, 1997).

It has been shown that referential strategies constitute a key link between grammar and discourse in adult language. Research indicates that adult speakers show sensitivity to discourse-pragmatic factors presumed to underlie the differential use of referring expressions in discourse (Chafe, 1994; Du Bois, 1985, 1987; Givón, 1983; Gundel et al., 1993; Kumpf, 1992). It has been suggested that informative arguments (i.e., arguments whose referents are not highly salient and accessible) are more likely to be realized overtly than uninformative arguments (i.e., arguments with highly salient and accessible referents) (Greenfield and Smith, 1976). For example, arguments with newly introduced referents, which are considered to be informative arguments, are more likely to be realized overtly than arguments with previously established referents, which are considered to be uninformative arguments. Chafe (1994) noted that during the unfolding of discourse, the accessibility of a given referent would change as a function of the level of activation state. He suggested that the choice of referring expressions is associated with three levels of activation states: active, semiactive, and inactive.<sup>1</sup> Referents with different levels of activation states are associated with different referential forms. Active referents are associated with high accessibility markers such as zero forms or unaccented pronominal forms. Semiactive or inactive referents are associated with relatively lower accessibility markers such as full noun phrases or proper names. The relationship between discourse and grammar was further explicated in Du Bois (1985, 1987). Du Bois formulated the 'Preferred Argument Structure' (PAS), which suggests that each clause contains no more than one lexical argument (the 'one lexical argument constraint'); that the lexical argument does not appear in the A role<sup>2</sup> (the 'non-lexical A constraint'); that each clause contains no more than one argument carrying new information (the 'one new argument constraint'); that new information is introduced into discourse through the non-A role, i.e., O or S, and that the A role typically carries old information (the 'given A constraint'). Gundel et al. (1993) proposed a Givenness Hierarchy to explicate the relationships between cognitive statuses and the choice of referring expressions in natural language discourse. This hierarchy consists of six cognitive statuses; from higher to lower, they are 'in focus', 'activated', 'familiar', 'uniquely identifiable', 'referential', and 'type identifiable'. The statuses are implicationally related, such that each status entails all lower statuses, but not vice versa. Each of the statuses is a necessary and sufficient condition for the use of one or more different forms. Similarly, Givón (1983) proposed an implicational hierarchy to explicate the correlation between the degree of continuity/accessibility of topic NPs and the marking devices. The scale ranges from zero anaphora as the most continuous/accessible topic to referential indefinite NPs as the most discontinuous/inaccessible topic. As seen above, the speaker's referential choice reflects the speaker's assumptions of the informative status of a given referent in the listener's mind. The correlation found between informativeness and argument realization reflects the speaker's attempt to be as explicit as possible for the listener to identify the referent in an unambiguous way.

Given the success of the discourse-pragmatic approach in explaining the choice of referring expressions in adult language, more studies have been conducted to investigate the adaptability of this approach to children's referential choice (Allen, 2000; Clancy, 1993; Clancy, 1997; Guerriero et al., 2006; Narasimhan et al., 2005; Serratrice, 2005). A similar correlation between informativeness and argument realization has been observed in child language cross-linguistically in English (Greenfield and Smith, 1976; Guerriero et al., 2006), Italian (Serratrice, 2005), Spanish (Paradis and Navarro, 2003), Korean (Clancy, 1993), Japanese (Guerriero et al., 2006), and Inuktitut (Allen, 2000). In Greenfield and Smith's (1976) seminal study, English-speaking children at the one-word stage tended to encode those aspects of event that were most informative (new information) and left unexpressed those elements that were presupposed (given information). In addition to newness, Clancy (1997) also included the features of query, contrast, and absence in her analysis of referential choice in Korean acquisition. The results showed that noun phrases were the preferred form for answering queries, for mentioning absent referents, and for introducing new referents. Both pronouns and nouns were common choices for contrasting referents. Allen (2000) added four more informativeness features in addition to the four features in Clancy (1997) to study Inuit children's

<sup>1</sup> An active referent is one that is salient in the hearer's consciousness at a particular moment in time, something that is the focus of interest. A semiactive referent is one that is in the hearer's peripheral consciousness; it is part of the background knowledge, but is not currently the focus of attention. An inactive referent is neither introduced linguistically nor is physically present.

<sup>2</sup> The A role refers to the subject of a transitive verb; the O role refers to the object of a transitive verb, and the S role refers to the subject of an intransitive verb.

argument representation: differentiation in context, differentiation in discourse, inanimacy, and third person. The results indicated that the Inuit children paid attention to discourse pragmatics in choosing whether to represent an argument as overt or null; increasing the informativeness value of a referent increased the likelihood of using an overt argument form. Similarly, Serratrice (2005) revealed that overt subjects were more likely than null subjects to represent third person, new, or ambiguous referents in Italian-speaking children's data. In addition, it was shown that there was increasing sensitivity to the informational value of referents as a function of language development. These studies indicated that children, like adults, are sensitive to the dynamics of information flow in discourse, and that their referential choices reflect their effort to reduce the potential uncertainty of the listener regarding the referents that they are talking about.

## 1.2. Argument realization in Mandarin

Mandarin is characterized by the phenomenon that both subjects and objects can be grammatically null. In general, subject referents and object referents that are understood from context do not need to be specified, as seen in the example below. As seen in B's utterance, the subject and the object are not specified because both are clear from the discourse context.

A: nei-chang dianying ni juede zenme-yang?

that-CL movie you feel how-manner

How did you feel about that movie?

B: \_\_\_\_ yidian dou bu xihuan \_\_\_\_

a: little all not like

(I) didn't like (it) a bit.

(Li and Thompson, 1981:658)

Since Mandarin does not have inflection or case markers, the pronominal system is relatively simple. The Mandarin pronominal system consists of personal pronouns and demonstrative pronouns. Demonstrative pronouns include the proximal demonstrative *zhè/zhèi* 'this' and the distal demonstrative *nà/nèi* 'that'. Personal pronouns include the first person pronouns *wǒ* 'I/me' and *wǒmen* 'we/us', the second person pronouns *nǐ* 'you (sg)' and *nǐmen* 'you (plural)', and the third person pronouns *tā* 'he/she/it/him/her' and *tāmen* 'they/them'. The Mandarin personal pronouns refer primarily to humans. Unlike English, Mandarin third person pronouns are rarely used to refer to animals and even more rarely to refer to inanimate entities. Since first and second person referents are usually clear from context, they are often associated with null forms unless there is some reason to highlight the reference to the speaker or hearer (Li and Thompson, 1981).

Mandarin NP types include bare nouns, and nouns used with demonstratives, quantifiers, or possessive constructions. It is suggested that numeral determiners (*yī* 'one') can mark newness, demonstrative ones (*nèi* 'that') givenness, but also that bare nominals (no determiner) can be used in both cases. Whatever the type of NP that is used, new information tends to be postverbal and topics must be sentence-initial, which is consistent with a universal tendency for new information to occur towards the end of utterances and for given information to occur towards the beginning (Hickmann and Hendriks, 1999).

Hickmann and Hendriks (1999) and Gundel et al. (1993) provided quantitative analyses of the proportions of the different referential forms used in Mandarin data. Hickmann and Hendriks (1999) examined narratives produced on the basis of two picture sequences by subjects of four ages (preschoolers, seven-year-olds, ten-year-olds, and adults) in four languages (English, German, French, and Mandarin Chinese). The analyses focused on all referential forms denoting the animate referents after their first mentions. Cross-linguistic variation was found. All other things being equal, the use of null forms occurred most frequently in Mandarin. In the data of the four Mandarin age groups, null forms accounted for 29%, 27%, 38%, and 31% of the total referential forms, respectively. Personal/demonstrative pronouns accounted for 30%, 18%, 26%, and 26% in the four age groups, respectively. As for nominals, the proportions were 41%, 54%, 36%, and 43%, respectively. Summing up the results for Mandarin speakers at all ages, pronominals<sup>3</sup> occurred significantly more frequently than nominals, preverbal pronominals more frequently than preverbal nominals, and post verbal nominals more frequently than post verbal pronominals.

Gundel et al. (1993) proposed six implicational related cognitive statuses to account for the use of referring expressions; they supported their proposal by reference to discourse data of five languages, including English, Japanese, Mandarin, Russian, and Spanish. The data used in the study came from a variety of spoken and written sources which differed in formality and degree of planning. These included novels, short stories, magazine articles, news broadcasts, interviews, casual conversations, and narrative file description. According to the analysis for Mandarin, zero forms were used 10.83% of the time, personal pronouns and demonstratives 17.5% of the time, and noun phrases 71.67% of the time. It was suggested that null forms, personal pronouns, and demonstratives required the referents to be at higher cognitive statuses, and that noun phrases required the referents to be at lower cognitive statuses.

As seen above, the use of zero forms in Mandarin was found to occur more frequently in the results of Hickmann and Hendriks (1999) than in those of Gundel et al. (1993); in contrast, nominals were found to occur less frequently in the results

<sup>3</sup> In Hickmann and Hendriks (1999), 'pronominals' included not only personal and demonstrative pronouns but also zero forms. However, in the present study, 'pronominals' referred to personal and demonstrative pronouns, and 'null forms' were analyzed separately, in order to derive a better understanding of argument omission in child language.

of Hickmann and Hendriks (1999) than in those of Gundel et al. (1993). The discrepancy may be due to the fact that the two studies investigated different types of discourse and analyzed different types of referents. In Hickmann and Hendriks (1999), the data were spoken data while in Gundel et al. (1993), the data consisted of both spoken and written data. In addition, in Hickmann and Hendriks (1999), only referential forms denoting the animate referents after their first mentions were analyzed while in Gundel et al. (1993), all of the referents were analyzed.

### 1.3. The present study

As seen above, Hickmann and Hendriks's (1999) cross-linguistic study included Mandarin-speaking children's data. However, the study focused on children's ability to organize cohesive anaphoric relations, by examining maintenance of reference to animate characters in narratives. Much still needs to be done to systematically investigate the referential choice of children acquiring Mandarin Chinese, especially from the discourse-pragmatic perspective. Mandarin Chinese appears to be a particularly interesting testing ground for the discourse-pragmatic account: Mandarin permits omitted arguments, and, unlike languages such as Inuktitut, Mandarin has no inflection, thus leaving no information trace at all. Moreover, since Mandarin allows omission of either subject or object arguments, it would be interesting to examine whether the referential strategies for both subject and object arguments are associated with the same discourse-pragmatic factors in Mandarin child discourse, and whether the distributions of referential forms in subject and object positions reflect any asymmetry in informativeness between subject and object arguments. Longitudinal investigations are also needed. Longitudinal data can provide more accurate information as to whether the discourse-pragmatic account can explain referential choice across developmental stages for children learning Mandarin Chinese.

To address these issues, we conducted a longitudinal study to explore Mandarin-speaking children's referential choice in natural conversation from the discourse-pragmatic perspective. We expected that argument representation in Mandarin child language is not random, but that it follows a systematic pattern, and that this pattern is predicated on informativeness features of discourse referents (Guerrero et al., 2006). In other words, we hypothesized that referential choices in early language, for both subject and object arguments, are made in accordance with the informativeness of discourse.

## 2. Methods

### 2.1. Participants and data

The participants of this study were two Mandarin-speaking girls, Lin and Jie (pseudonyms), and their mothers, who lived in the northern part of Taiwan. Lin had a younger sister and Jie was the only child. Both children's parents had received post-graduate education. The data used in this study consisted of eight hours of natural mother-child conversation video-recorded at the children's homes, with four one-hour sessions with each child. With the video-recorded data, we can observe the non-linguistic information which occurred together with the linguistic data. Lin's data were recorded at the ages of 2;2, 2;6, 2;10 and 3;1, and Jie's data were recorded at the ages of 2;2, 2;7, 2;10 and 3;1. Table 1 shows the mean length of utterance (MLU) of the children's language at each data session.

As seen in the table, both children's MLU became longer with age. However, while the children were about the same age at each data session, Lin's MLU was longer than Jie's MLU at every session: Lin's MLU at Session I was longer than that of Jie's at Session II; Lin's MLU at Session II was comparable to that of Jie's at Session III, and Lin's MLU at Session III was comparable to that of Jie's at Session IV. It appeared that Lin's language development was slightly more advanced than Jie's during the data collection period.

The communicative settings and activities of each data session are shown in Table 2. As seen in the table, all of the data were collected in the living rooms of the children's homes. The two children's data sessions included similar activities, such as eating, reading books, drawing pictures, and playing with toys.

The data collected were transcribed following the CHAT conventions and were analyzed with the CLAN program (MacWhinney, 2000).

### 2.2. Data analysis

Every utterance with an overt verb was identified for analysis. All subject and object arguments were coded for the following categories of referential forms and informativeness features.

**Table 1**  
The MLU of each data session.

| Session    | I                  | II                 | III                 | IV                 |
|------------|--------------------|--------------------|---------------------|--------------------|
| Lin<br>MLU | Lin (2;2)<br>2.635 | Lin (2;6)<br>2.802 | Lin (2;10)<br>3.168 | Lin (3;1)<br>3.541 |
| Jie<br>MLU | Jie (2;2)<br>2.046 | Jie (2;7)<br>2.353 | Jie (2;10)<br>2.800 | Jie (3;1)<br>3.115 |

**Table 2**

The communicative settings and activities of the data sessions.

| Session    | I  | II   | III  | IV  |
|------------|--|--|--|---|
| Lin        |  |  |  |   |
| Setting    | Living room  | Living room  | Living room  | Living room                                 |
| Activities | Eating<br>Reading books<br>Drawing pictures            | Eating<br>Reading books<br>Playing with toys<br>Drawing pictures | Role-playing<br>Playing games<br>Playing with toys | Eating<br>Playing games                     |
| Jie        |  |  |  |   |
| Setting    | Living room  | Living room  | Living room  | Living room                                 |
| Activities | Reading books<br>Playing with toys<br>Drawing pictures | Eating<br>Playing games<br>Playing with toys                     | Eating<br>Reading books<br>Role-playing            | Eating<br>Reading books<br>Drawing pictures |

### 2.2.1. Referential forms

While referential forms have been treated as binary in most of the previous studies, they are treated as such in different ways. Some investigators (e.g., Allen, 2000; Serratrice, 2005) analyzed referential forms in terms of overtness of the arguments, and grouped pronouns and demonstratives with lexical noun phrases as opposed to null forms (i.e., overt vs. null), whereas in other studies (e.g., Guerriero et al., 2006) pronouns and demonstratives were grouped with null forms as opposed to lexical forms (i.e., non-lexical vs. lexical). Such a difference in classification might have resulted in some of the differences in the results observed in these studies (Guerriero et al., 2006). Only Clancy (1997) used a three-way classification of referential forms, which consisted of the categories of ellipsis, pronoun, and noun. We adopted this three-way classification in the present study so as to avoid imposing a potential bias. In addition, by adopting this three-way classification, the analysis may reveal whether Mandarin pronominal forms were more similar to null forms or nominal forms in terms of their relationship with informativeness. The categories of referential forms used in this study were as follows:

- (a) Null forms: Absence of overt form.
- (b) Pronominal form: Including pronouns (e.g., *wǒ* 'I'), demonstratives (e.g., *zhè* 'this').
- (c) Nominal form: Including bare nouns (e.g., *māo* 'cat'), noun phrases (e.g., *hóngsè de huā* 'red flowers') and proper names (e.g., *Yìmíng Shúshu* 'Uncle Yiming').

### 2.2.2. Informativeness features

Following Allen (2000), we adopted a set of eight informativeness features, which have been shown to influence argument representations in many languages. These informativeness features determine how informative the children should be when referring to a referent. Each of the eight informativeness features has an informative value and an uninformative value. An informative value refers to the situation when the referent at hand is less certain (e.g., absent) and requires high informativeness in the linguistic form. In contrast, an uninformative value refers to the situation when the referent is more certain (e.g., present) and does not require high informativeness in the linguistic form. The eight informativeness features are named for the informative value of the features (e.g., Absence), and can be divided into three groups: *knowledge features*, *confusion features*, and *search-space features*. In the paragraphs below, each informativeness feature is defined based on Allen (2000) and illustrated with an excerpt of the informative value from the data of the present study.

**2.2.2.1. Knowledge features.** Knowledge features concern the presence of the referent in the joint knowledge of the speaker and the hearer, whether that knowledge derives from the physical or mental context.

- (a) Absence: The feature ABSENCE characterizes a referent that is not present in the physical context of the conversation. Since the hearer does not have knowledge of the referent from the physical context of the discourse, the identity of the referent is much less certain than it would be were the referent present in the physical context.

#### Example 1 (Lin 3;1)<sup>4</sup>

\*LIN: 今天 有 作 雪人 喔!  
 jīntiān yǒu zuò xuěrén o!  
 today there be make snowman PRT  
 '(We) made a snowman today.'

<sup>4</sup> Each utterance in the examples is presented in a set of four lines: Line 1 shows the Chinese characters; Line 2, the Pinyin Romanization; Line 3, a word-by-word gloss, and Line 4, a free translation. See Appendix A for the transcription conventions and the gloss abbreviations used in the examples.

\*LIN: 我們 老師 教 我們 作 # 作 好 漂亮  
 wǒmen lǎoshī jiāo wǒmen zuò # zuò hǎo piàoliàng  
 our teacher teach us make make very beautiful  
 的 玫瑰花 # 作成 - 個 雪人 喔。  
 de méiguīhuā # zuò -chéng yí ge xuěrén o.  
 NOM rose make-into one CL snowman PRT  
 'Our teacher taught us to make beautiful roses and to make a snowman.'

In this example, the child was talking to her mother about an event which had happened earlier that day in the kindergarten. In her utterances, the child referred to several referents which were not present in the physical context of the conversation, including *wǒmen lǎoshī* 'our teacher', *yí ge xuěrén* 'a snowman' and *hǎo piàoliàng de méiguīhuā* 'beautiful roses'.

- (b) **Newness:** The feature NEWNESS characterizes a referent that has not been previously talked about in the conversation at hand. Since the hearer has no mental knowledge of a new referent, its identity is much less certain than it would be were the referent already given in discourse. An argument is considered to be new if the referent it denotes has not been mentioned in the preceding 20 utterances.

Example 2 (Lin 2;2)

\*MOT: 海苔 有 另外 - 個 名字 啊。  
 hǎitái yǒu lìngwài yí ge míngzì a.  
 laver have another one CL name PRT  
 'There is another name for 'laver'.'

\*MOT: 叫 什麼?  
 jiào shénme?  
 call what  
 'What is it?'

\*LIN: 叔叔 家 在 哪裡? ←  
 shúshu jiā zài nǎlǐ?  
 uncle home at where  
 'Where's Uncle's home?'

In this example, the child was eating some laver. The mother asked the child a question about laver; however, the child did not answer the mother's question. Instead, the child changed the topic and talked about *shúshu jiā* 'Uncle's home'. The referent *shúshu jiā* 'Uncle's home' had not previously been referred to in the conversation, and so can be considered to be a newly introduced referent.

- (c) **Query:** The feature QUERY characterizes a referent that is the subject of or response to a question. Since the referent is either not yet identified or newly identified, the listener has little mental knowledge of this referent, and thus its identity is much less certain than it would be were the referent already given in discourse.

Example 3 (Lin 2;2)

\*MOT: 這 是 誰 [% 指著書本封面上的獅子]?  
 zhè shì shéi [% pointing at a lion on the cover of a book]?  
 this is who  
 'Who is this?'

\*LIN: 獅子. ←  
 shīzi.  
 lion  
 '(It's a) lion.'

In Example 3, the mother and the child were reading a story book. The mother was pointing at a lion on the cover of the book, and asked the child who it was. The child's utterance *shīzi* 'lion' constituted a response to the mother's question.

2.2.2.2. *Confusion features.* Confusion features concern the resolution of potential confusion about the identity of a referent when various potential referents are present either explicitly or implicitly in the discourse or the physical context.

- (d) **Contrast:** The feature CONTRAST characterizes a referent the speaker is explicitly contrasting with other potential referents in the discourse or in the shared physical or mental context, usually through tone of voice, gesture, or other contextual means.

## Example 4 (Jie 2;2)

\*JIE: 媽咪 褲子 破掉。  
 māmī kùzi pò-diào.  
 Mommy trousers worn out  
 'Mommy's trousers are worn out.'

\*MOT: +^+^對 啊。  
 duì a.  
 yes PRT  
 'Yeah'

\*MOT: 媽媽 沒 有 錢錢 買。  
 māmā méi yǒu qiánqián mǎi.  
 Mom not have money buy  
 'Mommy doesn't have (any) money to buy (any).'

\*MOT: 那 Jie 你 有 沒有 錢?  
 nà Jié nǐ yǒu-méi-yǒu qián?  
 then Jie you have-not-have money  
 'Then, do you have (some) money, Jie?'

\*MOT: 你 要 不要 幫 媽媽 買 褲子?  
 nǐ yào-bú-yào bāng māmā mǎi kùzi?  
 you want-not-want help Mom buy trousers  
 'Do you want to buy (some) trousers for Mommy?'

\*JIE: 爸爸 有 錢 [%手指爸爸]. ←  
 bàba yǒu qián [% pointing at the father].  
 Dad have money  
 'Daddy has money.'

\*MOT: 0 [=! laughing].

\*MOT: 爸爸 有 錢錢 [=! 笑].  
 bàba yǒu qiánqián [=! laughing].  
 Dad have money  
 'Daddy has money.'

In this example, the child noticed that the mother's trousers were worn out. The mother jokingly asked the child whether the child had money to buy trousers for her. However, the child pointed at the father, saying that the father had money. The child used the gesture of pointing to contrast the father with the child herself.

- (e) Differentiation in context: The feature DIFFERENTIATION IN CONTEXT characterizes a referent that is one of two or more referents in the immediate physical context that could fit the verb semantics and the identifying elements of the argument in question. Since there is more than one potential referent in the physical context fitting the characteristics of the argument, there is potential uncertainty on the part of the hearer in identifying the target referent.

## Example 5 (Jie 2;7)

\*MOT: 這 是 帆船鞋。  
 zhè shì fánchuánxié.  
 this be boat-shoes  
 'These are boat-shoes.'

\*MOT: 這 是 休閒鞋。  
 zhè shì xiūxiánxié.  
 this be sneakers  
 'These are sneakers.'

\*JIE: 我 不 要 休閒鞋. ←  
 wǒ bú yào xiūxiánxié.  
 I not want sneakers  
 'I don't want sneakers.'

\*JIE: 我 要 這個 鞋. ←  
 wǒ yào zhè-ge xié.  
 I want this-CL shoe  
 'I want these shoes.'

In this example, there were two focused referents in the immediate physical context, i.e., *fánchuánxié* 'boat-shoes' and *xīuxiánxié* 'sneakers'. As seen in Lines 3 and 4, the child needed to differentiate the intended referent from the other potential referent in the physical context in order to mention one of the two competing referents in the context.

- (f) Differentiation in discourse: The feature DIFFERENTIATION IN DISCOURSE characterizes a referent that is one of two or more referents already established in the discourse (i.e., in the five preceding utterances, following Givón, 1983) that could fit the verb semantics and identifying elements of the argument in question. Since there is more than one potential referent in the discourse context fitting the characteristics of the argument, there is potential uncertainty on the part of the hearer in identifying the target referent.

Example 6 (Jie 2;2)

\*MOT: Jie    你    要    爸爸 講    故事    還是  
 Jié    nǐ    yào    bàba jiǎng gùshì hái shì  
 Jie    you    want    Dad    read    story    or

媽媽 講    故事?  
 māma jiǎng gùshì?  
 Mom    read    story

'Jie, do you want Daddy or Mommy to read the story?'

\*JIE: Jie.    ←  
 Jié.  
 Jie  
 'Jie.'

\*MOT: 你    要    講    喔.  
 nǐ    yào    jiǎng    o.  
 you    want    read    PRT  
 'You want to read (it).'

\*MOT: 那    你    講.  
 nà    nǐ    jiǎng.  
 then    you    read  
 'Then you read (it).'

In the mother's utterance in Line 1, three person referents were introduced into the discourse: the child, the mother, and the father. In the child's response, the three person referents constituted three potential referents in the discourse context. In order to mention the intended referent, i.e., the child herself, the child needed to differentiate the intended referent from the other potential referents in the discourse context.

2.2.2.3. *Search-space features.* Search-space features concern differences in the relative size of the search space one must consider to find the referent in question.

- (g) Inanimacy: The feature INANIMACY characterizes referents that are not animate. In typical child discourse, the number of animate entities is relatively limited (e.g., child, mother, father, sibling, and dog) compared to the vast number of inanimate entities (e.g., table, cup, toy, juice, television, plant, and clothes). Thus, the search space for animate referents is relatively small, while the search space for inanimate referents is relatively much larger.

Example 7: (Lin 2;6)

\*LIN: 我    想    吃    糖果.    ←  
 wǒ    xiǎng    chī    tángguǒ.  
 I    want    eat    candy  
 'I want to eat (some) candies.'

\*MOT: 話    還    沒    講    完.  
 huà    hái    méi    jiǎng    wán.  
 words    still    not    talk    finish  
 '(We) haven't finished (our) talk yet.'

\*MOT: 講    完    才    可以    吃.  
 jiǎng    wán    cái    kǎyǐ    chī.  
 talk    finish    only    can    eat  
 '(You) can't eat (the candies) until (we) finish (our) talk.'



**Table 3**  
Informativeness features (Allen, 2000:490).

| Pragmatic features           | Informative value                                      | Uninformative value                                |
|------------------------------|--|--|
| Absence                      | Referent absent from physical context                  | Referent present in physical context               |
| Contrast                     | Contrast emphasized between potential referents        | No contrast emphasized between potential referents |
| Differentiation in context   | Two or more potential referents in physical context    | Only one potential referent in physical context    |
| Differentiation in discourse | Two or more potential referents in preceding discourse | Only one potential referent in preceding discourse |
| Inanimacy                    | Inanimate referent                                     | Animate referent                                   |
| Newness                      | Referent new to discourse                              | Referent not new to discourse                      |
| Query                        | Referent subject of or answer to query                 | Referent not subject of or answer to query         |
| Third person                 | Third person referent                                  | First or second person referent                    |

In Example 7, the mother and the child were reading a picture book and were talking about some of the pictures in the book. In Line 1, the child asked for some candies; the intended referent 'candies' in the child's utterance was an inanimate referent.

- (h) Third person: The feature THIRD PERSON characterizes a referent that is not first or second person. In typical child discourse, the number of first and second person entities is relatively limited compared to the vast number of potential third person entities. Thus, the search space for first and second person referents is relatively small, but the search space for third person referents is relatively much larger.

Example 8 (Jie 2;7)

\*MOT: 那 是 誰 送 你 的?  
 nà shì shéi sòng nǐ de?  
 that be who send you NOM  
 'Who gave you that?'

\*JIE: Yubin 哥哥. ←  
 Yùbīn gēge.  
 Yubin brother  
 'Big Brother Yubin.'

\*MOT: Yubin 哥哥 喔.  
 Yùbīn gēge o.  
 Yubin brother PRT  
 'Oh, Big Brother Yubin.'

In this example, the mother and the child were talking about a gift the child had received. The mother asked the child who had given her the gift. The child referred to *Yùbīn gēge* 'Big brother Yubin', which was a third-person referent.

As mentioned above, each informativeness feature has two values: an informative value and an uninformative value. The informative and uninformative values for each of the features (in alphabetical order) are summarized in Table 3.

The data were coded by a trained research assistant, who was a native speaker of Mandarin and a graduate student of linguistics. In addition, one and half hours of data from each child were randomly selected and were independently coded by another trained research assistant, who was also a native speaker of Mandarin and a graduate student of linguistics. Cohen's Kappa was used to determine the inter-rater reliability. The reliability for the coding of referential forms was 90%, and the reliability for the coding of informativeness features was 92%.

### 3. Results

Table 4 demonstrates the numbers of referential forms in the two children's speech. As seen in the table, the total number of referential forms in Lin's data was 1326 and the number in Jie's data was 2362. The proportions of the three referential forms ranged from 26% to 38% in Lin's data and from 27% to 38% in Jie's data.

**Table 4**  
Numbers of referential forms used by the two children.

|            | Lin  |        | Jie  |        |
|------------|------|--------|------|--------|
|            | N    | %      | N    | %      |
| Null       | 353  | 26.62  | 828  | 35.06  |
| Pronominal | 472  | 35.60  | 887  | 37.55  |
| Nominal    | 501  | 37.78  | 647  | 27.39  |
| Total      | 1326 | 100.00 | 2362 | 100.00 |

The children's data were also analyzed in relation to informativeness. Table 5 shows the numbers of informative and uninformative arguments with respect to each informativeness feature in Lin's and Jie's data. Both children's data contained much more uninformative arguments than informative arguments for each informativeness feature, except for the feature of Third person in both children's data, and Inanimacy in Jie's data.

### 3.1. Referential choice and informativeness

Further analysis was conducted to examine the relationship between the children's use of referential forms and informativeness. The results in the eight tables from Tables 6–13 demonstrate the analyses with respect to the eight informativeness features.

Table 6 shows the distributions of referential forms with respect to the feature of Absence in Lin's and Jie's data. The referential forms in the data were examined in terms of the informative value and the uninformative value of Absence: informative arguments refer to absent referents while uninformative arguments refer to present referents. As seen in Table 6, the distributions of referential forms for absent referents and for present referents revealed very different patterns of use. When referring to absent referents, the children used a high rate of nominal forms (74.39% in Lin, 62.04% in Jie); when referring to present referents, the percentage of nominal forms became much lower (29.44% in Lin, 23.90% in Jie). In contrast, both children used null forms and pronominal forms to refer to present referents more frequently than they used these forms to refer to absent referents. Chi-square analyses showed that the referential choices for absent referents and present

**Table 5**  
Numbers of informative and uninformative arguments with respect to each informative feature.

| Pragmatic features           | Informative |       | Uninformative |       |
|------------------------------|-------------|-------|---------------|-------|
|                              | No          | %     | No.           | %     |
| <b>Lin</b>                   |             |       |               |       |
| Absence                      | 246         | 18.55 | 1080          | 81.45 |
| Newness                      | 332         | 25.04 | 994           | 74.96 |
| Query                        | 182         | 13.73 | 1144          | 86.27 |
| Contrast                     | 18          | 1.36  | 1308          | 98.64 |
| Differentiation in context   | 141         | 10.63 | 1185          | 89.37 |
| Differentiation in discourse | 320         | 24.13 | 1006          | 75.87 |
| Inanimacy                    | 543         | 40.95 | 783           | 59.05 |
| Third person                 | 793         | 59.80 | 533           | 40.20 |
| <b>Jie</b>                   |             |       |               |       |
| Absence                      | 216         | 9.14  | 2146          | 90.86 |
| Newness                      | 400         | 16.93 | 1962          | 83.07 |
| Query                        | 362         | 15.33 | 2000          | 84.67 |
| Contrast                     | 43          | 1.82  | 2319          | 98.18 |
| Differentiation in context   | 152         | 6.44  | 2210          | 93.56 |
| Differentiation in discourse | 212         | 8.98  | 2150          | 91.02 |
| Inanimacy                    | 1240        | 52.50 | 1122          | 47.50 |
| Third person                 | 1630        | 69.01 | 732           | 30.99 |

**Table 6**  
Distributions of referential forms with respect to Absence.

|            | Absent |        | Present |        | $\chi^2$  | Post hoc |
|------------|--------|--------|---------|--------|-----------|----------|
|            | N      | %      | N       | %      |           |          |
| <b>Lin</b> |        |        |         |        |           |          |
| Null       | 37     | 15.04  | 316     | 29.26  | 175.49*** | A < P    |
| Pronominal | 26     | 10.57  | 446     | 41.30  |           | A < P    |
| Nominal    | 183    | 74.39  | 318     | 29.44  |           | A > P    |
| Total      | 246    | 100.00 | 1080    | 100.0  |           |          |
| <b>Jie</b> |        |        |         |        |           |          |
| Null       | 51     | 23.61  | 777     | 36.21  | 147.13*** | A < P    |
| Pronominal | 31     | 14.35  | 856     | 39.89  |           | A < P    |
| Nominal    | 134    | 62.04  | 513     | 23.90  |           | A > P    |
| Total      | 216    | 100.00 | 2146    | 100.00 |           |          |

\*\*\*  $p < .001$ .

**Table 7**  
Distributions of referential forms with respect to Newness.

|            | New |        | Old  |        | $\chi^2$  | Post hoc |
|------------|-----|--------|------|--------|-----------|----------|
|            | N   | %      | N    | %      |           |          |
| Lin        |     |        |      |        |           |          |
| Null       | 10  | 3.01   | 343  | 34.51  | 281.81*** | N < O    |
| Pronominal | 72  | 21.69  | 400  | 40.24  |           | N < O    |
| Nominal    | 250 | 75.30  | 251  | 25.25  |           | N > O    |
| Total      | 332 | 100.00 | 994  | 100.00 |           |          |
| Jie        |     |        |      |        |           |          |
| Null       | 7   | 1.75   | 821  | 41.85  | 294.73*** | N < O    |
| Pronominal | 171 | 42.75  | 716  | 36.49  |           | n.s.     |
| Nominal    | 222 | 55.50  | 425  | 21.66  |           | N > O    |
| Total      | 400 | 100.00 | 1962 | 100.00 |           |          |

n.s.: not significant.

\*\*\*  $p < .001$ .

**Table 8**  
Distributions of referential forms with respect to Query.

|            | Query |        | Non-query |        | $\chi^2$  | Post hoc |
|------------|-------|--------|-----------|--------|-----------|----------|
|            | N     | %      | N         | %      |           |          |
| Lin        |       |        |           |        |           |          |
| Null       | 4     | 2.20   | 349       | 30.51  | 128.24*** | Q < N    |
| Pronominal | 36    | 19.78  | 436       | 38.11  |           | Q < N    |
| Nominal    | 142   | 78.02  | 359       | 31.38  |           | Q > N    |
| Total      | 182   | 100.00 | 1144      | 100.00 |           |          |
| Jie        |       |        |           |        |           |          |
| Null       | 3     | 0.83   | 825       | 41.25  | 350.44*** | Q < N    |
| Pronominal | 128   | 35.36  | 759       | 37.95  |           | n.s.     |
| Nominal    | 231   | 63.81  | 416       | 20.80  |           | Q > N    |
| Total      | 362   | 100.00 | 2000      | 100.00 |           |          |

n.s.: not significant.

\*\*\*  $p < .001$ .

**Table 9**  
Distributions of referential forms with respect to Contrast.

|            | Contrastive |        | Non-contrastive |        | $\chi^2$ | Post hoc |
|------------|-------------|--------|-----------------|--------|----------|----------|
|            | N           | %      | N               | %      |          |          |
| Lin        |             |        |                 |        |          |          |
| Null       | 2           | 11.11  | 351             | 26.83  | n.s.     |          |
| Pronominal | 10          | 55.56  | 462             | 35.32  |          |          |
| Nominal    | 6           | 33.33  | 495             | 37.84  |          |          |
| Total      | 18          | 100.00 | 1308            | 100.00 |          |          |
| Jie        |             |        |                 |        |          |          |
| Null       | 0           | 0.00   | 828             | 35.71  | 44.69*** | C < N    |
| Pronominal | 13          | 30.23  | 874             | 37.69  |          | n.s.     |
| Nominal    | 30          | 69.77  | 617             | 26.61  |          | C > N    |
| Total      | 43          | 100.00 | 2319            | 100.00 |          |          |

n.s.: not significant.

\*\*\*  $p < .001$ .

**Table 10**

Distributions of referential forms with respect to Differentiation in context.

|            | Differentiation in context |        | Non-differentiation in context |        | $\chi^2$  | Post hoc |
|------------|----------------------------|--------|--------------------------------|--------|-----------|----------|
|            | N                          | %      | N                              | %      |           |          |
| Lin        |                            |        |                                |        |           |          |
| Null       | 2                          | 1.42   | 351                            | 29.62  | 136.77*** | D < N    |
| Pronominal | 23                         | 16.31  | 449                            | 37.89  |           | D < N    |
| Nominal    | 116                        | 82.27  | 385                            | 32.49  |           | D > N    |
| Total      | 141                        | 100.00 | 1185                           | 100.00 |           |          |
| Jie        |                            |        |                                |        |           |          |
| Null       | 1                          | 0.66   | 827                            | 37.42  | 285.57*** | D < N    |
| Pronominal | 20                         | 13.16  | 867                            | 39.23  |           | D < N    |
| Nominal    | 131                        | 86.18  | 516                            | 23.35  |           | D > N    |
| Total      | 152                        | 100.00 | 2210                           | 100.00 |           |          |

\*\*\*  $p < .001$ .**Table 11**

Distributions of referential forms with respect to Differentiation in discourse.

|            | Differentiation in discourse |        | Non-differentiation in discourse |        | $\chi^2$  | Post hoc |
|------------|------------------------------|--------|----------------------------------|--------|-----------|----------|
|            | N                            | %      | N                                | %      |           |          |
| Lin        |                              |        |                                  |        |           |          |
| Null       | 13                           | 4.06   | 340                              | 33.80  | 312.39*** | D < N    |
| Pronominal | 54                           | 16.88  | 418                              | 41.55  |           | D < N    |
| Nominal    | 253                          | 79.06  | 248                              | 24.65  |           | D > N    |
| Total      | 320                          | 100.00 | 1006                             | 100.00 |           |          |
| Jie        |                              |        |                                  |        |           |          |
| Null       | 2                            | 0.94   | 826                              | 38.42  | 363.43*** | D < N    |
| Pronominal | 35                           | 16.51  | 852                              | 39.63  |           | D < N    |
| Nominal    | 175                          | 82.55  | 472                              | 21.95  |           | D > N    |
| Total      | 212                          | 100.00 | 2150                             | 100.00 |           |          |

\*\*\*  $p < .001$ .**Table 12**

Distributions of referential forms with respect to Inanimacy.

|            | Inanimate |        | Animate |        | $\chi^2$  | Post hoc |
|------------|-----------|--------|---------|--------|-----------|----------|
|            | N         | %      | N       | %      |           |          |
| Lin        |           |        |         |        |           |          |
| Null       | 118       | 21.73  | 235     | 30.01  | 112.58*** | I < A    |
| Pronominal | 129       | 23.76  | 343     | 43.81  |           | I < A    |
| Nominal    | 296       | 54.51  | 205     | 26.18  |           | I > A    |
| Total      | 543       | 100.00 | 783     | 100.00 |           |          |
| Jie        |           |        |         |        |           |          |
| Null       | 369       | 29.76  | 459     | 40.91  | 116.38*** | I < A    |
| Pronominal | 415       | 33.47  | 472     | 42.07  |           | I < A    |
| Nominal    | 456       | 36.77  | 191     | 17.02  |           | I > A    |
| Total      | 1240      | 100.00 | 1122    | 100.00 |           |          |

\*\*\*  $p < .001$ .

**Table 13**  
Distributions of referential forms with respect to Third person.

|            | Third person |        | Non-third person |        | $\chi^2$  | Post hoc |
|------------|--------------|--------|------------------|--------|-----------|----------|
|            | N            | %      | N                | %      |           |          |
| Lin        |              |        |                  |        |           |          |
| Null       | 162          | 20.43  | 191              | 35.83  | 404.07*** | T < N    |
| Pronominal | 160          | 20.18  | 312              | 58.54  |           | T < N    |
| Nominal    | 471          | 59.39  | 30               | 5.63   |           | T > N    |
| Total      | 793          | 100.00 | 533              | 100.00 |           |          |
| Jie        |              |        |                  |        |           |          |
| Null       | 508          | 31.17  | 320              | 43.72  | 322.84*** | T < N    |
| Pronominal | 497          | 30.49  | 390              | 53.28  |           | T < N    |
| Nominal    | 625          | 38.34  | 22               | 3.01   |           | T > N    |
| Total      | 1630         | 100.00 | 732              | 100.00 |           |          |

\*\*\*  $p < .001$ .

referents were significantly different in both children's data ( $\chi^2(2) = 175.49, p < .001$  in Lin;  $\chi^2(2) = 147.13, p < .001$  in Jie), suggesting that the children were sensitive to the feature of Absence in their referential choices.

In order to understand which form(s) used by the children contributed to the significant differences, a post hoc multiple comparison test was conducted (Marascuilo and McSweeney, 1977). The results showed that (1) a significantly lower percentage of null forms were used for absent referents than for present referents in both children's data (15.04% < 29.26% in Lin, and 23.61% < 36.21% in Jie), (2) a significantly lower percentage of pronominal forms were used for absent referents than for present referents in both children's data (10.57% < 41.30% in Lin, and 14.35% < 39.89% in Jie), and (3) a significantly higher percentage of nominal forms were used for absent referents than for present referents in both children's data (74.39% > 29.44% in Lin, and 62.04% > 23.90% in Jie). Thus, the children's use of all the three types of referential forms contributed to the significant differences observed in the Chi-square analyses.

Similar distribution patterns and statistical results were also observed in the tables concerning the other informativeness features. As seen from Tables 7–13, all the Chi-square analyses reached statistical significance, except for Lin's data with respect to Contrast (Table 9). The results indicated that both children's referential choices were highly influenced by the informativeness features examined. Furthermore, the post hoc multiple comparison tests for the Chi-square analyses (Marascuilo and McSweeney, 1977) showed that Lin's data revealed rather consistent patterns in the use of referential forms. For all of the informativeness features analyzed, we observed that (1) a significantly lower percentage of null forms were used for referents with informative values (e.g., new, contrastive) than for referents with uninformative values (e.g., old, non-contrastive), (2) a significantly lower percentage of pronominal forms were used for referents with informative values than for referents with uninformative values, and (3) a significantly higher percentage of nominal forms were used for referents with informative values than for referents with uninformative values. That is, when Lin referred to a referent with an informative value, (i.e., a referent which was less certain), she tended to use a nominal form to provide the required high informativeness. In contrast, when she referred to a referent with an uninformative value (i.e., a referent which was more certain), a null form or a pronominal form would usually be the choice. Thus, Lin used null forms and pronominal forms in a similar way, which was distinct from the way in which she used nominal forms. However, Jie's data revealed a slightly different picture. The post hoc multiple comparison tests of Jie's data demonstrated that five of the features (i.e., Absence, Differentiation in context, Differentiation in discourse, Inanimacy, and Third person) presented patterns consistent with those observed in Lin's data. However, three of them (i.e., Newness, Query, and Contrast) showed different patterns regarding the use of pronominal forms. It was observed that for these three informativeness features, Jie's use of pronominal forms did not show significant differences between the informative values and the uninformative values; the results of significance observed in the Chi-square analyses were resulted from the distinctive distribution patterns of null forms and nominal forms.

As seen above, the results seemed to suggest that Jie did not use pronominal forms to differentiate the informative values from the uninformative values of Newness, Query, and Contrast. Since informative values usually require high informativeness in linguistic forms, further analysis was conducted to examine why Jie referred to these new, queried, or contrastive referents with pronominal forms, which were less specific and informative than nominal forms, and whether Jie's uses of these pronominal arguments were communicatively effective. A closer look at Jie's data revealed that (1) the majority of these pronominal forms were demonstratives; (2) these pronominal arguments usually represented referents which were present in the situational context, and (3) these pronominal arguments usually were accompanied by the use of non-linguistic strategies, such as deictic gestures (e.g., pointing, touching, and reaching), or eye gaze, to indicate the intended referents. Thus, it appeared that Jie in fact was sensitive to the lower specificity and informativeness of pronominal forms, and the need to supplement pronominal forms with non-linguistic information. With the non-linguistic strategies, these pronominal arguments were usually communicatively effective. In other words, Jie's use of pronominal forms also reflected

her sensitivity to the informativeness features. Examples 9, 10, and 11 illustrate Jie's use of pronominal forms and non-linguistic strategies to refer to new referents, queried referents and contrastive referents, respectively.

Example 9: (Jie 2;2)

\*MOT: 好 那 你 唱 給 媽 媽 聽。  
hǎo nà nǐ chàng gěi māma tīng.  
ok then you sing to Mom listen  
'Ok, then you sing for Mommy.'

\*MOT: 你 唱 歌。  
nǐ chàng gē.  
you sing song  
'You sing a song.'

\*MOT: <你 要 唱 > [>] +...  
<nǐ yào chàng> [>] +...  
you want sing  
'You want to sing...'

\*JIE: <這 是 > [<] 妙 妙 [% 翻開書, 看著書上的小女孩]  
<zhè shì > [<] Miàomiào [% opening a book and looking at a girl in the book].  
this be Miaomiao  
'This is Miaomiao.'

\*MOT: 哪 - +...  
nǎ yí +...  
which one  
'Which one...'

\*MOT: 喔 這 是 妙 妙。  
o zhè shì miàomiào.  
PRT this be Miaomiao  
'Oh, this is Miaomiao.'

Just prior to this excerpt, Jie had said to the mother that she wanted to sing a song. However, as seen in the excerpt, before the mother had finished asking Jie what song she wanted to sing, the child changed the topic in Line 4. The subject of the sentence in Line 4 was a pronominal form *zhè* 'this', which introduced a new referent. The pronominal form was used with the child's action of opening a book and looking at a girl in the book. The mother's immediate response in Line 5 'Which one...' indicated the inherent low informativeness of the pronominal form. However, when the mother followed the child's eye gaze, she was then able to identify the intended referent.

Example 10: (Jie 2;2)

\*MOT: 奇 奇 在 吃 什 麼 ?  
qíqí zài chī shénme?  
Qiqi DUR eat what  
'What is Qiqi eating?'

\*JIE: 奇 奇 在 吃 這 個 [% 手 指 書 上 的 牛 角 麵 包 ] ←  
qíqí zài chī zhè-ge [% pointing at a croissant in the book].  
Qiqi DUR eat this-CL  
'Qiqi is eating this.'

In Example 10, the mother and the child were reading a picture book. In Line 1, the mother asked Jie what Qiqi, a character in the book, was eating. To answer the mother's question, Jie used a pronominal form *zhè-ge* 'this', with a gesture of pointing at a

croissant in the book, to indicate that Qiqi was eating a croissant. With the child's use of the pronominal form and the pointing gesture, the mother appeared to understand the intended referent in the child's answer.

Example 11: (Jie 2;7)

\*RES: Jie           糖果    比較   重要       還是   玩具   比較  
 Jié           tángguǒ bǐjiào zhòngyào háishì wánjù bǐjiào  
 Jie           candy   more   important or   toy   more  
 重要?  
 zhòngyào?  
 important  
 'Jie, is the candy or the toy more important to you?'

\*RES: 兩    個    只   能   選    -   個   你   要  
 liǎng ge zhǐ néng xuǎn yí ge nǐ yào  
 two CL only can choose one CL you want  
 選    哪    -   個?  
 xuǎn nǎ yí ge?  
 choose which one CL  
 'If you can choose only one of the two, which one do you want to choose?'

\*JIE: 這個    [%眼睛看著糖果].   ←  
 zhè-ge [% looking at the candy].  
 this-CL  
 'This.'

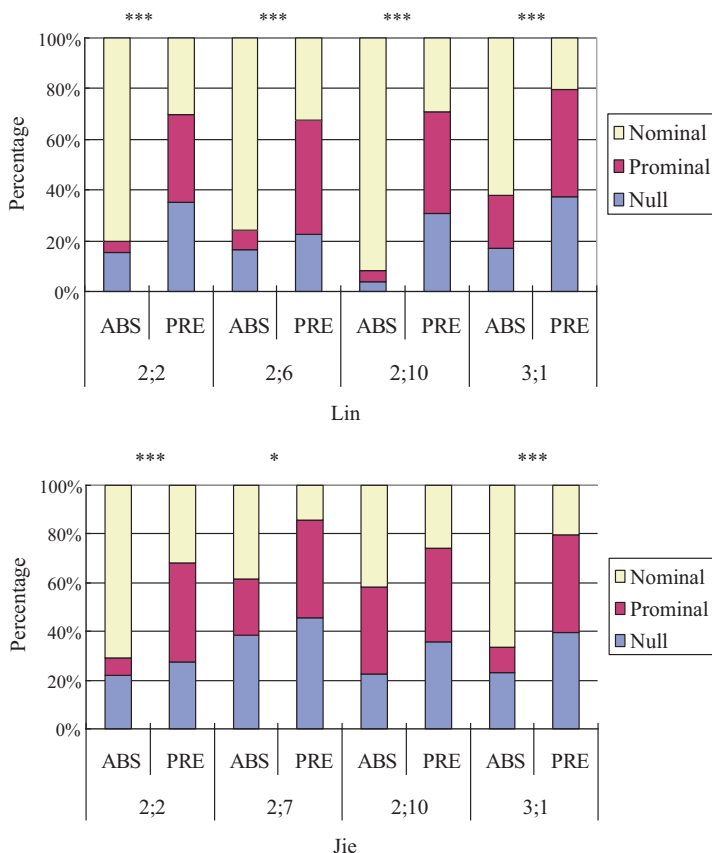


Fig. 1. Distributions of referential forms with respect to Absence in each session. \* $p < .05$ ; \*\*\* $p < .001$ . Note: ABS: absent; PRE: present.

\*RES: 她 要 這個 [%手指糖果].  
 tā yào zhè-ge [% pointing at the candy].  
 she want this-CL  
 'She wants this.'

\*RES: 她 要 糖果 [=! 笑].  
 tā yào tángguǒ [=! laughing].  
 she want candy  
 'She wants the candy.'

In this example, Jie was holding a lollipop in one hand. She was trying in vain to open a toy box with the other hand. The mother suggested that Jie put down the lollipop so that it would be easier for her to open the toy box, but Jie shook her head and kept licking the lollipop. In Line 1, the researcher then asked Jie which one was more important to her, the candy or the toy. As seen in her answer, Jie used a pronominal form *zhè-ge* 'this' with eye gaze to indicate that the candy was the intended referent.

As seen above, both Lin and Jie were sensitive to all of the informativeness features examined, and their sensitivity was reflected in their use of referential forms. They tended to use referential forms with high informativeness to represent referents with informative values and to use referential forms with low informativeness to represent referents with uninformative values. In addition to linguistic strategies, non-linguistic strategies also played an important role in Jie's referential systems. The two children's referential systems can be summarized in Table 14.

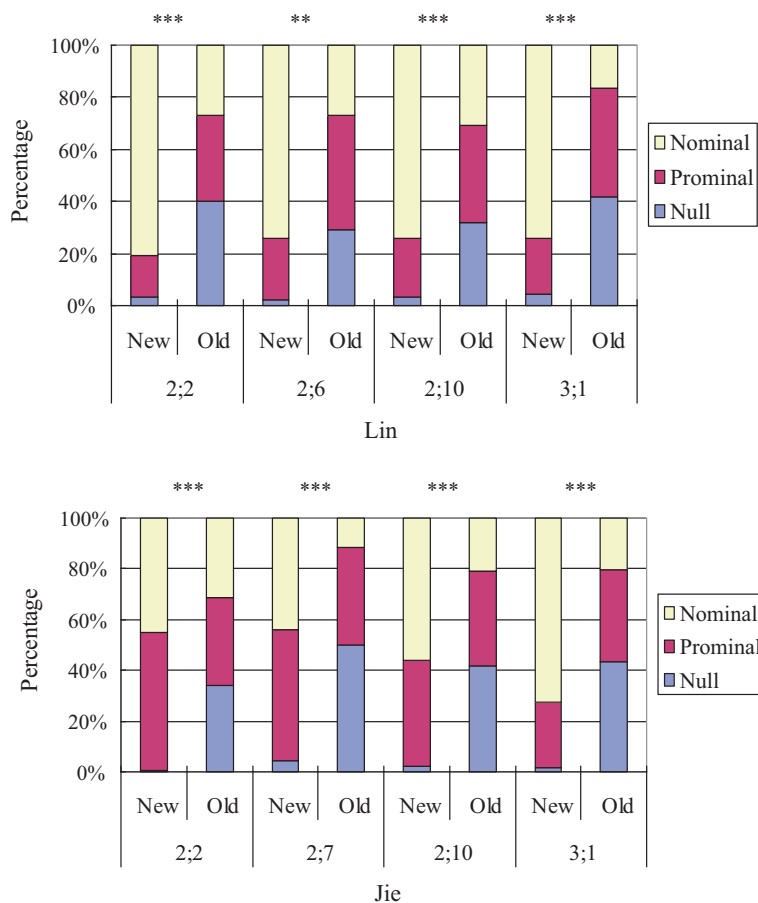


Fig. 2. Distributions of referential forms with respect to Newness in each session. \*\* $p < .01$ ; \*\*\* $p < .001$ .



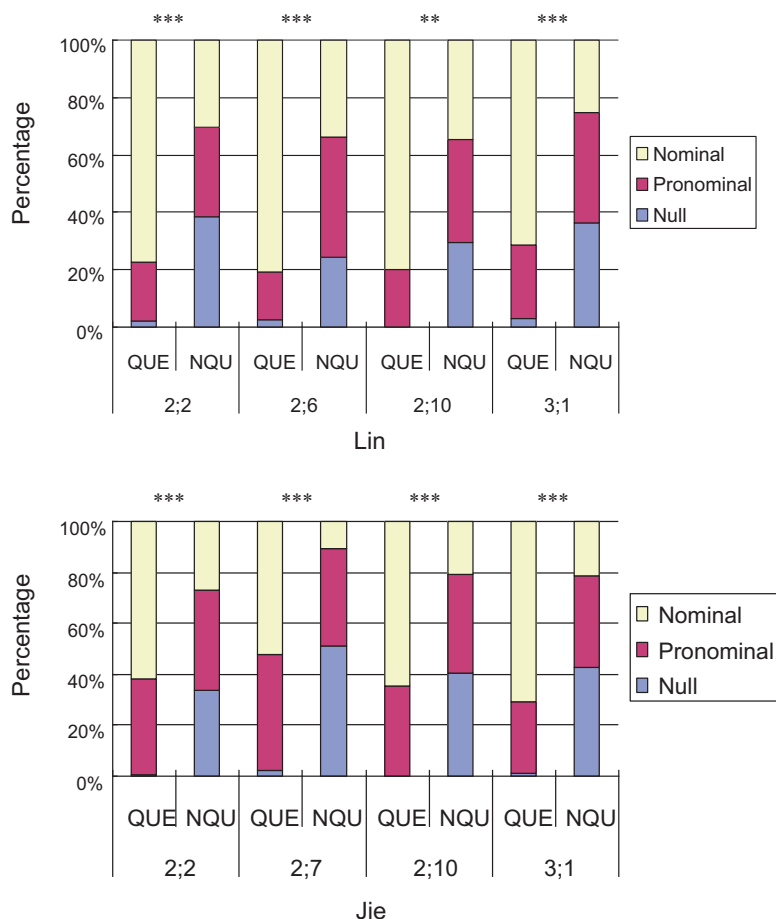
**Table 14**  
The two children's referential systems.

| Informativeness | Informative values                                | Uninformative values |
|-----------------|---|----------------------|
| Lin             | Nominal   | Null<br>Pronominal   |
| Jie             | Nominal<br>Pronominal + non-linguistic strategies | Null<br>Pronominal   |

### 3.2. Referential choice for subjects and objects

Since both subject arguments and object arguments can be omitted in Mandarin Chinese, further analysis was conducted to examine whether the children's subject and object realizations were associated with the same motivation of informativeness. The results are shown in Table 15.

As seen in Table 15, the children tended to use different referential forms for subject arguments and object arguments. Statistical analyses were conducted, and the results showed that the referential choices for subjects and objects differed significantly in both children's data ( $\chi^2(2) = 361.75, p < .001$  in Lin,  $\chi^2(2) = 529.82, p < .001$  in Jie). The results of the post hoc multiple comparison test (Marascuilo and McSweeney, 1977) showed that (1) a significantly higher percentage of null forms were used in the subject position than in the object position in both children's data (36.98% > 11.50% in Lin, and 48.84% > 15.61% in Jie), (2) a significantly higher percentage of pronominal forms were used in the subject position than in the object position in both children's data (46.12% > 20.22% in Lin, and 40.52% > 33.37% in Jie), and (3) a significantly lower percentage of nominal forms were used in the subject position than in the object position in both children's data (16.90% < 68.27% in Lin, and 10.64% < 51.02% in Jie).



**Fig. 3.** Distributions of referential forms with respect to Query in each session. \*\* $p < .01$ ; \*\*\* $p < .001$ . Note: QUE: query; NQU: non-query.

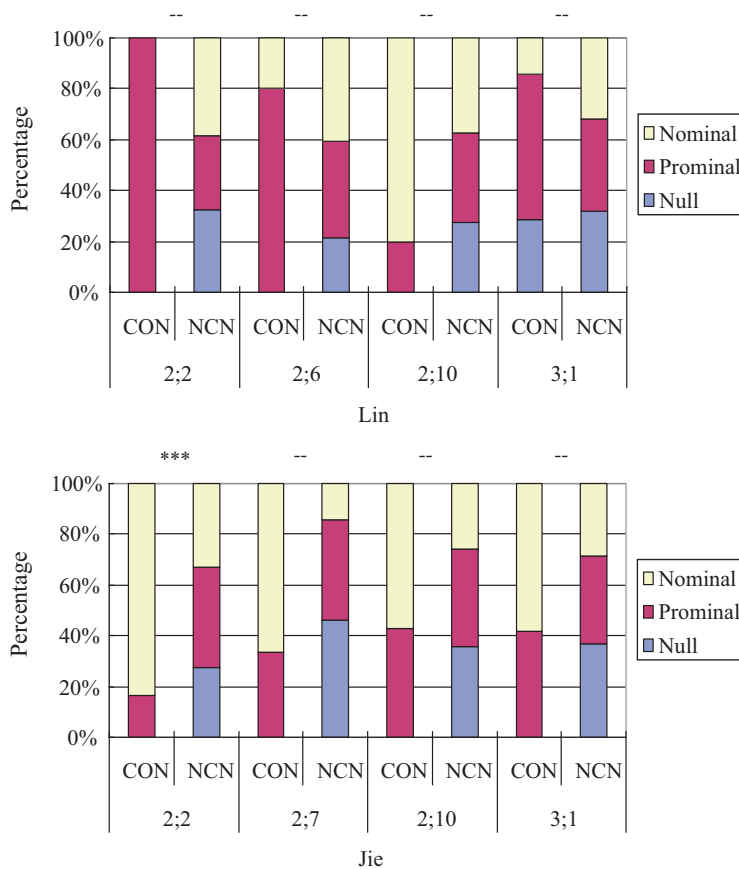
**Table 15**  
Referential forms for subject arguments and object arguments.

|     | Forms      | SUB  |        | OBJ |        | $\chi^2$  | Post hoc |
|-----|------------|------|--------|-----|--------|-----------|----------|
|     |            | N    | %      | N   | %      |           |          |
| Lin | Null       | 291  | 36.98  | 62  | 11.50  | 361.75*** | S > O    |
|     | Pronominal | 363  | 46.12  | 109 | 20.22  |           | S > O    |
|     | Nominal    | 133  | 16.90  | 368 | 68.27  |           | S < O    |
|     | Total      | 787  | 100.00 | 539 | 100.00 |           |          |
| Jie | Null       | 675  | 48.84  | 153 | 15.61  | 529.82*** | S > O    |
|     | Pronominal | 560  | 40.52  | 327 | 33.37  |           | S > O    |
|     | Nominal    | 147  | 10.64  | 500 | 51.02  |           | S < O    |
|     | Total      | 1382 | 100.00 | 980 | 100.00 |           |          |

\*\*\*  $p < .001$ .

From the above, it was found that the referential choice for subject arguments differed significantly from that for object arguments in the children's data. As we have previously observed, referential choice and informativeness were highly related in the children's data, and further analysis was thus conducted to examine whether there was an asymmetry in informativeness between subject arguments and object arguments. The analyses are shown in Table 16.

As seen in Table 16, for almost all of the informativeness features in both children's data, the distribution of informative and uninformative values differed significantly for subject arguments and object arguments. The table reveals that in both children's data, significantly higher percentages of informative values were associated with object arguments, and



**Fig. 4.** Distributions of referential forms with respect to Contrast in each session. \*\*\* $p < .001$ ; '-' not applicable. Note: CON: contrastive; NCN: non-contrastive.

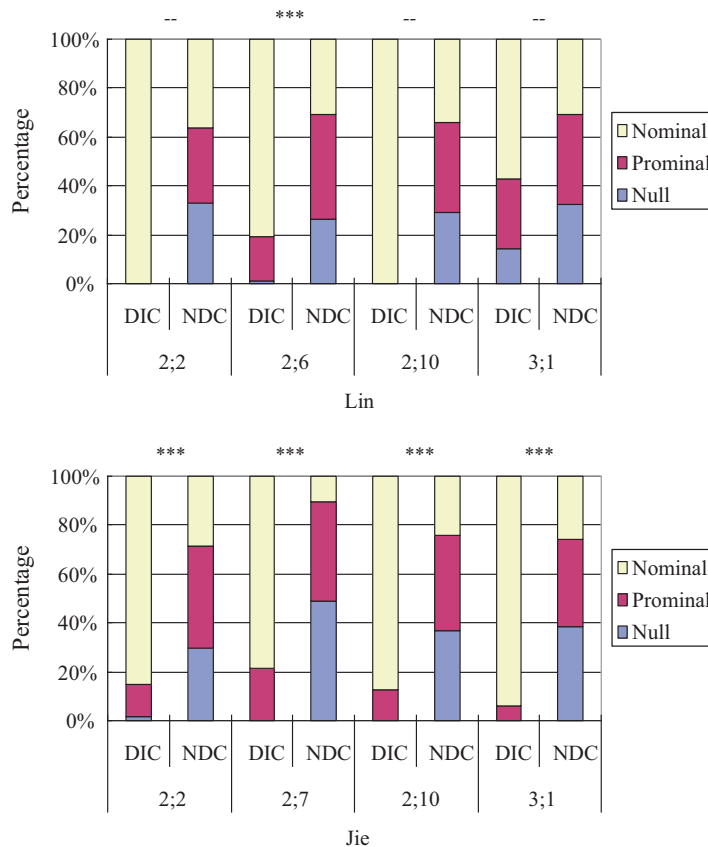
**Table 16**  
Informativeness in subject arguments and object arguments.

|     | Lin       |           |            | Jie        |           |            |
|-----|-----------|-----------|------------|------------|-----------|------------|
|     | SUB (787) | OBJ (539) | $\chi^2$   | SUB (1382) | OBJ (980) | $\chi^2$   |
| ABS | 8.01      | 33.95     | 142.53***  | 5.50       | 14.29     | 53.28***   |
| PRE | 91.99     | 66.05     |            | 94.50      | 85.71     |            |
| NEW | 9.28      | 48.05     | 256.28***  | 5.86       | 32.55     | 290.37***  |
| OLD | 90.72     | 51.95     |            | 94.14      | 67.45     |            |
| QUE | 2.67      | 29.87     | 199.896*** | 1.74       | 34.49     | 474.004*** |
| NQU | 97.33     | 70.13     |            | 98.26      | 65.51     |            |
| CON | 1.40      | 1.30      | n.s.       | 0.65       | 3.47      | 25.48***   |
| NCN | 98.60     | 98.70     |            | 99.35      | 96.53     |            |
| DIC | 6.35      | 16.88     | 37.33***   | 3.69       | 10.31     | 41.68***   |
| NDC | 93.65     | 83.12     |            | 96.31      | 89.69     |            |
| DID | 14.36     | 38.40     | 101.03***  | 4.12       | 15.82     | 95.94***   |
| NDD | 85.64     | 61.60     |            | 95.88      | 84.18     |            |
| INA | 20.20     | 71.24     | 344.64***  | 35.24      | 76.84     | 397.87***  |
| ANI | 79.80     | 28.76     |            | 64.76      | 23.16     |            |
| TRD | 36.85     | 93.32     | 424.40***  | 53.04      | 91.53     | 397.23***  |
| NTD | 63.15     | 6.68      |            | 46.96      | 8.47      |            |

n.s.: not significant.

Note: ABS: absent; PRE: present; QUE: query; NQU: non-query; CON: contrastive; NCN: non-contrastive; DIC: differentiation in context; NDC: non-differentiation in context; DID: differentiation in discourse; NDD: non-differentiation in discourse; INA: inanimate; ANI: animate; TRD: third person; NTD: non-third person.

\*\*\*  $p < .001$ .



**Fig. 5.** Distributions of referential forms with respect to Differentiation in context in each session. \*\*\* $p < .001$ ; '-' not applicable. Note: DIC: differentiation in context; NDC: non-differentiation in context.

significantly higher percentages of uninformative values were associated with subject arguments. It appeared that the asymmetry in informativeness between subjects and objects may contribute to the different referential choices observed for subject arguments and object arguments.

### 3.3. Referential choice at different ages

The children's data in this study were collected for a period of one year (from 2;2 to 3;1). In order to examine whether the same referential strategies were reflected throughout this period, further analysis was conducted to examine the children's referential choices at different ages. The results are presented from Figs. 1–8.

Fig. 1 presents the children's referential choices for absent/present referents in each data session. The results of Chi-square analyses showed that except for Jie's session at 2;10, the children's referential choices for absent referents and for present referents were significantly different in every data session. The results thus revealed the children's early sensitivity to the feature of Absence. That is, since the children were as young as 2;2 and throughout their development, they were able to use different linguistic forms to differentiate between absent referents and present referents.

Similar distribution patterns and statistical results were also observed in the figures concerning the other informativeness features. As seen in Figs. 2 (Newness), 3 (Query), 6 (Differentiation in discourse), 7 (Inanimacy), and 8 (Third-person), the results of the Chi-square analyses reached significant differences in all of the children's data sessions, suggesting that since the children were as young as 2;2 and throughout their development, these informativeness features were powerful variables influencing their referential choices.

As for Figs. 4 (Contrast) and 5 (Differentiation in context), some of the data sessions could not be analyzed statistically because of the limited numbers of CON referents or DIC referents. The data in all of the other sessions revealed significant differences in the children's referential choices. For those sessions which were not analyzed statistically, the distributions still revealed notable differences in referential choices for the different types of referents. Thus, it appeared that the informativeness features of Contrast and Differentiation in context were also important factors affecting the

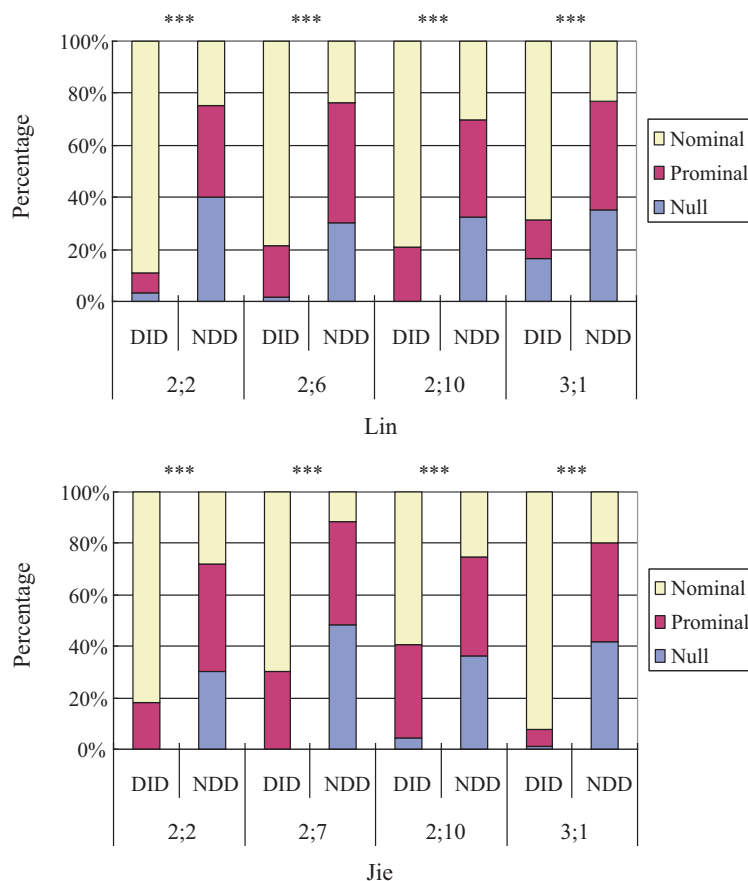


Fig. 6. Distributions of referential forms with respect to Differentiation in discourse in each session. \*\*\* $p < .001$ . Note: DID: differentiation in discourse; NDD: non-differentiation in discourse.

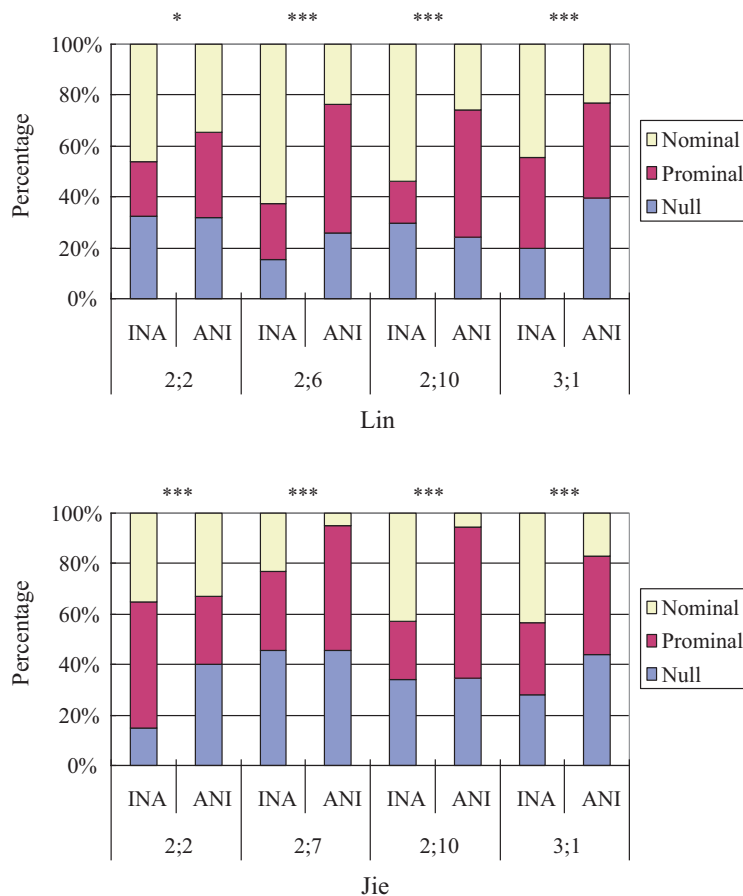


Fig. 7. Distributions of referential forms with respect to Inanimacy in each session. \* $p < .05$ ; \*\*\* $p < .001$ . Note: INA: inanimate; ANI: animate.

children's referential choices, and that the children were sensitive to these factors since the first sessions of the data collection.

#### 4. Discussion and Conclusion

The results indicated that the referential choices of both children were highly influenced by informativeness. The children were sensitive to the eight informativeness features, and their referential choices were made in accordance with discourse-pragmatic principles. The results are thus consistent with those of previous studies which showed that children learning null argument languages were sensitive to the informativeness status of discourse referents (Allen, 2000; Clancy, 1993; Narasimhan et al., 2005; Serratrice, 2005).

While both children were sensitive to informativeness, it was observed that the referential strategies of the two children were slightly different. Lin tended to use nominal forms to refer to referents with informative values, and null forms or pronominal forms for referents with uninformative values. Jie's referential strategies were largely consistent with Lin's; however, Jie also relied on pronominal forms in conjunction with non-linguistic strategies to represent referents with informative values for the features of Newness, Query and Contrast. It has been shown in Guerriero et al. (2006) that children and their mothers often demonstrated close similarity in their linguistic as well as non-linguistic referential patterns. Those children whose mothers were more consistent in using non-linguistic strategies to supplement null or pronominal arguments also used non-linguistic strategies in similar patterns. Thus, we may suspect that the differences observed between Lin's and Jie's non-linguistic strategies may have something to do with maternal input. That is, Jie's mother, in comparison with Lin's mother, may use deictic gestures and eye gaze more consistently to supplement pronominal forms for new, queried or contrastive referents. A preliminary observation of the mothers' data indicated that it seemed to be the case; however, an adequate understanding of the role of maternal input in the children's referential patterns will require systematic empirical investigations of the mothers' referential strategies.

The analyses also revealed the importance of examining the children's referential choices in terms of the three-way classification (null, pronominal and nominal), rather than imposing a binary system (null vs. overt, or non-lexical vs. lexical)

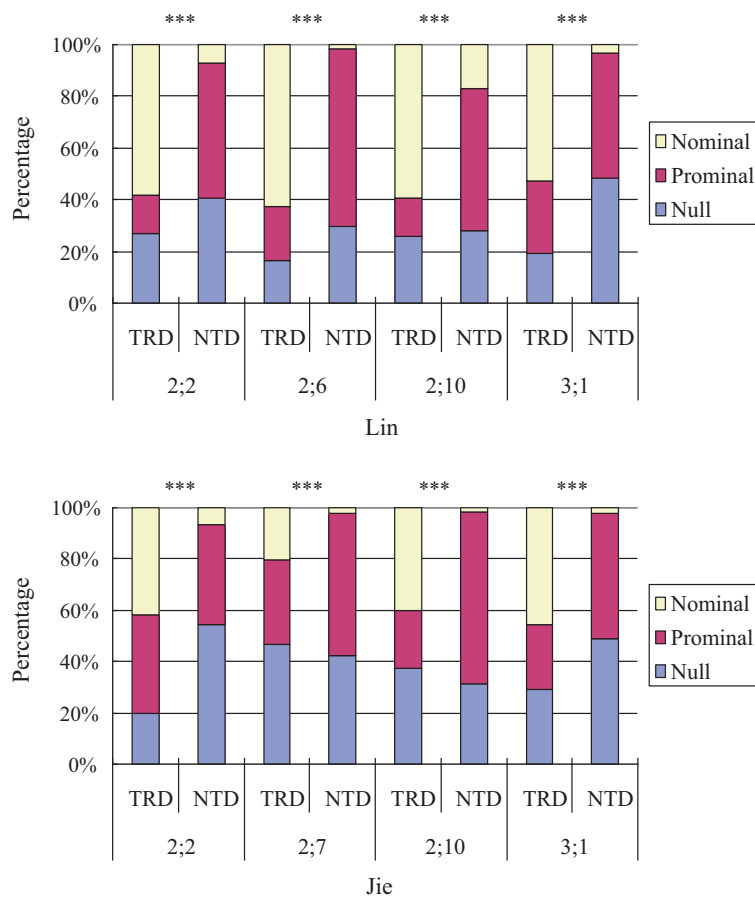


Fig. 8. Distributions of referential forms with respect to Third person in each session. \*\*\* $p < .001$ . Note: TRD: third person; NTD: non-third person.

in the first place. We observed that both children strongly differentiated between null forms and nominal forms with respect to informativeness; however, some individual variations were observed in the children's use of pronominal forms. The results showed that the distribution of pronominal forms was similar to that of null forms in general; however, Jie's use of pronominal forms, with the supplementary use of deictic gestures or eye gaze, also consistently functioned in a way similar to the way in which nominal forms functioned. As suggested by Clancy (1997), it would be an oversimplification to analyze referential choice in terms of a binary opposition because null, pronominal, and nominal forms may have different discourse functions.

Since Mandarin allows the omission of either subject or object arguments, this study also investigated whether the referential strategies for both subject and object arguments were associated with the same discourse-pragmatic factors in child discourse. The results showed that the referential choices for subjects and objects differed significantly in both children's data. Both children used a significantly higher percentage of null forms and pronominal forms in the subject position than in the object position; in contrast, they used a significantly higher percentage of nominal forms in the object position than in the subject position. In addition, the study also examined the relationship between informativeness and argument positions. The results showed that the distribution of informativeness differed significantly between subject arguments and object arguments: Subject arguments tended to be associated with uninformativeness while object arguments tended to be associated with informativeness. Thus, the different referential choices for subject and object arguments appeared to be associated with the asymmetry in informativeness between subject and object arguments. This discourse-pragmatic account provides an alternative explanation to the performance account (Bloom, 1993; Valian, 1991), which suggests that the processing load of a sentence is greater at the beginning of the sentence, thus resulting in more omitted subjects than omitted objects. From another perspective, the discourse-pragmatic account is perhaps complementary to the processing account. As suggested by Allen (2000), it may be that children initially omit subjects because of processing constraints, but that they make an extra effort to produce subjects in just those cases in which the subject is informative.

The analysis also demonstrated that pragmatic sensitivity was reflected in the children's referential choice since the first sessions of the data collection, i.e., since the children were as young as 2;2, and throughout their later development. This may imply that the emergence of such a sensitivity occurred even earlier than the age of 2;2. Some of the previous studies have

revealed that children are sensitive to discourse pragmatics at even earlier ages. Serratrice's (2005) six Italian-speaking children were at the ages of 1;8.5 to 1;11.16 at the onset of the study, and Clancy's (1997) two Korean-speaking children were at the ages of 1;8 and 1;10 at the start of data collection. These children appeared to demonstrate the ability to use different referential forms in a pragmatically sensitive way from the onset of the studies. Further studies of Mandarin-speaking children younger than 2;2 are needed in order to better understand when such a pragmatic sensitivity emerges in the course of Mandarin-speaking children's language acquisition.

This study has provided an important piece of cross-linguistic evidence for the discourse-pragmatic account for children's referential choice. From the perspectives of pragmatic development and cognitive development, children's sensitivity to the dynamics of information flow also reveals an ability to understand other people's perspectives, an ability which is crucial for children to acquire in order to become communicatively competent speakers. In other words, children need to develop the perspective-taking ability in order to be able to assess the informative status of a given referent in the listener's mind in deciding their referential choice in communicative interaction. Thus, the use of referential strategies appears to reflect the links between children's syntactic development, pragmatic development, and cognitive development.

## Acknowledgements

I would like to express my appreciation to the participants of this study for their kindly support and to the two anonymous reviewers for their insightful comments. The research reported in this paper was funded by a grant from the National Science Council of Taiwan (NSC 96-2411-H-004-043).

## Appendix A.

### Transcription conventions.

|           |                         |
|-----------|-------------------------|
| #         | Pause between words     |
| +^        | Quick uptake            |
| +...      | Trailing off            |
| [>]       | Overlap follows         |
| [<]       | Overlap precedes        |
| [≠! text] | Paralinguistic material |
| [% text]  | Comments on main line   |

### Gloss abbreviations.

|     |                 |
|-----|-----------------|
| CL  | Classifier      |
| DUR | Durative aspect |
| NOM | Nominalizer     |
| PRT | Particle        |

## References

- Allen, Shanley, 2000. A discourse-pragmatic explanation for argument representation in child Inuktitut. *Linguistics* 38, 483–521.
- Ariel, Mira, 1990. *Assessing Noun-Phrases Antecedents*. Routledge, London.
- Ariel, Mira, 1996. Referring expressions and the +/- coreference distinction. In: Fretheim, T., Gundel, J. (Eds.), *Reference and Referent Accessibility*. John Benjamins, Amsterdam, pp. 13–35.
- Bloom, Lois, 1970. *Language Development: Form and Function in Emerging Grammars*. MIT Press, Cambridge, MA.
- Bloom, Paul, 1990. Subjectless sentences in child language. *Linguistic Inquiry* 21, 491–504.
- Bloom, Paul, 1993. Grammatical continuity in language development: the case of subjectless sentences. *Linguistic Inquiry* 24, 721–734.
- Chafe, Wallace, 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In: Li, C. (Ed.), *Subject and Topic*. Academic Press, London, pp. 25–55.
- Chafe, Wallace, 1994. *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. University of Chicago Press, Chicago.
- Chafe, Wallace, 1996. Inferring identifiability and accessibility. In: Fretheim, T., Gundel, J. (Eds.), *Reference and Referent Accessibility*. John Benjamins, Amsterdam, pp. 37–46.
- Clancy, Patricia, 1993. Preferred argument structure in Korean acquisition. In: Clark, E. (Ed.), *The Proceedings of the Twenty-Fifth Child Language Research Forum*. CSLI, Stanford, CA, pp. 307–314.
- Clancy, Patricia, 1997. Discourse motivations for referential choice in Korean acquisition. In: Sohn, H., Haig, J. (Eds.), *Japanese/Korean Linguistics*, vol. 6. CSLI, Stanford, CA, pp. 639–657.
- Du Bois, John W., 1985. Competing motivations. In: Haiman, J. (Ed.), *Iconicity in Syntax*. John Benjamins, Amsterdam, pp. 343–365.

- Du Bois, John W., 1987. The discourse basis of ergativity. *Language* 63, 805–855.
- Givón, Talmy, 1983. Topic continuity in discourse: An introduction. In: Greenberg, J. H., Givón, T. (Series Eds.), Givón, T. (Vol. Ed.), *Typological Studies in Language: vol. 3. Topic Continuity in Discourse: A Quantitative Cross-Language Study*, John Benjamins, Amsterdam.
- Givón, Talmy, 1984. *Syntax: A Functional-Typological Introduction*, vol. 1. John Benjamins, Amsterdam.
- Greenfield, Patricia Marks, Smith, Joshua H., 1976. *The Structure of Communication in Early Language Development*. Academic Press, New York.
- Guerriero, Sonia, Oshima-Takane, Yuriko, Kuriyama, Yoko, 2006. The development of referential choice in English and Japanese: a discourse-pragmatic perspective. *Journal of Child Language* 33, 823–857.
- Gundel, Jeanette, Hedberg, Nancy, Zacharski, Ron, 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69, 274–307.
- Halliday, Michael, Hasan, Ruqaiya, 1976. *Cohesion in English*. Longman, London.
- Hamann, Cornelia, Plunkett, Kim, 1998. Subjectless sentences in child Danish. *Cognition* 69 (1), 35–72.
- Hickmann, Maya, Hendriks, Henriëtte, 1999. Cohesion and anaphor in children's narratives: a comparison of English, French, German, and Mandarin Chinese. *Journal of Child Language* 26, 419–452.
- Hirakawa, Makiko, 1993. Null subjects versus null objects in an early grammar of Japanese. *McGill Working Papers in Linguistics* 9, 30–45.
- Huang, Yan, 2000. *Anaphora: A Crosslinguistic Study*. Oxford University Press, Oxford.
- Hyams, Nina, 1986. *Language Acquisition and the Theory of Parameters*. Reidel, Dordrecht.
- Hyams, Nina, Wexler, Ken, 1993. On the grammatical basis of null subjects in child language. *Linguistic Inquiry* 24, 421–460.
- Kumpf, Lorraine E., 1992. Preferred argument structure in second language discourse: a preliminary study. *Studies in Language* 16, 369–403.
- Levinson, Stephen, 1987. Pragmatics and the grammar of anaphor: a partial pragmatic reduction of binding and control phenomena. *Journal of Linguistics* 23, 379–434.
- Levinson, Stephen, 1991. Pragmatic reduction of the binding conditions revisited. *Journal of Linguistics* 27, 107–161.
- Li, Charles N., Thompson, Sandra A., 1981. *Mandarin Chinese: A Functional Reference Grammar*. University of California Press, Berkeley, CA.
- MacWhinney, Brian, 2000. *The CHILDES Project: Tools for Analyzing Talk*, 3rd ed. Lawrence Erlbaum, Mahwah, NJ.
- Marascuilo, Leonard A., McSweeney, Maryellen, 1977. *Nonparametric and Distribution-Free Methods for the Social Science*. Brooks/Cole, Monterey, CA.
- Narasimhan, B., Budwig, N., Murty, L., 2005. Argument realization in Hindi caregiver-child discourse. *Journal of Pragmatics* 37, 461–495.
- Paradis, Johanne, Navarro, Samuel, 2003. Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: what is the role of input? *Journal of Child Language* 30, 371–393.
- Radford, Andrew, 1990. *Syntactic Theory and the Acquisition of English Syntax: The Nature of Early Child Grammars of English*. Blackwell, Oxford, UK.
- Serratrice, Ludovica, 2005. The role of discourse pragmatics in the acquisition of subjects in Italian. *Applied Psycholinguistics* 26, 437–462.
- Valian, Virginia, 1991. Syntactic subjects in the early speech of American and Italian children. *Cognition* 40, 21–81.
- Wang, Qi, Lillo-Martin, Diane, Best, Catherine T., Levitt, Andrea, 1992. Null subject versus null object: some evidence from the acquisition of Chinese and English. *Language Acquisition* 2 (3), 221–254.

**Chiung-chih Huang** is associate professor of Linguistics at National Chengchi University, Taipei, Taiwan. Her research interests include first and second language acquisition, psycholinguistics, discourse analysis, and TESL.