



Title	Argument omission in Cantonese preschool children: a discourse-pragmatics perspective
Other Contributor(s)	University of Hong Kong.
Author(s)	Ho, Ka-yan, Agnes; 何家欣
Citation	
Issued Date	2006
URL	http://hdl.handle.net/10722/50057
Rights	Creative Commons: Attribution 3.0 Hong Kong License

**Argument omission in Cantonese preschool children:
a discourse-pragmatics perspective**

Ho Ka Yan, Agnes

A dissertation submitted in partial fulfillment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, June 30, 2006

**Argument omission in Cantonese preschool children:
a discourse-pragmatics perspective**

Ho Ka Yan, Agnes

Abstract

Discourse-pragmatics accounts for argument omission assume that referents are more likely to be encoded as overt arguments when they are more informative due to the presence of alternatives or uncertainties in the discourse or physical context (Allen, 2000). This paper employed the set of informativeness features described in Allen (2000) to examine whether discourse-pragmatics accounts could explain argument omission in Cantonese children aged 2;06 - 5;06. Results showed that children omitted subject arguments more often than object arguments for two of the three verbs examined; and this pattern did not change with age. As predicted by the discourse-pragmatics accounts, children tended to use overt argument when it carried higher informativeness value. However, given the relatively low informativeness values obtained even for overt arguments, and the same informative value observed in overt and null arguments for instances, it is likely that discourse-pragmatics accounts alone could not explain argument omission in young children.

Introduction

Children at the early stage of language learning tend to omit the subject and/or object arguments of the verb in a sentence. This phenomenon occurs in languages which require overt arguments such as English, as well as in languages which permit null arguments such as Cantonese (Lee, 2000; Man, 1996) and Inuktitut (Allen, 2000).

Three types of accounts were proposed to explain the phenomenon of argument omission. They are grammatical accounts, performance accounts, and discourse-pragmatics accounts. Grammatical accounts assume that children are born with a grammar that permitted argument omission. Children who are exposed to languages that do not allow argument omission will mature and adjust their grammar to the setting of adult speakers between the ages of two and four. In one grammatical account, subject omission was found to relate to acquisition of I(inflectional)-feature parameter which governs finiteness marking (Sano & Hyams, 1994). Once children begin to use finite markers such as modal verbs and third person singular –s, argument omission becomes less likely.

Performance accounts are related to the limited sentence-processing ability of children. They assume that child's grammar is the same as the adults', but due to limited sentence-processing ability children omit certain words in a sentence, leading to argument omission (Wang, Lillo-Martin, Best, Levitt, 1992). Some of the performance accounts suggest that the processing load is heavier at the beginning of the sentence, where the subject is, than at the end of the sentence, where the object is; and therefore subject omission is more frequent than object omission in child language.

Discourse-pragmatics accounts argue that argument omission in child language can be explained by children's sensitivity to the informativeness of the event described in a sentence (Allen, 2000). The Principle of Informativeness (Greenheld & Smith, 1976), as reported in Allen (2000), states that aspects of events that are more informative due to the

presence of alternatives and uncertainties in the discourse or the physical context, which cause ambiguity to the listeners, are more likely to be encoded by child, and aspects of events which were less informative tend to be omitted. In the context of argument use and omission, the more informative the referent in the event is, the more likely it is to be overtly encoded. According to Allen (2000), informativeness of referents depends on three groups of features, including knowledge features, confusion features, and search-space features. Knowledge features, which concern the presence of the referent in the speaker's and listener's joint knowledge, include ABSENCE, NEWNESS, and QUERY. Confusion features, which concern the potential confusion of the identity of the referent, include CONTRAST, DIFFERENTIATION IN CONTEXT, and DIFFERENTIATION IN DISCOURSE. Search-space features, which concern the relative size of the search space, include INANIMACY and THIRD PERSON. Table 1 shows how these features can be defined as informative and uninformative.

Previous studies on null argument in child Cantonese included Lee (2000) and Man (1996). Lee (2000) used grammatical accounts to explain argument-drop in Cantonese children aged between 1;07 and 2;08, but he concluded that the principles of grammatical accounts such as I(inflexion)-feature parameter did not apply to child Cantonese. Man (1996) studied the same group of participants as Lee (2000) and found that children were more likely to omit the subject arguments for ditransitive verbs such as 'bei2' (give) than other transitive verbs; and young children showed less subject omission as they matured with age. She also concluded that semantic and pragmatic factors played a role in argument omission. The subject omission rate in the matrix clause with matrix 'mental' verb such as 'zung1ji3' (like), for which subject omission would cause ambiguity to the listener due to the absence of the event denoted by the mental verbs in the physical context, was lower than simple transitive sentences. Although Man (1996) examined the effect of pragmatic factors

Table 1. *Definition of informativeness features taken from Allen (2000)*

Features	Informative	Uninformative
1. ABSENCE	referent absent from the physical context	referent present in the physical context
2. NEWNESS	referent new to the discourse (within 20 preceeding utterances)	referent not new to the discourse (within 20 preceeding utterances)
3. QUERY	referent that is the subject of or answer to query	referent that is not the subject of or answer to query
4. CONTRAST	contrast emphasized between potential referents	no contrast emphasized between potential referents
5. DIFFERENTIATION IN CONTEXT	two or more potential referents in the physical context	only one potential referent in the physical context
6. DIFFERENTIATION IN DISCOURSE	two or more potential referents in preceeding 20 utterances	only one potential referent in preceeding 20 utterances
7. INANIMACY	inanimate referent	animate referent
8. THIRD PERSON	third person referent	first or second person referent

on argument drop, it covered only a part of the discourse-pragmatics accounts. Since there were previous studies on argument omission in child Cantonese concluding that pragmatic factors contributed in the decision of whether to represent the argument overtly (Man, 1996) and other accounts such as grammatical accounts did not adequately explain the phenomenon of argument omission (Lee, 2000), this study would further investigate whether discourse-pragmatics accounts could explain argument omission in child Cantonese. In the study of Inuit children (Allen, 2000), it was concluded that they paid attention to

informativeness features in choosing overt or null arguments to represent the arguments. Similar to the study on Inuit children (Allen, 2000), this study would examine whether Cantonese children decided the use of null argument using informativeness features.

In this paper, the use of null arguments produced by Cantonese preschool children was described. This study investigated whether the use of null arguments was affected by age and sentence position of the argument; and further investigated whether the informativeness values based on discourse-pragmatic accounts differed for overt and null arguments and changed with age of the children. In two sets of analyses, this paper attempted to answer these questions:

Analysis 1:

- (1) Does the percentage use of null arguments change with age?
- (2) Does the percentage of null arguments differ for the subject and the object position?

Analysis 2:

- (3) Do informative values change with age?
- (4) Do informative values differ for overt and null arguments in the subject position?
- (5) Do informative values differ for overt and null arguments in the object position?

The study employed some of the informativeness features used in the study of Allen (2000), which included NEWNESS, QUERY, CONTRAST, DIFFERENTIATION IN DISCOURSE, INANIMACY, and THIRD PERSON. Other features used in Allen (2000), including ABSENCE and DIFFERENTIATION IN CONTEXT were not employed in this study since the data in this study came from written transcripts of language samples in a database, and these features concerned the presence or alternatives of the events in the physical context which could not be reflected in these transcripts.

Methodology

Participants

Language samples in a cross-sectional database (Fletcher, Leung, Stokes, & Weizman, 2000) were used as the source of data in this study. The participants of this database were 70 typically-developing, Cantonese-speaking preschool children. They were recruited from a Cantonese-speaking pre-school in Hong Kong, and were predominantly middle-class though socioeconomic status was not taken into account during recruitment. These children, aged between 2;06 and 5;06, were classified into seven age groups, in which each consisted of an age range of 6 months. There were five boys and five girls in each age group. The language sample was collected when the child was engaged in a conversation with an investigator during play around familiar bath/dress/feed/sleep routines in about 20 minutes, after a warm-up task at the beginning of the session. Appendix A presents the mean number of utterances and MLU of each age group.

Procedure

To examine argument omission, there were two options. One was to identify whether the subject and the object arguments were used for each of the transitive verbs in the samples. Another option was to identify argument omission from utterances with certain transitive verbs only. The second option was adopted for this study because this could help the study to be more focused.

Initial screening was carried out to select verbs, which would be included in the study, according to the following criteria: (1) They had to be action transitive verbs, which typically entailed the use of subject and object arguments; (2) they had to show a high frequency of use in the samples for each age group, so as to ensure that there were sufficient tokens in each age group for analysis. The 'freq' command of the CLAN program (MacWhinney & Snow, 1990) was run to identify the most frequently-occurred verbs. The action transitive verbs were then ranked according to their frequencies. The verbs 'sik6' (eat) and 'zoek3' (wear) were chosen because they ranked at the first and the second places

respectively and their frequencies were evenly distributed across the age groups. During the selection of verbs, it was interesting to find that some action verbs often appeared in isolation without either the subject, the object, or both. The presence of argument omission in utterances with these verbs might be different from others. The verb ‘zit1’ (squirt), which was one of these verbs, was therefore chosen to examine whether there were verb specific patterns for argument omission. After the screening, three action transitive verbs, ‘sik6’ (eat), ‘zoek3’ (wear), and ‘zit1’ (squirt), were selected as target verbs.

All utterances with the target verbs were extracted from the samples. Utterances which were multiple sentences, serial verb constructions, questions, and non-spontaneous production (i.e. routines such as songs and rhymes, and repetition due to repair or self-correction) were excluded, because this study focused on simple declarative sentences only, and the pattern of use of argument omission could only be reflected from spontaneous production. The numbers of occurrence of ‘sik6’ (eat), ‘zoek3’ (wear), and ‘zit1’ (squirt) excluded across all age groups were 17 – 36, 8 – 24, and 1 – 13 respectively. Table 2 shows the numbers of occurrence of these in each age group after exclusion.

Table 2. *Number of occurrence of the three target verbs included in the analysis*

	Number of occurrence of the verb included		
	‘sik6’ (eat)	‘zoek3’ (wear)	‘zit1’(squirt)
Age Group 1 (2;06)	35	51	56
Age Group 2 (3;00)	50	55	23
Age Group 3 (3;06)	46	34	43
Age Group 4 (4;00)	82	36	24
Age Group 5 (4;06)	77	61	62
Age Group 6 (5;00)	73	52	38
Age Group 7 (5;06)	63	30	15

Coding

In Analysis 1, the subject and the object positions of each target verb token were coded for overtness, that is either overt or null. Examples were listed in (1) and (2). Arguments that were coded as overt arguments included nouns or pronouns, (1a), generics, (1b), deictic, (1c), and noun phrases with the nouns omitted, (1d). Arguments that were coded as null arguments included absence of argument, (2).

(1) a. sik6 min6bao1 (MCL 4;00)

eat bread

“Eat bread”

b. sik6 je5 (CKY 3;11)

eat things

“Eat food”

c. zoek3 li1 go3 (NMC 3;06)

wear this

“Wear this”

d. sik6 jat1 lap1 ˊˌ (YCY 3;11)

eat one piece

“Eat a piece (of candy)”

(2) zoek3 ˊˌ (NMC 3;06)

wear

“Wear (clothes)”

After each argument were coded as overt or null, the percentage of null arguments in the subject and object positions of utterances with each target verb were calculated, with the sum of number of overt and null arguments as denominators, for each participant.

In Analysis 2, six informativeness features, including CONTRAST, DIFFERENTIATION IN DISCOURSE, INANIMACY, NEWNESS, QUERY, and THIRD PERSON, were employed (Allen, 2000). The definitions of the features were the same as those described in Allen (2000), which are listed in Table 1 in the introduction section. Recall that, for the feature CONTRAST, the argument was informative when there was contrast emphasized with other potential referents and uninformative when there is no contrast between the potential referents. For the feature DIFFERENTIATION IN DISCOURSE, the argument was informative when there were more than one potential referent in the preceding 20 utterances and uninformative when only one potential referent existed in the preceding 20 utterances. For INANIMACY, the argument was informative if it was inanimate and uninformative if it was animate. For NEWNESS, the argument was informative when the referent was new to discourse in the preceding 20 utterances and uninformative when the referent was not new to discourse in the preceding 20 utterances. For QUERY, the argument was informative when the referent was the subject or answer to the query and uninformative when the referent was not. For THIRD PERSON, the argument was informative when the referent was not the first and second person referent and uninformative when the referent was the first or second person referent. The argument was scored as 1 if it was informative and 0 if it was uninformative. The informativeness value was determined for each of the features in the context of 20 preceding utterances for the subject and the object position of each target verb token regardless of whether the position was actually filled by an argument. The informativeness score for each token, which was the sum of information values of the six features, was then calculated. The maximum score for each token would be 6. The higher the informativeness score, the more informative the referent is and the more likely the referent is overtly encoded according to the discourse-pragmatics accounts.

To illustrate how the tokens are coded, an example extracted from the written transcript of a participant (YCY, 3;11) is shown as below:

CHI: e6 maai5 di1 ping4gwo2! (preceeding 17th utterance)

buy CL apples

“Buy some apples”

σ

CHI: caang2! (preceeding 15th utterance)

orange

“Orange”

σ

INV: nei5 zung1ji3 sik6 mat1je5 gaa3? (preceeding 1st utterance)

you like eat what sfp

“What do you like to eat?”

CHI: ˊ. zung1ji3 *sik6* joeng4cung1.

like eat onion

“(I) like eating onion”

In the example, the subject (I) was omitted and the object ‘joeng4cung1’ (onion) was overtly encoded. For each informativeness feature, the subject and the object were scored as 1 when the referent was informative and 0 when the referent was uninformative. In the subject position, ‘I’ was a first person and animate referent, so it was scored as 0 for the features THIRD PERSON and INANIMACY. It was also scored as 0 for the feature NEWNESS, because first person referent was classified as non-new (Allen, 2000). The subject ‘I’ was not the answer of the investigator’s question, and it was not used to emphasize the contrast with another potential referent, so it was scored as 0 for the features QUERY and CONTRAST. For the feature DIFFERENTIATION IN DISCOURSE, since

the investigator's question had the subject 'nei5' (you) that obviously indicated the investigator was asking the child, the null subject 'I' in the child's answer was scored as 0 because it was the only potential referent. The informativeness score of the subject 'I' was therefore 0, which was the sum of the informativeness values of the six features. In the object position, 'joeng4cung1' (onion) was not a first or second person referent and was an inanimate referent, so it was scored as 1 for each of the features THIRD PERSON and INANIMACY. It was not mentioned in the preceding 20 utterances, so it was new to the listener and therefore was scored as 1 for the feature NEWNESS. 'joeng4cung1' (onion) was the answer of the investigator's 'what' question, so it was scored as 1 for the feature QUERY. The referents 'ping4gwo2' (apple) and 'caang2' (orange) mentioned in the preceding 15th and 17th utterances could fit the verb semantics of 'sik6' (eat), so there were more than one potential referent and the object 'joeng4cung1' (onion) was scored as 1 for the feature DIFFERENTIATION IN DISCOURSE'. For the feature CONTRAST, the referent 'joeng4cung1' (onion) was not used to emphasize the contrast with another referent, so it was scored as 0. The informativeness score of the object in this example was therefore 5. Table 3 shows the information scores for the null subject 'I' and the overt object 'onion' in the example from YCY, 3;11.

Table 3. *Informativeness scores for the subject and object in the example (YCY, 3;11)*

	Informativeness features						Informativeness score
	C	D	I	N	Q	T	
Subject: '· ·' (I)	0	0	0	0	0	0	0
Object: 'joeng4cung1' (onion)	0	1	1	1	1	1	5

C = Contrast D = Differentiation in discourse I = Inanimacy
N = Newness Q = Query T = Third person

Reliability

Inter-rater reliability analysis included two parts, one on the inclusion of utterances and another on the informativeness scores. For the first part, the inter-rater reliability was completed on 20% of the utterances for a total of 1333 utterances. Utterances were randomly chosen from transcripts of all the participants. The inter-rater agreement for this part of inter-rater reliability analysis was 85.17%. For the second part, the inter-rater reliability level was completed on 15% of the tokens for a total of 1468 tokens. Tokens were randomly chosen from all the tokens from both of the verbs 'sik6' (eat) and 'zoek3' (wear) in subject and object positions. The inter-rater agreement for this part was 85.86%.

Data analysis

There were two sets of analyses. The first analysis examined whether the percentage of null arguments changed with age and whether it differed in the subject and object positions. The second analysis examined whether informativeness scores changed with age and whether they differed for overt and null arguments in the two argument positions. Descriptive statistics were used to determine the mean, standard deviation, and the percentage range of null arguments and informativeness scores.

For inferential statistics, two-way repeated measures ANOVA was administered for the first analysis to confirm the effects of age and argument position on the percentage of null arguments; while two-way ANOVA was used in the second analysis to confirm whether the informativeness scores changed with age and the overtness of the arguments.

Results

Occurrence of null arguments in the subject and object positions

The first analysis examined whether the percentage use of null arguments changed with the age of the children and with the position of the argument. For the verb 'sik6' (eat), the mean, standard deviation, and percentage range of null arguments in the subject and

object positions for each age group are summarized in Table 4. The means ranged from 52.11% to 88.89% in the subject position and from 18.28% to 38.83% in the object position.

Table 4. *Mean (standard deviation) and percentage range of occurrence of null subjects and null objects in utterances with the verb 'sik6' (eat) for each age group*

	% Occurrence of null arguments in			
	Subject Position		Object Position	
	Mean (SD)	Range ($\pm 95\%$)	Mean (SD)	Range ($\pm 95\%$)
Age Group 1 (2;06)	88.89%	63.27% -	31.02%	6.99% -
(n = 9)	(33.33%)	100.00%	(31.26%)	55.05%
Age Group 2 (3;00)	79.19%	61.23% -	33.55%	8.64% -
(n = 9)	(23.36%)	97.15%	(32.40%)	58.46%
Age Group 3 (3;06)	66.37%	36.70% -	38.83%	6.06% -
(n = 10)	(41.48%)	96.04%	(45.81%)	71.61%
Age Group 4 (4;00)	79.21%	65.97% -	27.98%	10.36% -
(n = 9)	(17.22%)	92.45%	(22.92%)	45.60%
Age Group 5 (4;06)	81.02%	67.61% -	18.28%	5.40% -
(n = 10)	(18.75%)	94.43%	(18.01%)	31.17%
Age Group 6 (5;00)	74.60%	57.33% -	30.31%	19.34% -
(n = 8)	(20.66%)	91.87%	(13.13%)	41.29%
Age Group 7 (5;06)	52.12%	33.62% -	37.29%	9.78% -
(n = 9)	(24.07%)	70.62%	(35.78%)	64.80%

The mean percentage of occurrence of the null subject was consistently higher than that of the null object across the age groups. In both of the subject and object positions, the percentage of occurrence of null argument was different among the age groups, but within

group variability appeared very high for some age groups, and Group 3 in particular. Two-way repeated measures ANOVA, argument position (2) and age group (7), was implemented to statistically evaluate these observations. The position of arguments had a significant main effect on the mean percentage occurrence of null arguments ($F(1,57) = 79.70, p = 0.00$), but age differences did not turn out to be significant ($F(6,57) = 0.04, p = 0.84$). There was no significant interaction between age and argument position ($F(6,57) = 1.78, p = 0.12$).

The percentage of occurrence of null arguments in utterances with the verb ‘zoek3’ (wear) was also analyzed. Table 5 shows the mean, standard deviation, and percentage range of null arguments for the verb ‘zoek3’ (wear) in the subject and object positions for each age group. The mean percentage of occurrence of null arguments in the subject position ranged from 67.04% to 96.34%, and those in the object position ranged from 18.81% to 44.28%. The data set for the verb ‘zoek3’ (wear) showed similar pattern as that for the verb ‘sik6’ (eat). The mean percentage of occurrence of null subject for the verb ‘zoek3’ (wear) was also consistently higher than that of the null object across the age groups. The percentages of occurrence of null arguments in the subject and object positions were quite different among the age groups though they were smaller than those in the data set of the verb ‘sik6’ (eat). Similar to the verb ‘sik6’ (eat), the within group variability was very high, and Group 4 and 7 in particular.

Two-way repeated measures ANOVA, argument position (2) and age group (7), was implemented to statistically evaluate the above observations. The position of arguments had a significant main effect on the mean percentage occurrence of null arguments ($F(1,55) = 94.75, p = 0.00$); but age differences were not significant ($F(6,55) = 1.70, p = 0.14$). There was no significant interaction between age and argument position ($F(6,55) = 0.67, p = 0.67$).

Table 5. Mean (standard deviation) and percentage range of occurrence of null subjects and null objects in utterances with the verb ‘zoek3’ (wear) for each age group

	% Occurrence of null arguments in			
	Subject Position		Object Position	
	Mean (SD)	Range ($\pm 95\%$)	Mean (SD)	Range ($\pm 95\%$)
Age Group 1 (2;06)	92.29%	82.98% -	44.28%	15.68% -
(n = 9)	(12.10%)	100.00%	(37.21%)	72.88%
Age Group 2 (3;00)	77.28%	59.86% -	20.74%	1.52% -
(n = 9)	(22.66%)	94.69%	(25.00%)	39.96%
Age Group 3 (3;06)	96.34%	90.46% -	28.79%	0.00% -
(n = 8)	(7.04%)	100.00%	(34.94%)	58.00%
Age Group 4 (4;00)	67.04%	35.00% -	25.56%	0.00% -
(n = 9)	(41.68%)	99.08%	(33.37%)	51.21%
Age Group 5 (4;06)	77.96%	60.33% -	35.84%	15.59% -
(n = 10)	(24.64%)	95.59%	(28.30%)	56.09%
Age Group 6 (5;00)	85.75%	72.64% -	18.81%	0.00% -
(n = 10)	(18.33%)	98.87%	(26.26%)	37.59%
Age Group 7 (5;06)	83.33%	56.63% -	40.69%	14.40% -
(n = 7)	(28.87%)	100.00%	(28.43%)	66.99%

The utterances with the third target verb ‘zit1’ (squirt) were also examined for the occurrence of null arguments. In most age groups, only about half of the participants produced utterances with this verb that were included in the study (Table 6). In addition, the individual variation on the number of the verb ‘zit1’ (squirt) produced by each participant was large (0 – 24 tokens) in all age groups (Table 6). Individual participant data were

therefore collapsed together to form group percentage of occurrence for further discussion and analysis.

Table 6. *Number of participants who produced utterances with the verb 'zit1' (squirt) and the range of the verb 'zit1' (squirt) produced*

	N	Range of the verb 'zit1' produced
Age Group 1 (2;06)	5	0 – 20
Age Group 2 (3;00)	5	0 – 11
Age Group 3 (3;06)	4	0 – 17
Age Group 4 (4;00)	4	0 – 16
Age Group 5 (4;06)	5	0 – 24
Age Group 6 (5;00)	6	0 – 14
Age Group 7 (5;06)	3	0 – 12

N = number of participants who produced utterances with the verb 'zit1' (squirt)

The percentage of occurrence of null arguments in the subject and object positions in each age group are listed in Table 7. They ranged from 80.00% to 100.00% in the subject position and from 87.50% to 100.00% in the object position.

The percentage of occurrence of null arguments for the verb 'zit1' (squirt) was very high for all age groups in both subject and object positions. The youngest and the oldest groups consistently omitted both arguments while the other groups did so between 80% and 98.28%. Wilcoxon matched pairs test, which was nonparametric test, was implemented to statistically confirm the above observations on the differences of the percentages in the subject and object positions. The percentages of null arguments in the subject position (mean = 94.83%, s.d. = 7.39%) and in the object position (mean = 93.82%, s.d. = 5.90%) were not significantly different ($T = 5.00, p = 0.50$).

Table 7. *Percentage of occurrence of null arguments in the utterances with the verb ‘zit1’ (squirt) for each age group*

	% Occurrence of null arguments in	
	Subject Position	Object Position
Age Group 1 (2;06)	100.00% (58/58)	100.00% (58/58)
Age Group 2 (3;00)	100.00% (24/24)	87.50% (21/24)
Age Group 3 (3;06)	90.91% (40/44)	88.64% (39/44)
Age Group 4 (4;00)	80.00% (24/30)	100.00% (30/30)
Age Group 5 (4;06)	98.28% (57/58)	91.38% (53/58)
Age Group 6 (5;00)	94.59% (35/37)	89.19% (33/37)
Age Group 7 (5;06)	100.00% (15/15)	100.00% (15/15)

Informativeness scores for overt and null arguments in the subject and object positions

The second analysis examined whether informativeness scores for the subject and object arguments changed with age, and whether informativeness scores differed for overt subject/object arguments and null subject/object arguments. In this analysis, the informativeness scores of arguments in the subject and object positions were analyzed separately. From the data set of the verb ‘zit1’ (squirt), it was found that there were no overt arguments produced in three of the age groups in both subject and object positions, making it impossible to examine the differences of informativeness scores between overt arguments and null arguments. Therefore, only utterances with the target verbs ‘sik6’ (eat) and ‘zoek3’ (wear) were analyzed. Table 8 and Table 9 show the mean, standard deviation, and range of informativeness scores, and number of overt and null arguments in the subject and object positions in the utterances with the verb ‘sik6’ (eat) for each age group. Recall that the maximum informativeness score was 6, and it was assumed that arguments were more likely to be overtly encoded when the informativeness scores are high. The mean

informativeness scores of overt arguments ranged from 0.88 to 1.35 and those of null arguments ranged from 0.59 to 1.03 in the subject position; while the mean informativeness scores of overt arguments ranged from 3.19 to 3.47 and those of null arguments ranged from 2.05 to 2.38 in the object position.

Table 8. *Mean (standard deviation) and range of informativeness scores, and number of overt and null subjects in utterances with 'sik6' (eat) for each age group*

	Subject Position					
	Overt Argument			Null Argument		
	N	Informativeness Score		N	Informativeness Score	
		Mean (SD)	Range (±95%)		Mean (SD)	Range (±95%)
Age Group 1 (2;06)	7	1.00 (1.15)	-0.07 – 2.07	27	0.59 (0.64)	0.34 – 0.84
Age Group 2 (3;00)	11	1.18 (1.08)	0.46 – 1.91	39	0.72 (0.86)	0.44 – 1.00
Age Group 3 (3;06)	17	0.94 (0.90)	0.48 – 1.40	30	0.67 (0.92)	0.32 – 1.01
Age Group 4 (4;00)	19	1.21 (1.13)	0.66 – 1.76	63	0.59 (0.78)	0.39 – 0.78
Age Group 5 (4;06)	17	1.35 (0.93)	0.87 – 1.83	59	1.05 (0.78)	0.85 – 1.25
Age Group 6 (5;00)	18	1.11 (0.96)	0.63 – 1.59	50	1.00 (0.70)	0.80 – 1.20
Age Group 7 (5;06)	25	0.88 (0.78)	0.56 – 1.20	34	0.79 (0.64)	0.57 – 1.02

N = number of arguments

The mean informativeness score in the object position was higher for overt arguments than null arguments, but those for overt and null arguments in the subject position were quite similar across all age groups. The mean informativeness scores were comparable across all age groups. Two-way ANOVA, argument position (2) and age group (7), was implemented to statistically confirm the above observations. For the subject position, a significant main effect of overtness was shown ($F(1,402) = 11.18, p = 0.00$)

Table 9. Mean (standard deviation) and range of informativeness scores, and number of overt and null objects in utterances with ‘sik6’ (eat) for each age group

	Object Position					
	Overt Argument			Null Argument		
	N	Informativeness Score		N	Informativeness Score	
		Mean (SD)	Range (±95%)		Mean (SD)	Range (±95%)
Age Group 1 (2;06)	22	3.32 (0.95)	2.90 – 3.74	12	2.33 (0.49)	2.02 – 2.65
Age Group 2 (3;00)	35	3.40 (0.85)	3.11 – 3.69	15	2.20 (0.41)	1.97 – 2.43
Age Group 3 (3;06)	31	3.42 (1.09)	3.02 – 3.82	16	2.38 (0.72)	1.99 – 2.76
Age Group 4 (4;00)	58	3.43 (0.94)	3.18 – 3.68	24	2.25 (0.53)	2.03 – 2.47
Age Group 5 (4;06)	55	3.36 (0.93)	3.11 – 3.62	21	2.05 (0.22)	1.95 – 2.15
Age Group 6 (5;00)	47	3.19 (0.74)	2.97 – 3.41	21	2.19 (0.40)	2.01 – 2.37
Age Group 7 (5;06)	36	3.47 (1.03)	3.12 – 3.82	23	2.39 (0.66)	2.11 – 2.68

N = number of arguments

while age had no significant effect ($F(6,402) = 1.54, p = 0.16$). Interaction between age and overtness was not significant ($F(6,402) = 0.72, p = 0.63$). For the object position, overtness also had a significant main effect ($F(1,402) = 155.01, p = 0.00$), but age had no significant effect ($F(6,402) = 0.67, p = 0.67$). Interaction between age and overtness was also not significant ($F(6,402) = 0.28, p = 0.95$).

For the utterances with another target verb ‘zoek3’ (wear), the mean, standard deviation, range of informativeness scores, and the number of overt and null arguments in the subject and object positions are shown in Table 10 and Table 11. The mean informativeness score of overt arguments ranged from 0.33 to 1.40 and that of null arguments ranged from 0.90 to 1.16 in the subject position; while the mean informativeness

score of overt arguments ranged from 3.00 to 3.57 and that of null arguments ranged from 2.00 to 2.27 in the object position.

Table 10. *Mean (standard deviation) and range of informativeness scores, and number of overt and null subjects in utterances with ‘zoek3’ (wear) for each age group*

	Subject Position					
	Overt Argument			Null Argument		
	N	Informativeness Score		N	Informativeness Score	
		Mean (SD)	Range (±95%)		Mean (SD)	Range (±95%)
Age Group 1 (2;06)	9	0.56 (0.73)	0.00–1.11	41	1.12 (0.71)	0.90–1.35
Age Group 2 (3;00)	10	1.40 (0.97)	0.71–2.09	45	1.16 (0.74)	0.93–1.38
Age Group 3 (3;06)	3	0.33 (0.58)	-1.10–1.77	31	0.90 (0.94)	0.56–1.25
Age Group 4 (4;00)	8	1.13 (0.83)	0.43–1.82	28	0.96 (0.96)	0.59–1.34
Age Group 5 (4;06)	13	0.62 (0.65)	0.22–1.01	43	0.91 (0.78)	0.67–1.15
Age Group 6 (5;00)	8	0.75 (0.71)	0.16–1.34	41	0.95 (0.84)	0.69–1.21
Age Group 7 (5;06)	3	1.33 (1.53)	-2.46–5.13	27	1.00 (0.92)	0.64–1.36

N= number of arguments

A similar pattern to the data set for the verb ‘sik6’ (eat) was observed in the data set for the verb ‘zoek3’ (wear). In the object position, the mean informativeness score was higher for overt arguments than null arguments, but those for overt and null arguments in the subject position were fairly similar across all age groups. The mean informativeness scores were similar across almost all age groups in the subject and object positions, except for particular exceptions such as the informativeness score of overt objects in group 4 which had a comparatively high value. Two-way ANOVA was also used to examine statistically the above observations. For the subject position, both age ($F(6,296) = 1.82, p = 0.09$) and

Table 11. Mean (standard deviation) and range of informativeness scores, and number of overt and null objects in utterances with ‘zoek3’ (wear) for each age group

	Object Position					
	Overt Argument			Null Argument		
	N	Informativeness Score		N	Informativeness Score	
		Mean (SD)	Range (±95%)		Mean (SD)	Range (±95%)
Age Group 1 (2;06)	34	3.03 (1.00)	2.68 – 3.38	16	2.00 (0.37)	1.81 – 2.19
Age Group 2 (3;00)	41	3.12 (0.78)	2.88 – 3.37	14	2.21 (0.43)	1.97 – 2.46
Age Group 3 (3;06)	23	3.42 (0.78)	3.01 – 3.68	11	2.27 (0.65)	1.84 – 2.71
Age Group 4 (4;00)	28	3.57 (0.79)	3.27 – 3.88	8	2.25 (0.71)	1.66 – 2.84
Age Group 5 (4;06)	35	3.11 (0.76)	2.85 – 3.37	21	2.10 (0.30)	1.96 – 2.23
Age Group 6 (5;00)	36	3.00 (0.76)	2.74 – 3.26	13	2.00 (0.00)	2.00 – 2.00
Age Group 7 (5;06)	18	3.44 (0.70)	3.09 – 3.79	12	2.17 (0.39)	1.92 – 2.41

N = number of arguments

overtness ($F(1,296) = 0.82, p = 0.36$) had no significant main effects on informativeness scores. For the object position, similar to the results of the data set for the verb ‘sik6’ (eat), overtness had a significant main effects on informative scores ($F(1,296) = 142.26, p = 0.00$), but age did not have significant main effect ($F(6,296) = 1.60, p = 0.15$). Interaction between age and overtness was not significant ($F(6,296) = 0.35, p = 0.91$).

Summary

The first set of results revealed that young children used higher percentage of null subject than null object for the verbs ‘sik6’ (eat) and ‘zoek3’ (wear), but they used similar percentage of null subject and null objects for the verb ‘zit1’ (squirt). The percentages of

null subject and null object for the verb ‘zit1’ were very high. All these patterns did not change with age.

The second set of results showed that the informativeness score for overt subject was higher than null subject, and that for overt object was higher than null object for the verb ‘sik6’ (eat). For the verb ‘zoek3’ (wear), the informativeness score for overt object was higher than null object, but that for overt subject and null subject was similar. The pattern did not change with age for both of the verbs ‘sik6’ (eat) and ‘zoek3’ (wear).

Discussion

The results of Analysis 1 presented in the above section indicated that null arguments were used more frequently in the subject position than the object position by Cantonese preschoolers for two of the three target verbs. This pattern did not change with age. This finding was shown in the analysis on utterances with the verb ‘sik6’ (eat) and was further supported from another transitive verb ‘zoek3’ (wear). The results for the third target verb ‘zit1’ (squirt), however, suggested this pattern of null marking might not be consistent across all transitive verbs.

The results of the data set for the verb ‘zit1’ (squirt) were different from those for the verbs ‘sik6’ (eat) and ‘zoek3’ (wear). The verbs ‘sik6’ (eat) and ‘zoek3’ (wear) belong to the class ‘achievement’ according to Vendler’s classification, while the verb ‘zit1’ (squirt) belongs to ‘semelfactives’. Semelfactives are a class of eventualities that are instantaneous but atelic (Rothstein, 2004). In the language samples, the participants tend to use the semelfactive ‘zit1’ (squirt) in isolation when describing their ongoing instantaneous action. The results revealed that null arguments, both in the subject and object positions, were of high percentages (nearly 100%); but percentages of null arguments were lower in the object position than in the subject position in the results for utterances with the achievements

'sik6' (eat) and 'zoek3' (wear). This suggested that the preference of use of null arguments might be affected by the aspectual classes of the verb in the utterance.

According to the results, the pattern of use of argument omission did not change with age. Cantonese preschool children aged from 2;06 to 5;06, showed a similar percentage use of null arguments in the subject and object positions. In a previous study on Cantonese children (Man, 1996), it was suggested that age, which was a nongrammatical factor, had an effect on the use of subject omission, but object omission was more resistant to age factor. The difference between the findings in this study and in Man (1996) might be explained by the age difference of the participants in the two studies. The age of the participants in Man (1996) was between 1;05 and 2;07, and the age of the participants of this study was between 2;06 and 5;06 which was higher than those in Man (1996). Besides, adults' percentage use of null arguments in the two argument positions should be examined in future studies, in order to compare it with the child data to see if there would be any change in the percentage use of null arguments when approaching adulthood. Any differences between the adult's and child's percentage use of null arguments might help to determine the accounts that explained argument omission in Cantonese preschool children. Grammatical accounts (Sano & Hyams, 1994) and processing accounts (Bloom, 1990) assume argument omissions are due to limitations of children's language or processing ability at the early stage of language development. There should be gradual or sudden change in percentage use of null arguments as children matured with age, if these two accounts could explain the use of argument omissions in Cantonese preschool children.

According to the results, null subject was used more often than null object. In the analysis of utterances with 'sik6' (eat) and 'zoek3' (wear), the percentage of null arguments was higher in the subject position than in the object position. Subject/object asymmetry was

observed. This will be further discussed with the informativeness scores in the later part of the discussion.

The results of Analysis 2 indicated that informativeness scores were higher for overt arguments than null arguments in the subject and object positions for the verb 'sik6' (eat) and in the object position for the verb 'zoek3' (wear). This suggested that arguments that were overtly encoded were more informative than those omitted. Recall that informativeness scores were the sum of features that were informative. This finding was consistent with research on argument representations of Inuit children, which suggested that discourse-pragmatics features of informativeness could be used to explain the use of argument omission (Allen, 2000). Allen (2000) also proposed through preliminary investigation that the effect of informativeness features was cumulative, that is the use of null argument might not be guided by one informativeness feature alone. The decision on the use of null argument might be guided by combination of informativeness features. The results of this study, which showed that the argument with higher informativeness score was more likely to be overtly encoded, might support the cumulative effect of informativeness features in deciding whether to represent the argument as overt or null marking.

As mentioned in the previous paragraphs, the results of Analysis 1 showed subject/object asymmetry, in which the percentage of null subject was much higher than that of null object, for the verbs 'sik6' (eat) and 'zoek3' (wear). In Analysis 2, it was found that the informativeness scores for both overt and null arguments were much lower in the subject position than in the object position. This appeared to match with the pattern observed in the results of Analysis 1 based on discourse-pragmatics accounts, because the accounts assume that arguments are more likely to be omitted when they are less informative (Allen, 2000), that is having lower informativeness scores. The informativeness scores for the subject arguments were relatively lower than the object arguments, so null

subjects were more frequently used than null objects as observed in the results. For the low informativeness scores in the subject position, there was a possible explanation about the information value of the features in a conversation situation. When the investigator engaged in a conversation with the child, they often referred to themselves or one another, instead of somebody else in the subject position. Reference to first or second person was considered as uninformative for the features THIRD PERSON, INANIMACY, and NEWNESS, which were already half of the total six informativeness features. However, in the object position, they referred to not only themselves and each other, but also other people or things. This led to lower informativeness scores in the subject position than in the object position. Discourse-pragmatics accounts therefore could explain why the subject position was more likely to have null arguments than the object position.

Discourse-pragmatics accounts could explain the use of argument omission in Cantonese preschoolers but not completely. This could be accounted by the following three reasons. The first reason was that the pattern of higher informativeness scores for overt arguments than null arguments in the subject position observed in the result was not consistent in the two argument positions. The informativeness scores for overt and null arguments were not different in the subject position for the verb 'zoek3' (wear). This was different from the pattern observed in the object position for the verb 'zoek3' (wear) and in the subject and object positions for the verb 'sik6' (eat). Second, the informativeness scores for overt arguments in the object position were relatively low, mostly around 3 to 4 out of a total of 6, indicating that only about half of the informativeness features were informative. The third reason was that the informativeness scores for some overt tokens were the same as those for some null tokens. According to discourse-pragmatics accounts, it was expected that the overt arguments would have higher informativeness scores than the null arguments. For those overt tokens which had the same score as some null tokens, discourse-pragmatics

accounts alone could not explain why they were overtly encoded. For these reasons, it was concluded that the number of informativeness features that were informative could not solely be used in determining whether the argument should be encoded overtly or omitted. Therefore, discourse-pragmatics accounts might not be the only explanation for argument omission.

Besides, the pattern of higher informativeness score for overt argument than null argument was consistent in the object position for both of the verbs 'sik6' (eat) and 'zoek3' (wear), but it was not consistent in the subject position. Also, the informativeness scores for overt arguments in the subject position were much lower than those in the object position. Based on these two observations, it was proposed that the use of null subjects and null objects might not be determined by the same factors, or might be determined by the same factors but to a different extent. Discourse-pragmatics accounts might provide a weaker explanation for the use of null subjects than for the use of null objects. The use of null subjects might also be explained by other accounts such as processing accounts. Allen (2000) proposed that discourse-pragmatics accounts and processing accounts might be complementary to each other, suggesting that subjects were omitted by children due to processing limitations, but could still be used when the subject carried high information value. This might explain the subject-object asymmetry in the results of this study, in which the percentages of null arguments were higher in the subject position than in the object position but the subjects were not all omitted.

In conclusion, the results of this study suggested that Cantonese preschool children's use of argument omission differed for the two argument positions, and could be determined by how informative the referent is given by the informative features. The use of null argument was not affected by age factor in Cantonese preschool children aged beyond 2;06. In this study, it was found that discourse-pragmatics accounts could explain the use of

argument omission in Cantonese preschool children aged between 2;06 and 5;06, but they were not very strong and not likely to be the only explanation. Combination with other accounts, which needed to be further investigated, might explain the use of null arguments by Cantonese preschool children better.

In this study, certain limitations were encountered. In the analysis on informative features, informativeness of the features that were related to physical context could not be determined, since the data used in this study came from written transcripts of language samples in a database. No video record was available to examine the physical context during the conversation. In the study of Allen (2000), which was a null subject study on discourse-pragmatic perspective, there were eight informative features. This study employed only six of them, and excluded two features which were related to the physical context and could not be reflected from the transcripts.

It is suggested that some areas can be investigated in further research. Firstly, video recording can be used to include informativeness features which involved physical context to examine whether Cantonese children use these features to guide the use of null arguments as well. Secondly, verb specific pattern which was observed in this study can be further investigated in future studies by examining more classes of verbs and more verbs in each class. Thirdly, the pattern of use of argument by adults in the same conversation context as in this study could be examined in future studies, to investigate the effect of language input on the child's pattern of argument use through the comparison on adult's and child's pattern of argument use.

Acknowledgement

I would like to express my appreciation to Dr. Anita Wong from the Department of Speech and Hearing Sciences of the University of Hong Kong for her advice and supervision on the dissertation. I wish to thank my classmate Natalie Cheng Kit Man for participating in the inter-rater reliability check. I would also like to thank my family and friends for their support.

References

- Allen, S.E.M. (2000). A discourse-pragmatic explanation for argument representation in child Inuktitut. *Linguistics* 38, 3, 483-521
- Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry* 21(4), 491-504.
- Fletcher, P., Leung, C-S. S., Stokes, S., & Weizman, Z. (2000). *Cantonese Pre-School Language Development: A Guide*. HKU: Department of Speech and Hearing Science
- Lee, T. H-T. (2000). Finiteness and null arguments in child Cantonese. *The Tsinghua journal of Chinese studies, new series*, 30, 1, pp. 101-128
- MacWhinney, B., & Snow, C. (1990). The child language data exchange system: an update. *Journal of child language*, 17, 457-472
- Man, P. (1996) The realization and omission of subject and object in early Cantonese. In T. H-T. Lee et al. (Ed.). *The development of grammatical competence in Cantonese-speaking children*. Report of a project funded by RGC earmarked grant CUHK. pp 107-124
- Rothstein, S. D. (2004). *Structuring events: a study in the semantics of lexical aspect*. Malden, MA: Blackwell Pub.
- Sano, T., & Hyams, N (1994). Agreement, finiteness, and the development of null arguments. In M. Gonzalez (ed.), *Proceedings of the NELS 24*, 2, pp 543-558.
- Wang, Q., Lillo-Martin, D., Best, C.T., & Levitt, A. (1992). Null subject versus null object: some evidence from the acquisition of Chinese and English. *Language acquisition*, 2(3), 221 - 254

Appendix A: Mean number of utterances and MLU of participants in each age group

Age Group	Mean number of utterances	Mean MLU of each group
1 (Age: 2;06)	91.3	2.1966
2 (Age: 3;00)	128.5	3.1654
3 (Age: 3;06)	128.3	3.5667
4 (Age: 4;00)	155.8	3.5189
5 (Age: 4;06)	163.0	4.1116
6 (Age: 5;00)	164.7	4.2856
7 (Age: 5;06)	152.4	4.0199