J. Child Lang. **31** (2004), 633–660. © 2004 Cambridge University Press
DOI: 10.1017/S030500090400604X Printed in the United Kingdom

The development of discourse referencing in Cantonese-speaking children*

ANITA M.-Y. WONG
University of Hong Kong

JUDITH R. JOHNSTON
University of British Columbia

(*Received* 8 *April* 2002. *Revised* 22 *July* 2003)

ABSTRACT

The ability to make clear reference in connected discourse was examined in children learning Cantonese, a Chinese language where noun phrase constituents, whatever their grammatical role, are omissible from sentences under discourse conditions that are not well-understood. Forty-three typically developing children aged 3;0, 5;0, 7;0 and 12;0 told 16 stories based on picture sequences. A panel of adult native Cantonese speakers was asked to judge the referential adequacy of each child's stories by identifying the character the child was talking about in 32 targeted referential acts. The targeted acts were of three sorts: MAINTENANCE of a known character, INTRODUCTION of a second new character, and REINTRODUCTION of a known character. Reference was judged to be adequate when 3 out of 4 'listeners' could successfully identify the character. Children's referential expressions were most adequate for Maintenance, less adequate for Introduction, and least adequate for Reintroduction. The twelve- and seven-year-olds approached ceiling on all three functions. The five-year-olds scored poorly on Reintroduction, and the three-year-olds failed both Introduction and Reintroduction, despite knowledge of at least one of the possible linguistic forms required for these acts as evidenced in a sentence imitation task. Viewed within the framework of Levelt's (1989)

[*] This research is based on the first author's doctoral dissertation, completed under the supervision of the second author. D. Geoffrey Hall and Carolyn Johnson offered invaluable and much appreciated suggestions at all stages of the project. The authors would also like to thank Cliff Mok for the story illustrations, Astrid Lau and Isabel Chan for their transcription of the stories, and the twelve Cantonese speakers for scoring the transcripts. Finally, this project would not have been possible without the participation of the children, and the support of their families and schools. We thank them all. Address for correspondence: Anita M.-Y. Wong, Division of Speech and Hearing Sciences, University of Hong Kong, Prince Philip Dental Hospital, 34 Hospital Road, Sai Ying Pun, Hong Kong-SAR-China. e-mail: amywong@hkusua.hku.hk

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

discourse model, the data improve our understanding of the developmental period during which children learn to make appropriate presuppositions about the listener's knowledge and attentional states.

INTRODUCTION

Narrative discourse provides a number of windows through which we can look into the complex interactions of children's language, social and cognitive development (e.g. Berman & Slobin, 1994). One particularly revealing window is provided by the way children refer to different story characters (e.g. Hickmann & Liang, 1990; Hickmann, 1995). Through discourse referencing, we can examine children's repertoire of nominal and pronominal forms, and the degree to which these forms are successfully used to meet the needs of the listener. The patterns of use indicate the nature of children's presuppositions about the listener's knowledge, and the degree to which they are able to track changes in this knowledge over time.

Levelt (1989) provides a scheme for understanding presupposition and how changes in presuppositions are taken into account during discourse referencing. In his proposal adult speakers create and maintain a mental model of a given discourse, and use this model to make decisions about the content and form of utterances. Decisions about referential expressions are based on the speaker's presupposition about the listener's knowledge of the referent as it is represented in this mental model. Frameworks for classifying the speaker's presupposition have changed over time, from the early dichotomy of old vs. new information (e.g. Clark & Haviland, 1974) to more elaborate lists of activation states (Chafe, 1976), accessibility levels (Givón, 1983), and hierarchical cognitive states (Gundel, Hedberg & Zacharski, 1993). These more elaborate characterizations emphasize the cognitive nature of presupposition, which is of two sorts: the representation of the listener's current knowledge of the intended referent, and the listener's relative attention state towards it.

In Levelt's (1989) scheme, the mental model evolves as the discourse develops. Every utterance that the speaker or the conversational partner contributes changes the model. To achieve referential adequacy, the speaker needs to continually update his representation of the listener's knowledge of, and attention state towards the referents. For example, at some initial point, a given entity, X, is not represented in the discourse model at all and thus cannot be presupposed as information that is known to the listener. When X is first introduced into the discourse, it is represented as immediately known to the listener and in the listener's focus of attention because of its novelty. However, if the discourse continues and X is not mentioned further, it remains represented in the discourse model as known information, but no longer as the focus of the listener's attention. This situation will

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

change again if X re-enters the discourse. Thus over time, entity X changes from being potentially unknown, to being in-focus, to being back-grounded, and to once again being in-focus. A speaker who is successful in discourse referencing keeps track of all the changes about all of the referents in his discourse model, each and every time he makes a reference.

Learning to create and maintain the discourse model is a formidable developmental task, one that is accomplished in conjunction with the learning of referential forms. Many factors can influence the nature of children's discourse models and their subsequent choice of expressions. For example, studies by Kail & Hickmann (1992) on referent introduction and Hickmann, Kail & Roland (1995) on reference maintenance suggest that children may proceed quite differently when their listeners have perceptual access to the referents than when they do not. To introduce a referent, Kail & Hickmann (1992) reported that children used definite as well as indefinite forms when the listener shared their access to the referent, but preferred indefinite forms when the listener did not.

In the present study we investigated discourse referencing in situations where the listener had NO perceptual access, and in which the child as speaker needed to both introduce a new referent and maintain reference to already established referents. Prior developmental studies that looked at this same communication task and context are in good agreement as to the developmental trends, though they disagree as to the age at which the child first succeeds. Initially children use definite expressions (e.g. personal pronouns) to maintain known referents more often than they use indefinite expressions (e.g. nouns or noun phrases with an indefinite article) to introduce new referents (Warden, 1976; Emslie & Stevenson, 1981; Hickmann & Liang, 1990; Clancy, 1992; Wong, 1998). New referents are essentially treated as if they were already known to their listeners. Fully contrastive use of language to distinguish between new and known referents appears somewhat later, with children typically using pronouns to refer to entities already known to the listener and nominal forms to refer to new referents. If we set 60% correct contrastive use as the criterion for success, children learning English (Emslie & Stevenson, 1981), Mandarin (Hickmann & Liang, 1990) and Cantonese (Wong, 1998) achieve success between four and a half years and five years of age, suggesting that they have developed some ability to infer and track the listener's knowledge states.

In comparison to the three studies just cited, Clancy's (1992) work on Japanese acquisition suggested that contrastive reference was achieved even prior to age four, while Warden's work on English acquisition indicated that contrastive reference may not be achieved until well after age five. Direct cross-linguistic comparisons will be needed to determine whether structural properties of Japanese can account for the early success. The protracted acquisition reported by Warden (1976) seems to reflect methodological

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

factors. Children in that study did not have an opportunity to examine all of the pictures successively before they were asked to tell a story. It is plausible that they might not fully appreciate the continuity of known characters and the appearance of new characters in the picture sequence.

Recall that the mental model changes as the discourse evolves, and it is in a constantly active state where new information is integrated with known information, and known information is managed with regard to their current relevance to the discourse. From an information processing perspective, all these integration and management activities require cognitive resources, and hence we might expect that the more entities that need to be monitored, and/or the greater the length of the discourse, the more difficult the referential task. The current study was designed to explore these issues as well as to describe the early phases of the development of reference in Cantonese-speaking children. We expected that the high level of NP optionality in Cantonese (Huang, 1984) would make the development of reference more challenging for Cantonese-speaking children than for children learning European languages. On the one hand, not only do they need to discover the particular forms that serve various referential functions, but they must do so with fewer exemplars. On the other hand they must learn the discourse conventions of the language in relation to the ellipsis of NPs. In the next paragraph, features of Cantonese, particularly forms for making reference, are described.

Referring expressions in Cantonese

In Cantonese nouns, there are no inflectional or derivational morphemes for marking number, gender, or case role, and Cantonese nouns can occur without any modifiers at all, i.e. as 'bare' nouns. Nouns can also be modified by a combination of demonstrative determiners, numeral determiners, classifiers, and adjectives, resulting in a number of different noun phrases. In a noun phrase, when a demonstrative or a numeral is used, a classifier obligatorily precedes the noun. A relative clause can also modify a noun, and such clauses, like all modifiers, precede the noun. Here is how the various modifiers are ordered in a Cantonese noun phrase:

Noun phrase: demonstrative–numeral–classifier–adjective–noun.

Cantonese has personal pronouns, demonstrative pronouns, and zero pronouns (ellipsis). Cantonese personal pronouns have no gender marking or overt case marking. The same pronoun is used for both masculine and feminine individuals and for subject and object case roles. The postposed morpheme *deih* is added to a pronoun to indicate plurality. Contexts often determine case in Cantonese, as in Mandarin (Erbaugh, 1992). The personal pronouns are: *ngo5* 'I', *nei5* 'you', *keoi5* 's/he, it' and the demonstrative

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

pronouns are *nei1*, *go2* 'this, that'. A personal pronoun or noun can combine with the head noun to form a possessive noun phrase. The linking particle *ge3* or classifier is optional when the noun phrase refers to personal relationships.

Like Mandarin, the basic unmarked word order in Cantonese is SVO, i.e. subject–verb–object. But Chinese word order is quite variable (Li, Bates & MacWhinney, 1993) and Matthews & Yip (1994) suggest that Cantonese has rather greater freedom in word order than English. Deviations from the basic word order include patterns such as: (a) VS, which involves mostly ergative verbs, (b) dislocation of the subject noun or pronoun to the end of a sentence in colloquial speech (right dislocation), and (c) topicalization of various elements in the sentence by moving them to the beginning of a sentence, resulting in SOV or O(S)V word order.

Like English, Cantonese uses different forms of referring expressions for marking the given and new pragmatic distinction in discourse, including the different noun phrases, lexical and zero pronouns. Lexical and zero pronouns function as definite referring expressions used for marking known referents. Some noun phrases also function as definite expressions, others as indefinite, depending on their internal structure. Unlike English, these noun phrases are not contrasted by determiners such as definite and indefinite articles. Numeral NPs (those which contain a numeral) can mark new referents, whereas Demonstrative NPs and Relative NPs can mark known referents. Classifier NPs and bare nouns can mark either new or known referents, depending on the context in which they are used.

There has been no systematic study of the type(s) of referring expressions used for each of the different information statuses (Brown & Yule, 1983) or cognitive status (Gundel *et al.*, 1993) in Cantonese. Gundel *et al.* (1993) presented evidence that Mandarin shows the same hierarchical relationship between cognitive status and referring expressions, as in French, German and English. For example, zero forms are used to denote entities that the listener knows and is currently attending to and Numeral NPs are used to denote entities that the listener does not know. Numeral NPs refer to noun phrases that contain a numeral, to be followed by a classifier and a noun. Given that the grammatical structure of Mandarin and Cantonese Chinese are similar in most aspects (Matthews & Yip, 1994), it is possible that Cantonese referring expressions correspond to cognitive status, and hence information status, in very similar ways.

In English, the marking of the given and new distinction is primarily by means of NP types. In Mandarin Chinese, in addition to NP types, word order is also involved. It has been observed that the preverbal position is reserved for definite expressions, and hence that indefinite expressions cannot be placed in the preverbal position (Chao, 1968; Li & Thompson, 1981; Zhu, 1982). Hickmann & Liang (1990) confirmed that Mandarin adult speakers marked the given and new distinction maximally, using

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

definite expressions in the preverbal position for marking known referents and indefinite expressions in the postverbal position for marking new referents. Again, given that the grammatical structure of Mandarin and Cantonese Chinese are similar in most aspects (Matthews & Yip, 1994), it is assumed that the given and new distinction is also marked by a combination of word order and NP types in Cantonese (Lee & Szeto, 1996). In Cantonese, a classifier NP or bare noun can function as a definite or an indefinite referring expression, and whether it marks a new or a known referent depends on its position relative to the verb.

In Cantonese Chinese, as in Mandarin Chinese, there is one common construction for introducing new referents. Given that the preverbal position is reserved for definite referring expressions marking known referents, an indefinite numeral NP cannot be placed in the preverbal subject position as in *jat1 go3 naam4zai2 fan3 gan2 gaau3* (one CL boy ASP sleep, 'a boy is sleeping'). With the use of the existential verb *jau5* before the numeral, an existential presentative construction puts the indefinite Numeral NP in the postverbal position, therefore allowing it to be used to signal a new referent (for Mandarin, see Chao, 1968; Li & Thompson, 1981), as in this example, *jau5 jat1 go3 naam4zai2 fan3 gan2 gaau3* (have one CL boy ASP sleep, 'there is a boy sleeping').

In this study, we used a picture-based narrative task to elicit samples of reference from a set of 16 experimentally designed stories, and then asked a panel of native Cantonese speaking adults to make empirical judgements of the adequacy of each referential act. Adequacy here was defined narrowly, referring to whether the listeners can identify the referent of interest. This methodological strategy was necessary because current proposals concerning pro-drop do not yet seem to account for the widespread use of zero forms by young Cantonese speaking children (Lee, 2000). This, as well as the fact that adequacy can only be determined in the linguistic context, made it impossible to determine *a priori* the expressions that would be considered 'adequate' at a particular point in the story. Using the story texts and the panel's judgements of adequacy, we attempted to determine (a) the age at which Cantonese-speaking children succeed in referring clearly to characters in a story, (b) the relative difficulty of various referential functions, e.g. Introduction vs. Maintenance vs. Reintroduction, (c) the specific referential expressions that are used for each function at each age, and (d) whether referential adequacy is dependent upon discourse length.

METHOD

Participants

Participants were forty-seven children, aged 3;0, 5;0, 7;0 and 12;0, who were learning Cantonese as their first language. They were recruited from

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

four kindergartens and two primary schools in different parts of the city of Hong Kong, China. All of the three- and five-year-old children demonstrated age-appropriate language comprehension and production abilities. All but one child received scores no lower than 1·0 standard deviation (*S.D.*) below the mean for their age on both the Receptive and Expressive sections of the Reynell Developmental Language Scale (Reynell & Huntley, 1985; The Committee on Standardization of the Hong Kong version, 1987; hereafter the Reynell). One child in the three-year-old group received an expressive score of −1·5 *S.D.*, but his receptive score was 1·5 *S.D.* above the mean. This latter fact coupled with the teacher's judgment that the child's language skills were comparable to those of age peers suggested that it was reticence rather than language delays that pushed his Expressive score lower. His demeanour during the experimental sessions was compatible with this explanation.

All of the three-, five-, and seven-year-old children received a pass on nonverbal cognitive abilities, with 'pass' defined as no lower than 1·0 *S.D.* below the mean score on the Columbia Mental Maturity Scale (Burgemeister, Blum & Lorge, 1972; hereafter the CMMS) as normed for North American children of the same age. The seven- and twelve-year-old children did not receive a formal language test; nor did the twelve-year-old receive a formal test of their nonverbal cognitive abilities because they were older than the normative samples for the Reynell and the CMMS. However, they, as well as the younger children, showed average abilities in class, and had no documented history of speech, language, or learning problems, or hearing or psychosocial difficulties, as reported by their classroom teachers. Of the 47 children, four were eventually excluded from the analysis. Three of the three-year-olds refused to participate in the experimental task, and data was lost from one five-year-old due to technical difficulties. This left 43 children in four developmental groups: three-year-olds ($n=10$), five-year-olds ($n=13$), seven-year-olds ($n=12$) and twelve-year-olds ($n=8$). Each age group had approximately equal number of boys and girls.

Task

Each child was seen individually in a quiet room in their schools. He/she was asked to tell 16 short stories to a blindfolded puppet, each story based on a different wordless picture book with 4 to 6 pages of coloured pictures. The experimenter first read out the title of the book, e.g. 'This story is called Washing the Car,' and leafed through the pictures once with the child. Then she prompted the child to make up a story for the puppet with support of the pictures. Each of the expected stories consisted of one simple episode with the basic elements of setting, event, and resolution. Story titles were always given as verb forms without any noun phrases referring to

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

the story characters. All story characters were humans. The stories were designed to give a range of complexity. Two of the stories featured only a single character. The other 14 stories featured both a major and a minor character who, although they were of the same gender and age, could be distinguished by differences in appearance, e.g. hair colour and outfit. The two characters did not appear together in the same picture except when a second character entered, or re-entered, the story. At these points, the new character would be portrayed in the foreground of the picture appearing larger and closer to the child, while the continuing character moved to the background away from the child.

Each child told three practice stories to become familiar with the task. The experimental story books were randomized, and presented in the same order for all. Possible order effects were mitigated by the relatively large number of stories used. Throughout the session, the blind-folded puppet made no comments and asked no questions while a story was being told, but would periodically compliment the child at the end of a story.

All stories were tape recorded and transcribed for analysis by an adult native Cantonese speaker. Ten percent of them were independently transcribed by a second adult native speaker of Cantonese. Word-for-word agreement in transcription was 97·5%.

Analysis: discourse reference and referential functions

The analysis focused on a sample of all possible referential function and character combinations, referential functions for the main character in particular. The main character is the only character appearing in the first picture of all the stories. Only referential acts serving one of the functions as specified below were included in this analysis:

Maintenance: the referential function in which the speaker continues to refer to the MAIN character in the second picture. The main character continues to be the only character in the story, and hence remains in the foreground.

Introduction: the referential function in which the speaker refers to the MINOR character as he/she appears for the first time after the main character has been introduced in the first picture. The minor character appears in the foreground while the main character moves to the background of the picture.

Reintroduction: the referential function in which the speaker once again refers to the major character after referring to the MINOR character in the previous picture. The major character appears in the foreground while the minor character moves to the background.

For the introduction function, we chose to focus on the minor character in order to avoid possible stereotyped introductions such as the English

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

'Once upon a time there was a …'. The stories were designed such that opportunities for Introduction and Reintroduction occurred either early or late in a given story relative to the first picture when the major character first appears. Early introduction of the minor character occurred immediately in the second picture, and late introduction in the fourth picture. Early re-introduction of the main character occurred in the third picture while late re-introduction in the fourth picture. Across the set of stories, the number of expected referential acts for each function varied: 6 for Maintenance, 14 for Introduction (10 early and 4 late) and 12 for Reintroduction (4 early and 8 late) in the discourse.

Analysis: ratings of referential adequacy

As noted earlier, our analysis of the development of reference made use of empirical ratings of communicative adequacy. This methodological strategy was necessary due to the high degree of NP optionality in spoken Cantonese and the key role of context in determining adequacy. In this study, adequacy was determined by having twelve adult native speakers of Cantonese read transcripts of the children's stories and attempt to identify the characters to which the children were referring. These expert 'listeners' were undergraduate students and recent immigrants to Canada from Hong Kong, spoke Cantonese at home, and passed a screening assessment of spoken Cantonese and written Chinese. As most immigrants from Hong Kong, they learned and speak English as a second language.

Each story was rated independently by four expert 'listeners'. To reduce the workload for any given judge, the rating task was spread across 12 judges, as follows. Each judge was asked to rate 224 (or 240) stories, i.e. all the stories produced by 14 (or 15) children. The stories rated by any one judge spanned the entire age range, but age and any other identifying information was removed from the transcript prior to rating. The stories from a given child were considered in the order they had been told, but across children the sets of stories were randomized. The 'listeners' finished their rating task within a week in two to three sessions of about two hours each.

Prior to being rated, the story transcripts were marked to indicate the points at which one of the targeted referential functions actually occurred. The 'listeners' were asked to stop at these points in the story text and note in the margin their answer to two questions about the adequacy of reference:

1. Compare the action underlined here to the action in the previous utterance. Is the child talking now about the same character, or a different character?

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

2. If your answer to question 1 is 'different,' has the child talked about this character in some earlier sentence? If 'yes,' please identify the person who is performing the underlined action.

Since the blindfolded-puppet listener had not been able to see the pictures, the expert human listeners were likewise given no access to the story pictures. They were encouraged to give an immediate and intuitive answer to each of these questions without extensive analysis. Before proceeding with their own set of transcripts, judges were trained with transcripts of stories from the training items. As part of this instruction, they were assured that they could base their answer on any information in the text, including but not limited to, the linguistic form chosen for that referential act.

Scoring for referential adequacy and dominant forms

Each targeted referential act attempted by a given child was rated by four 'listeners'. Across all children, the level of agreement among judges was high. At least three of them made the same adequacy decision for an average of 30·3 out of the 32 referential acts per child. All age groups combined, only 5% of the targeted acts had disputed status, where the judges split their adequacy decisions 2 vs. 2. The number of disputed acts was higher for the younger children, but even for the three-year-olds no more than 9% of the ratings were split in this way. A referential act was scored as being 'adequate' if at least three of the four judges answered the question(s) in a way that was appropriate for that type of referential function. For example, in the case of an Introduction act, the appropriate answers would be 'different' and 'no'. Referential acts which did not meet this adequacy judgement criteria were treated as not adequate. A child received zero point for adequacy when she/he failed to produce an expected referential act, and the total number of targeted acts thus be held constant for all children. The child's adequacy score for each of the three referential functions consisted of the percentage of targeted acts that was judged to be referentially adequate. All age groups combined, 80·5% of the referentially adequate acts were based on a 4–0 vote, and 19·5% based on a 3–1 vote.

This approach to determining adequacy is frankly operational, and does not fully reveal the bases of each judge's decision. It remains possible, for example, that even when the four judges agreed, they did so for different reasons. As Orsolini, Rossi & Pontecorvo (1996) pointed out, listeners can use information other than the referring expressions, such as the semantic content of the verbs, to identify the referent unambiguously. However, this scoring strategy had the advantage of removing researchers from interpretations of adequacy, and in the final analysis does not seem so different from everyday language use.

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

For each type of referential act, the referring expressions used by each child were identified and tallied. The forms that were used fell into one of the following eight categories:

1) Existential presentative expression (the numeral *jat1* is optional)
   jau5(jat)1 go3 naam4zai2
   have (one) Cl boy
   'there is a boy'
2) Proper noun and common noun
   mei5 jin1
   'Mei Yin'
   ze4ze1
   'sister' (a term for an older acquaintance as well as an older sibling)
   neoi5zai2
   'girl'
3) Classifier NP
   go3 naam4zai2
   CL boy
   'boy'
4) Demonstrative NP
   go2 go3 naam4zai2
   that CL boy
   'that boy'
5) Determiner *ling6ngoi6* 'another' NP
   ling6ngoi6 jat1 go3 naam4zai2
   another one CL boy
   'a different boy'
6) Relative NP
   sik6 jyun4 je5 go2 go3 neoizai2
   eat VPrt thing that CL girl
   'the girl who has finished eating'
7) Pronoun
   keoi5
   'he/she'
8) Possessive NP
   naam4zai2 go3 (ge3) pang4jau5
   boy CL (POSS) friend
   'the boy's friend'

Classifier NPs, Demonstrative NPs, and Pronouns, lexical or zero, are definite, while existential expressions are generally indefinite. Proper nouns, Determiner NPs, Relative NPs, and Possessive NPs can be definite or indefinite, depending on the information they convey, or referents

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

mentioned earlier. Hence the classification of this latter set of forms were considered in their given discourse context.

Once the responses were classified, the forms that were successfully used most often for a given referential function were identified for each age group. These 'dominant forms' were initially defined as any form used by the group as a whole in at least 20% of their referentially adequate responses, and used by at least half of the children in the group. Actual levels of use for the 'dominant' forms varied by function and age group as will be discussed in the Results section.

Sentence imitation task

In addition to the story task, each child completed a sentence imitation probe task designed to ascertain his/her knowledge of the referring expressions necessary for the targeted referential functions. Children were asked to repeat a set of nine sentences of equal length, all but one consisting of eight monosyllabic words and one bisyllabic word (see Appendix A). Each sentence began with one of four common types of referring expression: Existential presentative expressions (2 items), Demonstrative NP (1), Relative NP (4), Determiner *ling6ngoi6* NP (2). The remainder of each sentence consisted of a simple predicate adjective, or verb with or without a noun. There were no sentence-final particles. All children were able to repeat at least two out of the three shorter phrases that were used for training. For purposes of later analysis, a child received 1 point when the referring expression portion of the item was repeated verbatim.

RESULTS

Referential adequacy : function

The first analysis looked at developmental changes in referential adequacy. Table 1 indicates for each age group the mean percentage of targeted referential acts that were judged to be referentially adequate. The age-related progress is clear: the three-year-olds seldom made clear reference to the story characters, while the twelve-year-olds virtually always did so. ANOVA procedures are robust to unequal sample sizes as long as the unequal cell frequencies which result are random (Winer, Brown & Michels, 1991) and therefore are used for this and subsequent analyses. A two-way mixed-model ANOVA, '(4) group × (3) function', confirmed a significant main effect of Group, $F(3, 39) = 81{\cdot}58$, $p < 0{\cdot}05$. *Post hoc* analysis indicated that there were reliable performance differences between all but the oldest two groups, Scheffe, $p < 0{\cdot}05$.

Children's ability to make clear reference changed markedly with the type of referential function. All age groups combined, children provided the clearest reference when they were Maintaining reference to a character, were

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

TABLE 1. *Mean* (S.D.) *percent of targeted referential acts judged as referentially adequate, by function and age group*

		Referential function			
Age	*n*	Maintenance	Introduction	Reintroduction	Combined
3	10	70 (25)	13 (20)	2 (4)	28
5	13	97 (6)	65 (23)	38 (20)	67
7	12	94 (13)	92 (7)	66 (17)	84
12	8	100 (0)	92 (7)	84 (12)	92
All		90	65	46	

less effective when they were Introducing a second character, and were least clear when they were Reintroducing a character. The two-way mixed-model ANOVA, (4) group × (3) function, confirmed this main effect for Function, $F(2, 78) = 91{\cdot}78$, $p < 0{\cdot}05$. *Post-hoc* analysis indicated that each of the differences between referential functions was statistically reliable, Scheffe, $p < 0{\cdot}05$. There was also, however, a statistically reliable interaction effect between group and function, $F(6, 78) = 10{\cdot}72$, $p < 0{\cdot}05$. Inspection of the means indicates a developmental ceiling effect: differential success on the three referential functions was seen primarily in the younger children.

In order to further explore the group findings, an analysis of the pattern of referential adequacy scores was conducted on a child-by-child basis. A cut-off point was set at 50%, and each child received a pass (P) when s/he scored at or above the cut-off on a given referential function, i.e. Maintenance, Introduction, Reintroduction. If the group means reflect a true order of acquisition, we would expect to find each individual child showing one of the following performance patterns: PPP, PPF, PFF, FFF. All of the observed 'pass/fail' patterns across the three functions did in fact correspond to one of these four developmental patterns. There was no child, for example, who Reintroduced a character clearly but failed in Introduction or Maintenance. This scalar analysis thus confirmed the developmental order of success on the three referential functions that had been indicated by the group findings: Maintenance, then Introduction, then Reintroduction. Table 2 indicates the age distribution of children achieving each of the developmental patterns. As would be predicted, the proportion of children passing all three functions increases with age, with virtually all of the seven- and twelve-year-olds reaching this level.

Referential adequacy : discourse position

The next set of analyses investigated the possibility that it would be more difficult for children to provide clear reference as the length of the discourse

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

TABLE 2. *The age distribution of children showing various patterns of success on the three referential functions*

Performance/Age group pattern[a]	3	5	7	12
FFF	2	0	0	0
PFF	7	2	0	0
PPF	1	6	1	0
PPP	0	5	11	8

[a] P = Pass, F = Fail; scores for the three functions listed in the order Maintenance, Introduction, Reintroduction.

TABLE 3. *Mean* (S.D.) *percentage of targeted referential acts judged to be adequate, by function, discourse position, and age*

		Introduction		Reintroduction	
Age	*n*	Early	Late	Early	Late
3	10	14 (20)	10 (24)	0	3 (7)
5	13	61 (28)	7 (22)	37 (28)	44 (23)
7	12	89 (10)	98 (7)	75 (15)	67 (28)
12	8	90 (9)	97 (9)	81 (17)	92 (13)

increased and there was more information to hold and consider in the discourse model. For the Introduction and Reintroduction functions, targeted contexts were designed to occur either early or late in the discourse. Table 3 gives the mean (*S.D.*) referential adequacy scores for the Introduction and Reintroduction functions in each of these positions. Contrary to our prediction, for both Introduction and Reintroduction, children generally made more adequate reference when the referential act occurred later in the discourse. This effect of discourse position was found to be statistically significant in a three-way mixed-model ANOVA, '(4) group × (2) function × (2) position', $F(1, 39) = 5{\cdot}96$, $p < 0{\cdot}05$. Although the seven-year-olds showed a reverse pattern in the Reintroduction function, there were no significant interaction effects for position.

Referential adequacy : linguistic forms

We next looked at the specific linguistic forms that children used to refer to new and known characters in different age groups. Sample transcripts for the story DANCING are given in Appendix B. In this story, the child was required to introduce the minor character who is playing the piano, and to reintroduce the main character who is putting on her ballet shoes. As listed earlier, Cantonese provides eight major types of expressions for referential

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

TABLE 4. *Percentage of referentially adequate acts using a given language form, by age and function*

Form	Age group in years			
	3	5	7	12
I. Maintenance of a known character				
Zero pronoun	81	30	32	33
Pronoun		68	63	63
Existential expression				
Determiner NP				
Possessive NP				
Proper noun				
Demonstrative NP				
Relative NP				
Classifier NP				
II. Introduction of a second new character early (and late) in the discourse				
Zero pronoun				
Pronoun				
Existential expression		48 (70)	47 (60)	47 (42)
Determiner NP		24		
Possessive NP			21	33 (32)
Proper noun				
Demonstrative NP				
Relative NP				
Classifier NP				
III. Re-introduction of a known character early (and late) in the discourse				
Zero pronoun				
Pronoun		44 (29)	22	38
Existential expression				
Determiner NP				
Possessive NP				
Proper noun			(25)	(30)
Demonstrative NP		(38)	25 (20)	23 (20)
Relative NP			25	(20)
Classifier NP				

acts, along with the 'zero' option. Table 4 indicates the dominant forms used to achieve referential adequacy for each type of function, in the early and late discourse positions. The values in the table are for the percentage of adequate responses making use of a given form. Recall that 'dominant forms' are any expressions used by the age group as a whole in at least 20% of their referentially adequate responses, and by at least half of the children in the group.

Two dominant forms were used to Maintain reference to a given character, and these forms accounted for virtually all adequate responses. The three-year-olds primarily used the zero pronoun, i.e. their utterance for the second picture typically included just a predicate. The older groups of

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

children primarily used definite pronouns to maintain reference to the protagonist, with zero pronoun as the alternative.

For Introduction to the second character, no dominant forms could be identified for the three-year-olds since their referential adequacy scores for this function were low. For the five-, seven- and twelve-year-olds, the indefinite Existential presentative expression, e.g. *jau5 jat1 go3 neoi5zai2* ('there is a girl') as illustrated in the transcript from child VII3, was the preferred dominant form and was used in at least 40% of the referentially adequate acts in both discourse positions. However, two other expressions were also used at levels above 20%, i.e. Determiner *ling6ngoi6* ('another') NP, e.g. *ling6ngoi6 jat1 go3 neoi5zai2* ('*another girl*') as illustrated in the transcript from child V1 or Possessive NP, e.g. *keoi5 ge3 ga1ze1* ('*her sister*') as illustrated in the transcript from child XII2. These forms can be either definite or indefinite, but were interpreted by the judges as indefinite in these story contexts for the introduction of the minor character. Reference made by the use of the pronoun *keoi5* was often considered inadequate by the judges.

For Reintroduction of a referent, again no dominant forms could be identified for the three-year-olds due to their minimal success with this function. The five-, seven-, and twelve-year-olds tended to use one or more of three definite forms at levels of 20% or more, i.e. Pronoun, e.g. *keoi5* ('*she*') as illustrated in the transcript from child V1, Demonstrative NP, e.g. *go2 go3 neoi5zai2* ('that girl') as illustrated in the transcript from child XII2, or Relative NP, e.g. *tai2 syu1 go2 go3 neoi5zai2* ('the girl who was reading') as illustrated in the transcript from child VII3, with no clear trend related to age or position. The seven-and twelve-year-olds also used proper nouns in the late discourse position, e.g. 'Helen', that could be considered definite in the narrative context. Reference made by the pronoun *keoi5* in most cases, or by existential presentative expressions, e.g. *jau5 go3 neoi5zai2* ('*there is a girl*'), for reintroduction of the main character was considered inappropriate. Here is an example of an inadequate reintroduction from a five-year-old (child V4).

Jin4hau6 ling6ngoi6 jat1 go3 neoi5zai2 le1 keoi5 zau6 taan4 kam4
And then another one CL girl Prt s/he then play piano
(Prt: topic marking particle)
Jin4hau6 jau5 go3 neoi5zai2 jau6 zoek3 haai4
And then have CL girl also wear shoes

In sum, the dominant forms used by children in their narratives varied with age and function. Over time, as children grew adept at different referential functions, they made use of specific sets of expressions for each. The only exception to this pattern of specificity was the use of pronouns for both maintenance and reintroduction of a known character, which will be addressed in the Discussion section.

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

TABLE 5. *Number of children evidencing knowledge of specific referring expressions, by age*

Age	*n*	Existential presentatives	Determiner *ling6ngoi6* NP	Relative NP	Demonstrative NP
3	10	8	6	4	6
5	13	13	12	13	13
7	12	12	11	12	12
12	8	8	8	8	8

Sentence repetition probe

Recall that in addition to the story telling task, children were asked to repeat nine sentences containing one of four common types of referring expressions (see Appendix A), and were awarded one point for each exact repetition of the referential phrases. Mean scores increased with age from 4.2 at 3;0 to 8.9 at 12;0. The three-year-olds were able to imitate fewer than 50% of the subject noun phrases accurately, the seven- and twelve-year-olds were almost always accurate.

The primary purpose of this task was to ascertain children's knowledge of specific referring expressions, i.e. Existential presentatives, Relative NPs, Determiner *ling6ngoi6* expressions, and Demonstrative NPs. Elicited imitation has been used widely by researchers who wish a controlled, efficient way to assess grammatical knowledge (e.g. Brown, 1973; Gerken, 2000). Although there is some uncertainty about the level of processing that underlies success, e.g. syntactic vs. phonological, this uncertainty can be minimized by presenting sentences that are relatively long. Table 5 lists the number of children at each age who demonstrated knowledge of a given form, as evidenced by the accurate repetition of at least one item. With the exception of two three-year-olds, all of the children showed knowledge of at least one of the two forms that could be used for Introduction of a new character, i.e. Existential presentatives, or Determiner *ling6ngoi6* NPs. Most of the children successfully imitated both. Likewise, with the exception of four three-year-olds, all of the children evidenced knowledge of both of the expressions that could be used to Reintroduce a character, i.e. Relative NPs and Demonstrative NPs.

Referring expressions used for initial Introduction of the main character

Although the study was not designed to investigate referential adequacy in the initial Introduction of the protagonist, we looked briefly at the referring expressions used for this function. The findings were in harmony with the patterns reported earlier. The two three-year-olds who were not able to imitate Existential presentative expressions in the probe and did not

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

adequately Introduce the second character, also did not use these forms for the initial Introduction of the protagonist.

DISCUSSION

The narratives collected for this study show clear age-related increases in referential adequacy between 3;0 and 12;0. The three-year-olds were communicatively adequate in very few of their referential acts, while the twelve-year-olds achieved near-perfect reference. Three aspects of this development invite our attention. First, children in this study used systematically different referring expressions for the three referential functions. Second, some referential functions proved to be more difficult than others, with Maintenance of reference to a character being the least difficult and first acquired, followed by Introduction of a second character, then Reintroduction of a known character. Finally, across all age groups, reference was generally more adequate when the Introduction or the Reintroduction function came later rather than earlier in the discourse. In the discussion to follow, we will consider the ways in which each of these findings point to the competencies that underlie the development of discourse referencing.

As summarized earlier, Levelt (1989) argues that these speaker competencies include (1) the ability to construct a mental model of the discourse which incorporates the listener's knowledge of, and attention state towards, the referents, and (2) the ability to modify this model throughout the discourse, and keep track of the changes. These competencies are realized in the choice of an appropriate linguistic expression and are implicit in adequate reference to the different characters at different points in the story.

Linguistic patterns of reference

The older children in this study tended to use different expressions for the three different referential functions, as evidenced in the data on dominant forms. This selectivity can be seen to reflect two different aspects of communicative competence, (1) consideration of the listener's knowledge and attentional state, that is, presuppositional ability, and (2) narrative organization, e.g. attention to coherence.

Presuppositional ability is implicit in the emerging contrast between indefinite and definite forms; referents that were familiar to the listener were introduced with greater specificity than those that were new. Although contrastive use of definite and indefinite expressions was primarily seen in the older two groups of children, it was first manifest in the stories of a subgroup of the five-year-olds, a result that is consistent with prior studies by Clancy (1992) with Japanese speakers, Wong (1998) with Cantonese Chinese speakers, and Hickmann & Liang (1990) with Mandarin Chinese speakers. Children in the present study used indefinite referring expressions,

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

i.e. Existential expressions and Determiner *ling6ngoi6* NPs, to Introduce a new (minor) character, and definite referring expressions to Maintain and Reintroduce a character. The definite expressions were further differentiated, with pronouns and zero forms used for Maintenance, and pronouns, Demonstrative NPs, proper nouns and Relative NPs used for Reintroduction. This consistent mapping between form and referential function indicates that as children enter the school years, their choice of a particular referential expression is in part guided by presuppositions that derive from their discourse models.

Other aspects of form variation seemed to reflect narrative organization rather than presupposition. Children in this study evidenced contrastive use of nominal and pronominal forms, with pronominal forms preferred for Maintaining reference and nominal forms preferred for Introductions. This pattern would be expected from speakers who were appropriately judging the listener's knowledge states. However, data for the Reintroduction function include BOTH pronominal and nominal forms, a pattern which is not as easily explained by presupposition. The source of this pattern is suggested by several earlier studies of reference switching. Children between the 3;0 and 9;0 have been reported to use pronouns when switching reference back to the main protagonist, and nominals when switching reference back to minor characters (Karmiloff-Smith, 1981; Bamberg, 1987; Gomme & Johnson, 1997). This 'thematic subject' strategy, first introduced by Karmiloff-Smith (1981), reveals children's attempt to impose a coherent structure onto their narratives through reference to characters of different thematic importance in different ways. Pronoun use for Reintroduction of the main character by some of the older children in our study may indicate that they too were using such a 'thematic subject' strategy as a narrative organizing device. If so, children's choice of referring expressions could be seen as guided not only by their presuppositions about the listener's knowledge, but also by their own narrative schemes. On the one hand, they decide whether the listener has knowledge of, and or currently attending to the character to be referred to. On the other hand, they decide whether the character of concern is the central protagonist, or a minor character. Since our stories are much shorter than would be adequate to examine the use of the thematic subject strategy, and also they did not require Reintroduction of a minor character, this possibility of the concurrent use of the and presuppositions remains to be confirmed.

Finally, again returning to form variations, there was a tendency for the older children to use more different forms. While this in itself is not remarkable, the specific forms that appeared only in the stories of the seven- and twelve-year-olds are intriguing: Possessive NPs for Introduction of the second character, and proper names for Reintroduction of the main character. These are not forms that present linguistic challenges. Proper

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

nouns are among the earliest words learned, and Possessive NPs have been reported in Cantonese-speaking children at as early as 5;0 (Fletcher, Leung, Stokes & Weizman, 2000). However, to use a proper noun for a character, the child has to first name the character and then to refer to him/her with this name, and to use a possessive NP, the child has to make up a relationship between the characters. We speculate that, like the thematic subject strategy, these acts too imply narrative skills deriving from exposure to literacy or practice with storytelling.

Systematic variation in the use of referential expressions was primarily seen in those children who succeeded in all three referential functions, i.e. in the seven- and twelve-year-olds, and a subset of five-year-olds. These children demonstrated linguistic knowledge of the necessary nominal and pronominal forms, a certain degree of narrative sophistication, and the cognitive ability to make presuppositions about, and keep track of changes in, their listener's knowledge and attentional states. We turn now to the performance of the younger children to gain some insight into the emergence of these discourse competencies.

The emergence of presuppositional abilities

Our analyses indicated scalar performance on the three referential functions, i.e. Maintenance, Introduction, and Reintroduction. Across groups, Maintenance of reference to a character was the least difficult and first acquired, followed by Introduction of a second character, then Reintroduction of a known character. The first step in this developmental sequence reflected the performance of the three-year-olds, who were referentially adequate only on the Maintenance function. They were not referentially adequate on the Introduction and Reintroduction functions, even though 80 and 60% of the children, respectively, demonstrated knowledge of the linguistic forms that are typically used for these functions. Regardless of referential context, these children consistently referred to a character as if it were the same as had been talked about in the prior utterance, using the zero pronoun form. Peterson (1990) reports similar findings in a study in which children ages 2;0 to 3;6 explicitly Introduced people that were unknown to their listener only 13% of the time in personal narratives. Instead they tended to use pronouns or common names, which erroneously presumed prior knowledge of the individuals being referred to.

Some three-year-olds' undoubtedly had difficulty interpreting the pictures and constructing a coherent story line, as has been reported by other investigators (Berman & Slobin, 1994). For example, one three-year-old child remarked 'Why is (he) making a call' when shown the last picture in the 'accident' story in which a passer-by is calling for help. The child was drawn to the new character but failed to understand how the actions he was

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

describing could be related to prior events. The absence of a coherent story line would, of course, call into question any analysis of discourse referencing. With this in mind, we narrowed our analyses to look only at those three-year-old stories that had the strongest narrative content. Of the 160 stories, there were 120 that seemed to convey a connected story-line, in which an event and its resolution could be identified. Even in these stories, 76% of the expressions used for the three referential acts combined were zero pronouns. The absence of contrastive referential expressions even in those stories with coherent content may merely reflect downstream sentence processing limitations. But when we also consider the pervasiveness of the absence coupled with evidence that the necessary forms were known, it seems likely that the three-year-olds lacked the cognitive ability to create and maintain a discourse model that incorporated their listener's knowledge and attentional states.

This latter possibility invites us to revisit the apparent early successes of the three-year-old speakers in Maintenance contexts. This appears to be a situation where successful communication is possible in spite of the limitations of the speaker. For adult speakers of Cantonese, zero pronouns are appropriately used for reference to characters that are already known and in focus, conditions that occurred only in the Maintenance contexts. When three-year-olds used zero pronouns in these contexts, they met the expectations of the 'expert' listeners and reference was judged to be adequate. Use of the same form in other contexts was judged to be inadequate. Note, however, that the resulting pattern of differential 'success' could reflect the minds of the listeners more than the minds of the children. We can objectively conclude that Maintenance contexts led to the highest levels of referential adequacy of the three we investigated, but this fact does not imply that the three-year-olds were actually successful in creating and consulting a discourse model.

By 5;0, the picture is quite different. One subgroup of the five-year-old children, like the seven- and twelve-year-olds, showed referential success in all three functions, implicating the combination of linguistic knowledge and presupposition skill required for each. The remaining five-year-olds showed referential adequacy when Introducing and Maintaining characters, but not when Reintroducing characters, even though they again demonstrated knowledge of the forms needed to do so.

Given that they could Maintain and Introduce characters adequately, these children seemed generally able to construct, and consult with, a mental model of the discourse which incorporated some aspects of the listener's knowledge. However, the Reintroduction function is unique in demanding judgements not only about whether a character is already known to the listener, but also about the listener's relative attention to that character. Discourse models require ongoing adjustment of the 'attentional status' of

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

all referents. The child must decide whether the referent is a character that has just been mentioned and is thus in the listener's attentional foreground, or is a character that was mentioned in an earlier utterance and has since been backgrounded by other story material. Difficulty with these decisions about attentional focus would lead to an inappropriate presupposition and ultimately to inadequate reference in Reintroduction contexts, even for children who knew the necessary linguistic forms.

Taken together, these findings indicate that Cantonese speaking children first use referential expressions that accurately presuppose their listener's knowledge around 5;0. As reviewed earlier, this is in accord with other studies of reference in children learning either Mandarin (Hickmann & Liang, 1990) or Cantonese (Wong, 1998). However, using similar story material, Emslie & Stevenson (1981) found that for children learning English, even three-year-olds could demonstrate contrastive use of definite and indefinite expressions. This difference between language groups suggests that, as predicted, the optionality of the NP in Cantonese may impede children's understanding of listener's needs, and their relationship to referring expressions. Direct cross-linguistic comparisons with controlled story materials will be needed to confirm this possibility.

The findings in this study represent an interesting contrast to the body of literature which examines the development of discourse referencing in spontaneous conversations (e.g. Clancy, 1996; Allen, 2000). Korean children between 1;8 and 2;10 (Clancy, 1996) and Inuktitut children between 2;0 and 3;6 (Allen, 2000) were found to use different referring expressions for known and new entities, and in different discourse-pragmatic contexts (e.g. the presence and absence of the referents) as well. This demonstration of early presuppositional abilities is not likely to be attributed to cross-linguistic differences. Korean and Inuktitut are similar to Cantonese in that they also allow extensive use of zero pronouns, and furthermore, like Cantonese, Korean has no verbal inflections. Given that children are conversationalists before they become story-tellers, it is plausible that narratives pose a higher cognitive, processing and linguistic demands on young children, particularly preschoolers, and hence their developing presuppositional abilities are compromised. This plausibility needs to be confirmed by studies looking at the contrastive use of referring expressions in conversation, and in stories, from the same group of children learning the same first language.

Processing demands on discourse referencing

As well as indicating the development of presuppositional abilities, findings from this study invite comment on another, less studied, aspect of discourse referencing, namely its cognitive processing requirements. Our data indicate that children's ability to make clear reference to characters is affected by the

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

discourse position at which a referential act is required. Given that speakers structure their messages based on a mental model of the discourse and that the management of changes in this model entails complex cognitive processes, it is not surprising that the context can affect referential success. We had expected that referential acts occurring later in the discourse would be less adequate since there would be more information to manage across a longer history of change. Given that our working memory is a finite and limited capacity information processing system, so when it is stretched to its limits, some kind of trade-offs or errors are expected (Kail & Bisanz, 1982). But such was not the case. Children were MORE successful with Introduction or Reintroduction of a character when the referential act occurred late rather than early in the discourse. There was virtually no evidence that the accumulation of referents and other meaning relations in the speaker's mental model resulted in inaccurate presuppositions, and hence inadequate reference, at least not for stories that were simple by design.

It seems possible that the observed referential success was due to the relative saliency of a character change when it occurred later in the discourse. When the Introduction function occurs early in the story, the child must refer to the new character right after the protagonist has been introduced. However, when Introduction occurs late, the child refers to the new character only after he has been Maintaining reference to the protagonist for several story frames. Utterance after utterance, the child finds no need to change his mental model, as the just mentioned referent is already in his model and in the listener's attentional focus. This extended history of no-change must highlight the differences created by the arrival of a new character and draw the child's attention to the detail of his mental model. Successful presupposition would thus be facilitated. Alternatively, the arrival of a new character may simply capture the child's attention, with fuller reference to the character coming from his own renewed interest rather than from any assessment of the listener's need. This explanation would find support in studies which report that very young children are more likely to encode the changing, and hence interest drawing, elements in the environment (Greenfield & Westerman, 1978; Greenfield & Zukow, 1978). Our data do not allow us to choose between these explanations. It does seem likely, however, that the expected cognitive processing demands of extended narrative were overridden by a saliency effect. It remains possible that in longer and more complex stories, the expected processing costs will be found.

CONCLUSION

This study reports clear developmental patterns of reference in the narratives of Cantonese speaking children from 3;0 to 12;0. Success at

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

Maintaining and Introducing a character was seen at an earlier age than success at Reintroducing a character, and even after children showed considerable referential ability, they were influenced by the nature and number of preceding referential acts.

There are two aspects of this developmental picture that seem new and particularly worth pursuing. First, we found evidence for the successive emergence of two sorts of cognitive presupposition underlying adequate reference, presupposition about the listener's current knowledge of a referent, and presupposition about the listener's relative attention to a referent. Secondly, for Cantonese speaking children, the ability to incorporate these presuppositions into a discourse model and use this model to ensure referential adequacy appears to emerge during the late preschool years, somewhat later than has been reported for English. If confirmed by more closely controlled cross-linguistic studies, this finding would indicate that the high rate of NP optionality in Cantonese does have a language learning cost.

REFERENCES

Allen, S. E. M. (2000). A discourse-pragmatic explanation for argument representation in child Inuktitut. *Linguistics* **38**, 483–521.

Bamberg, M. G. W. (1987). *The acquisition of narratives: learning to use language*. Berlin: Mouton de Gruyter.

Berman, R. A. & Slobin, D. I. (1994). *Relating events in narrative: a crosslinguistic developmental study*. Hillsdale, NJ: Erlbaum.

Brown, G. & Yule, G. (1983). *Discourse analysis*. Cambridge: CUP.

Brown, R. (1973). *A first language – the early stages*. Middlesex, England: Penguin Books.

Burgemeister, B., Blum, H. & Lorge, L. (1972). *The Columbia Mental Maturity Scale*. New York: Psychological Association.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (ed.), *Subject and topic*. New York: Academic Press.

Chao, Y.-R. (1968). *A grammar of spoken Chinese*. Berkeley & Los Angeles: University of California Press.

Clancy, P. M. (1992). Referential strategies in the narratives of Japanese children. *Discourse Processes* **15**, 441–67.

Clancy, P. M. (1996). Discourse motivations for referential choice in Korean acquisition. In H.-M. Sohn & J. Haig (eds), *Korean/Japanese linguistics*, Vol. 6. Stanford: Stanford Linguistics Association by the Center for the Study of Language and Information.

Clark, H. H. & Haviland, S. E. (1974). Psychological processes as linguistic explanation. In D. Cohen (ed.), *Explaining linguistic phenomena*. Washington, DC: Hemisphere.

Emslie, H. & Stevenson, R. (1981). Preschool children's use of the articles in definite and indefinite referring expressions. *Journal of Child Language* **8**, 313–28.

Erbaugh, M. (1992). The acquisition of Mandarin. In D. I. Slobin (ed.), *The crosslinguistic study of language acquisition*, Vol. 3. Hillsdale, NJ: Erlbaum.

Fletcher, P., Leung, C.-S. S., Stokes, S. & Weizman, Z. (2000). *Cantonese pre-school language development: a guide*. Hong Kong: The University of Hong Kong.

Gerken, L. (2000). Examining young children's morphosyntactic development through elicited production. In L. Menn & N. B. Ratner (eds), *Methods for studying language production*. Mahwah, NJ: Erlbaum.

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

Givón, T. (1983). *Topic continuity in discourse: a quantitative cross-linguistic study*. Amsterdam: Benjamins.
Gomme, N. J. & Johnson, C. E. (1997). Pronominal reference in 3-year-olds' narratives. In E. V. Clark (ed.), *The proceedings of the Twenty-eighth Annual Child Language Research Forum*. Stanford: Centre for the Study of Language and Information.
Greenfield, P. M. & Westerman, M. A. (1978). Some psychological relations between action and language structure. *Journal of Psycholinguistic Research* **7**, 453–75.
Greenfield, P. M. & Zukow, P. G. (1978). Why do children say what they say when they say it? An experimental approach to the psychogensis of presupposition. In K. E. Nelson (ed.), *Children's Language*, Vol. 1. New York: Gardner Press, Inc.
Gundel, J. K., Hedberg, N. & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language* **69**(2), 274–307.
Hickmann, M. (1995). Discourse organization and the development of reference to person, space and time. In P. Fletcher & B. MacWhinney (eds), *The Handbook of Child Language*. Oxford: Blackwell.
Hickmann, M., Kail, M. & Roland, F. (1995). Cohesive anaphoric relations in French children's narratives as a function of mutual knowledge. *First Language* **15**, 277–300.
Hickmann, M. & Liang, J. (1990). Clause-structure variation in Chinese narrative discourse: a developmental analysis. *Linguistics* **28**, 1167–1200.
Huang, C.-T. J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry* **15**, 531–74.
Kail, M. & Hickmann, M. (1992). French children's ability to introduce referents in narratives as a function of mutual knowledge. *First Language* **12**, 73–94.
Kail, R. & Bisanz, J. (1982). Information processing and cognitive development. In L. Lipsitt & C. Spiker (eds), *Advances in child development and behaviour*. New York: Academic Press.
Karmiloff-Smith, A. (1981). The grammatical marking of thematic structure in the development of language production. In W. Deutsch (ed.), *The child's construction of language*. London: Academic Press.
Lee, T. H. T. (2000). Finiteness and null arguments in child Cantonese. *The Tsinghua Journal of Chinese Studies*, New Series **30**, 101–28.
Lee, T. H.-T. & Szeto, K. (1996). The acquisition of (in)definiteness in Cantonese. In T. H.-T. Lee, C. H. Wong, S. Leung, P. Man, A. Cheung, K. Szeto & C. S.-P. Wong (eds), *The development of grammatical competence in Cantonese-speaking children*. Report of a project funded by Hong Kong Research Grant Committee CUHK 2/91.
Levelt, W. J. M. (1989). *Speaking – from intention to articulation*. Cambridge, MA: MIT.
Li, C. & Thompson, S. A. (1981). *Mandarin Chinese: a functional reference grammar*. Berkeley: University of California Press.
Li, P., Bates, E. & MacWhinney, B. (1993). Processing a language without inflections: a reaction time study of sentence interpretation in Chinese. *Journal of Memory and Language* **32**, 169–92.
Matthews, S. & Yip, V. (1994). *Cantonese: a comprehensive grammar*. London: Routledge.
Orsolini, M., Rossi, F. & Pontecorvo, C. (1996). Re-introduction of referents in Italian children's narratives. *Journal of Child Language* **23**, 465–86.
Peterson, C. (1990). The who, when and where of early narratives. *Journal of Child Language* **17**, 433–55.
Reynell, J. & Huntley, M. (1985). *Reynell Developmental Language Scales: Revised*. Windsor: NFER-Nelson.
Warden, D. A. (1976). The influence of context on children's use of identifying expressions and references. *British Journal of Psychology* **67**(1), 101–12.
Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* **13**, 103–28.
Winer, B. J., Brown, D. R. & Michels, K. M. (1991). *Statistical principles in experimental design*. New York: McGraw Hill.

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

Wong, M. Y. A. (1998). *The contrastive use of referring expressions in Cantonese-speaking children*. Paper presented at the The First Asia-Pacific Conference on Speech, Language and Hearing, Hong Kong, October 1998.
Zhu, D. (1982). *Yufa jiangyi* (*Grammar Lecture Notes*). Beijing: Commercial Press.

APPENDIX A

SAMPLE SENTENCES FOR THE IMITATION PROBE

Existential expression

jau5 jat1 zek3 si1zi2 hai2dou6 fan3gaau3
have one CL lion here sleep
'There is a lion sleeping.'

Relative NP

Coeng3 gan2 go1 go2 zek3 zoek3zai2 hou2 leng3
sing ASP song that CL bird very pretty
'The bird which is singing is very pretty.'

Determiner ling6ngoi6 'another' NP

ling6ngoi6 jat1 tiu4 gam1jyu2 m4 sik6 je5
another one CL gold fish not eat things
'A different gold fish is not eating.'

Demonstrative NP

ji1 zek3 tou3zai2 sik6 gan2 hung4lo4baak6
this CL rabbit eat CL carrot
'This rabbit is eating carrot.'

APPENDIX B

TRANSCRIPTS OF THE STORY DANCING FROM A THREE-YEAR-OLD, FIVE-YEAR-OLD, SEVEN-YEAR-OLD AND TWELVE-YEAR-OLD

A three-year-old (*Child III1*):

duk6 syu1
read book
'read a book.'

taan4 kam4
play piano
'play the piano.'

zoek3 haai4
put-on shoe
'put on shoes.'

tiu3mou5
dance
'dance.'

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

A five-year-old (*Child V1*):

cung4cin4 le1 zau6 jau5 jat1 go3 neoi5zai2 hai2dou6
once-upon-a-time PRT then have one CL girl here
tai2 gan2 syu1
read ASP book
'Once upon a time, there was a girl reading a book.'

gan1zyu6 ling6ngoi6 jat1 go3 neoi5zai2 zau6 hai2dou6 taan4 gan2 kam4
afterward another one CL girl then here play ASP piano
'Then another girl was playing the piano.'

gan1zyu6 keoi5 zau6 zoek3 deoi3 ba1leoi4mou5 haai4
afterward she then put-on CL ballet shoe
'Then she put on her ballet shoes.'

gan1zyu6 keoi5 zau6 hai2dou6 tiu3mou5
afterward she then here dance
'Then she danced.'

A seven-year-old (*Child VII3*):

cung4cin4 jau5 go3 neoi5zai2 hai2dou6 tai2 syu1
once-upon-a-time have CL girl here read book
'Once upon a time, there was a girl reading a book.'

jin4hau6 jau5 jat1 go3 neoi5zai2 hai2dou6 taan4 kam4
afterward have one CL girl here play piano
'Then, there was a girl playing the piano.'

jin4hau6 tai2 syu1 go2 go3 neoi5zai2 haang4 zo2 lei4 zoek3
afterward read book that CL girl walk ASP here put-on
haai4 tiu3mou5 la3
shoe dance PRT
'Then the girl who was reading walked over and put on her shoes and danced.'

jin4hau6 keoi5 hai2dou6 tiu3mou5
afterward she here dance
'Then she danced.'

A twelve-year-old (*Child XII2*):

jau5 jat1 go3 neoi5zai2 le1 tai2 gan2 jat1 di1 gwaan1jyu1 lik6si2
have one CL girl PRT read ASP one CL concerning history
ge3 syu1bun2
PRT book
'There is a girl reading some history books.'

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press

gam2 le1 keoi5 ge3 ga1ze1 zau6 hai2dou6 taan4 gan2 di1 kam4
so PRT she POSS sister then here play ASP CL piano
hou2 mei5miu6
very lovely
'Then her sister is playing some nice music.'

gam2 go2 go3 neoi5zai2 le1 zau5 dat6jin4zi1gaan1 le1 wun6
so that CL girl PRT then suddenly PRT change
zo2 keoi5 zoek3 go2 deoi3 ba1leoi4 haai4
ASP she wear that pair ballet shoes
'Then that girl suddenly puts on her ballet shoes.'

ceoi4 zoek6 di1 go1sing1 hai2dou6 tiu3 mou5
follow CL song here dance
'(She) is dancing to the music.'

https://doi.org/10.1017/S030500090400604X Published online by Cambridge University Press