

**The Development of *bei2* Dative Constructions  
in Early Child Cantonese**

CHAN Wing Shan Angel

A Thesis Submitted in Partial Fulfillment  
of the Requirements for the Degree of  
Master of Philosophy  
in  
Linguistics

© The Chinese University of Hong Kong  
November 2003



## Acknowledgements

I first gratefully acknowledge the graduate school of the Chinese University of Hong Kong (CUHK) for awarding me a postgraduate studentship in Linguistics at the Department of Modern Languages and Intercultural Studies of this institution. I whole-heartedly thank my thesis supervisor Professor Virginia Yip and her husband Dr. Stephen Matthews for their open-mindedness in encouraging me to pursue the usage-based approach to language acquisition and to start applying it to my native language, Cantonese- something exciting enough to motivate myself in pursuing it as a long-term commitment. When I was much paralyzed by thesis-induced stress in the earlier days, Prof. Yip kept on reminding me that the road to glory is full of twists and turns and that the process demands years of hardwork and dedication. She keeps on providing assurance and encouragement all along the way. Dr. Matthews is always ready and quick to give helpful and insightful input, and is always generous in sharing his time. I feel blessed to be a student of this couple. I am also deeply grateful to Prof. Gladys Tang and Prof. Gu Yang in my thesis committee for their invaluable comments on my study, and Prof. Jiang Ping and Prof. Foong Ha Yap for their teaching and support along the way.

Special thanks go to Professor Michael Tomasello for sending me his papers and book manuscript before the book was published, Professor Elena Lieven for answering my questions on the usage-based approach, Professors Adele Goldberg, Bill Croft, Laura Michaelis, Mirjam Fried for answering my questions and suggesting references on construction grammar. Sincere thanks to all those who either have pointed to me related references from different theoretical perspectives or

provided me with a copy of their related work, among them are Professors Thomas Lee, Tang Sze Wing, Gu Yang, Stephen Matthews, One-Soon Her, Ken Wexler, Eve Clark, Beth Levin, Peter Gordon, Bhuvana Narasimhan, John Newman, Foong Ha Yap, Doug Davidson, Alain Peyraube, Letitia Naigles, Kai Kiekhoefer, Thea Cameron-Faulkner and Nitya Sethuraman.

For data sharing and use, I acknowledge Professor Thomas Lee and his research team for establishing the Cantonese Child Language Corpus; and Professors Virginia Yip and Stephen Matthews for establishing the Hong Kong Bilingual Child Language Corpus- the two substantial child language corpora that have largely enriched the local and international research community. Special thanks to Drs. Law Sam Po and Leung Man Tak for letting me look at their Cantonese adult spoken corpus data before it was released, and to Cheung Sik Lee for her generosity in sharing the diary data of her bilingual son, Siu Bou. Sincere thanks to my speech therapist friends Suki Lam, Joyce Choi, Kim Yan and Frankie Lui for their help on the sharing of clinical data. Their senior colleagues Mr. Peter Au-Yeung, Assistant Director of the Heep Hong Society; Miss Yvonne Lai, Senior Speech Therapist of the Heep Hong Society; Mr. Liu, Principal of the B.M. Kotewall Memorial School; Mr. Yu, Principal of the CCC Mongkok Church Kai Oi School are also duly acknowledged for their consent in releasing the clinical data of their organization/school.

I gratefully acknowledge the valuable input from the audience of the following conferences. The bilingual subject Sophie's developmental data was first presented at the Summer Linguistics Workshop 2001 held at Yunnan University, China in May 2001. Parts of the findings on the adult input and the early developmental findings

of the monolingual Cantonese subjects were presented at the 2002 joint meeting of the Symposium on Research in Child Language Disorders (SRCLD) and the International Congress for the Study of Child Language (IASCL) held at the Madison-Wisconsin in July 2002. I thank the following people for their input, in particular, Professors Brian MacWhinney, Michael Tomasello, Thomas Lee, Yang Xiaolu, Dedre Gentner, Anna Theakston, Elena Lieven; and the financial support from the CUHK Postgraduate Student Grants for Overseas Academic Activities 2001-2002.

Heartfelt thanks go to my family, my father Chung Kit Chan, my mother Annette Lau, my aunt Chor San Wong, my sister Susanna and my brother Jason for showing me what unfailing love and care is. Thanks are also due to my brother for his technical support.

Thanks to all my CUHK colleagues Uta, Felix, Ee San, Scholastica, Queenie, Prudence, Sarah, Catherine Wong, Gordon, Candice, Wience, Jafi, Mikyong, Roy and Ming and undergraduate students whose company has made my three years of postgraduate education at CUHK more exciting and fun. I also gratefully thank all my friends and my brothers and sisters in cell and at church for their prayers and support throughout the years.

Last but most importantly, this thesis is dedicated to GOD who has been guiding me and supporting me through every step of the process. I have never been able to imagine that the painstaking process of doing this thesis is able to grant me so many blessings along the way- the Madison conference presentation, the DAAD scholarship to do my PhD with Professors Tomasello and Lieven at the Max Planck

Institute in Leipzig, Germany, and the various worthwhile openings in construction grammar, usage-based approach, experimentation and event structures/ cognition, the positive comments from examiners, and the possible publication opportunities. GOD has been treating me so well and has been so gracious to me. When I think of how much HE has given me, I kneel down before HIM and praise HIM. HE has been faithfully transforming my fear and confusion into more courage and understanding. I am thankful for all that HE has provided for me.

Glory to HIM,

Angel Chan

November 2003

# Contents

- Acknowledgements
- List of Abbreviations
- List of Tables and Figures
- Abstract

## **Chapter One Introduction**

- 1.0 Introduction
- 1.1 The Target Construction
  - 1.1.1 The Canonical [*bei2*-T-R] Double Object Form
  - 1.1.2 The Non-Canonical [*bei2*-R-T] Double Object Form
  - 1.1.3 The Non-Canonical [*bei2*-T-*bei2*-R] Serial Verb Form
  - 1.1.4 The Extended *bei2*-Dative
- 1.2 A Review of Cantonese Dative Constructions

## **Chapter Two Theoretical Background**

- 2.0 Introduction
- 2.1 The Markedness Hypothesis
- 2.2 The Iconicity Hypothesis
- 2.3 The Input Frequency Hypothesis
- 2.4 Relevance to Cantonese
  - 2.4.1 The Markedness Hypothesis: Empirical Predictions
  - 2.4.2 The Iconicity Hypothesis: Empirical Predictions
  - 2.4.3 The Input Frequency Hypothesis: Empirical Predictions
  - 2.4.4 An Interim Summary of Empirical Predictions
- 2.5 The Null Dative Marker Hypothesis
- 2.6 Conceptualization of End-State Knowledge: Construction Grammar
- 2.7 Introducing the Usage-Based Theory to Child Language Acquisition

## **Chapter Three The Input Properties Hypothesis and Adult Cantonese Input**

- 3.0 Introduction
- 3.1 Schematization
- 3.2 The Input Properties Hypothesis
- 3.3 Two Empirical Hypotheses on the Theme-Recipient Asymmetry
  - 3.3.1 Unexpressed Arguments: The Theme Versus The Recipient

- 3.3.2 Susceptibility to Displacement: The Theme Versus The Recipient
- 3.4 A Corpus Study Of Adult Input
  - 3.4.1 Methodology For Adult Input Analysis
  - 3.4.2 Corpus Findings
    - 3.4.2.1 The Missing Theme: *bei2*-Datives with Frequent Null Theme
    - 3.4.2.2 The Non-Canonical [*bei2*-R-T] Form
    - 3.4.2.3 The Non-Canonical [*bei2*-T-*bei2*-R] Form
    - 3.4.2.4 The Frequent [*bei2*-R] Sequence
- 3.5 Cantonese Adult Input Properties: Implications for Early Schematization
- 3.6 Chapter Summary

## **Chapter Four Methodology and Early Developmental Findings**

- 4.0 Introduction
- 4.1 Methodology
  - 4.1.1 Longitudinal Corpus Data
    - 4.1.1.1 Monolingual Child Data: The Hong Kong Cantonese Child Language Corpus (CANCORP)
    - 4.1.1.2 Cantonese-English Bilingual Child Data: The Hong Kong Bilingual Child Language Corpus
  - 4.1.2 Cantonese-English Bilingual Diary Data: Cheung (2002, p.c.)
  - 4.1.3 Clinical Child Data: Local Speech Therapists in Hong Kong
  - 4.1.4 Procedures for Data Analysis
- 4.2 Early Developmental Findings
  - 4.2.1 Non-Full *bei2*-Datives Before Full *bei2*-Datives
  - 4.2.2 The First Spontaneous Use of Full *bei2*-Datives
  - 4.2.3 All Full *bei2*-Datives Attested
  - 4.2.4 Early Preference for Non-Canonical Forms
    - 4.2.4.1 Possible Priming Effects
    - 4.2.4.2 Placement of Pauses
  - 4.2.5 The Late Acquisition of the Canonical [*bei2*-T-R] Form
  - 4.2.6 The Non-Target Use of *bei2*-Datives
    - 4.2.6.1 The Non-Target [*bei2*-R-T] Form
    - 4.2.6.2 The Non-Target [*bei2*-T-*bei2*-R] Form
- 4.3 Usage Patterns in Older Children
- 4.4 Summary of Major Findings



## **Chapter Five      Discussion of Findings**

- 5.0      Introduction
- 5.1      A Review of Established Empirical Predictions
- 5.2      The Markedness Hypothesis
- 5.3      The Iconicity Hypothesis
- 5.4      The Input Frequency Hypothesis
- 5.5      The Input Properties Hypothesis
- 5.6      Markedness From the UG perspective
- 5.7      The Early Preference for Non-Canonical Forms: A Functional Perspective
- 5.8      The Source of the Early Non-Canonical *bei2*-datives: A Usage-Based Perspective
  - 5.8.1    The Early [*bei2*-R-T] Form
    - 5.8.1.1    Against Learning Directly From The Adult Speech Models
    - 5.8.1.2    Against Generating Directly From The [*bei2*-R-T] Verb Specific Schema
    - 5.8.1.3    Against Overgeneralizing the Abstract [V-R-T] Schema
  - 5.8.2    The Early [*bei2*-T-*bei2*-R] Form
    - 5.8.2.1    Against Learning Directly From The Adult Speech Models
    - 5.8.2.2    On Overgeneralizing The [V-T-*bei2*-R] Schema
- 5.9      Remaining Questions
- 5.10     Chapter Summary

## **Chapter Six      Conclusions and Further Research**

- 6.0      Introduction
- 6.1      Principal Conclusions
- 6.2      Contributions
- 6.3      Suggestions for Further Research
  - 6.3.1    Elicited Production Studies
  - 6.3.2    Comprehension Studies
  - 6.3.3    Cross-Linguistic Investigations

Appendices

References

## List of Abbreviations

CL	Noun Classifier
DET	Determiner
DO	Direct Object
DOC	Double Object Dative Construction
IO	Indirect Object
MM	Modifier Marker
NUM	Numeral
PDC	Prepositional Dative Construction
PERF	Perfective Aspect Marker
PRT	Particle
PVT	Postverbal Particle
T	Theme Argument
TOP	Topic Marker
R	Recipient Argument
Vf	Function Verb
3sg	Third Person Singular

# List of Tables and Figures

## Tables

- Table 1.1 Tang (1998)'s classification of Cantonese dative verbs
- Table 1.2 Three dative constructions and their associated verb classes in Cantonese
- Table 2.1 Predicted word order preferences in Cho et al. (2002)
- Table 2.2 A summary of iconic motivation for the canonical and the non-canonical *bei2* forms
- Table 2.3 An interim summary of empirical predictions
- Table 2.4 Four basic sets of acquisition processes in Tomasello (2003: 295)
- Table 3.1 Utterances containing *bei2* in CANCELP
- Table 4.1 Age span and the number of data files created in CANCELP
- Table 4.2 Background information of the eight monolingual Cantonese children in CANCELP
- Table 4.3 Age span and the number of Cantonese and English data files created in the bilingual corpus
- Table 4.4 Background information of the three bilingual Cantonese-English children
- Table 4.5 Background information of the seven children with language impairments
- Table 4.6 The first use of [*bei2*-T] and [*bei2*-R] non-full *bei2*-datives attested in CANCELP
- Table 4.7 All the uses of [*bei2*-T] and [*bei2*-R] non-full *bei2*-datives attested in CANCELP before children's first spontaneous use of a full

*bei2*-dative

- Table 4.8 Monolingual children's first spontaneous use of a full dative with the verb *bei2* in CANCECORP
- Table 4.9 The placement of the theme argument in early *bei2*-full datives in CANCECORP
- Table 4.10 Full *bei2*-datives attested in CANCECORP
- Table 4.11 The early developmental non-target [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms in different acquisition contexts
- Table 4.12 Prevalence of the non-target [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms in early child Cantonese
- Table 4.13 The elder siblings' use of the Cantonese *bei2*-datives as attested in their young siblings' corpus data

## Figures

- Figure 2.1 The internal symbolic structure of a construction
- Figure 3.1 Adult use of Cantonese *bei2*-datives in CANCECORP in child directed speech

## Abstract

This thesis investigates the development of dative constructions with the verb *bei2* ‘give’ in child Cantonese. The study uses a monolingual Cantonese child language corpus (CANCORP, Lee et al., 1996) to trace the early longitudinal development of the Cantonese *bei2*-datives among eight children from the age of 1;05.22 to 3;04.14. These empirical developmental findings are supplemented by naturalistic Cantonese-English bilingual corpus data (Yip, Matthews and Huang, 2001), bilingual diary data from Cheung (2002, p.c.) as well as clinical data from speech therapists in Hong Kong. The following hypotheses relevant to the acquisition of these constructions are considered: the markedness hypothesis (O’Grady, 2000), the iconicity hypothesis (Cho et al., 2002), the input frequency hypothesis (O’Grady, 2000), and the input properties hypothesis in relation to schematization in the usage-based theory of language acquisition (Tomasello, 2003). The findings show that the canonical [*bei2*-T(heme)-R(ecipient)] double object form is acquired late, while the non-canonical [*bei2*-R-T] double object form and [*bei2*-T-*bei2*-R] serial verb form emerge early and are used in a non-target manner from an adult language perspective, despite the fact that the canonical double object [*bei2*-T-R] form is motivated by iconicity and is more frequent in the adult input than the non-canonical forms. However, analysis of the adult input also shows that the canonical [*bei2*-T-R] double object form is used alongside other *bei2*-datives with unexpressed or displaced theme, and other related serial verb dative constructions which instantiate the [*bei2*-R] sequence. This suggests that early schematization of the canonical [*bei2*-T-R] form might be difficult in such a linguistic environment from a usage-based perspective. Moreover, the [V-T-R] double object form is

cross-linguistically marked, while the [V-R-T] double object form and the [V-T-dative marker-R] form are cross-linguistically unmarked (Michaelis and Haspelmath, 2003). The findings on the late acquisition of the canonical [V-T-R] double object form support the input properties hypothesis, because the Cantonese language specific input properties are structured in a way that makes schematization of the canonical [V-T-R] double object form far from straightforward. The findings on the early emergence of the non-canonical forms prior to the canonical form also support the markedness hypothesis, since the non-canonical forms are cross-linguistically less marked than the canonical form. The present adult input findings raise questions for the usage-based approach to language acquisition (Tomasello 2003) concerning how children could schematize the full-fledged canonical [V-T-R] structure from the input given the prevalence of null and displaced arguments. The parallels between cross-linguistic distribution and developmental preferences also call for more cross-linguistic acquisition studies of the marked [V-T-R] double object construction.

## 摘要

本論文研究兒童粵語中動詞「畀」的與格結構的發展。本研究採用一個單一語兒童的粵語語料庫 (CANCORP, Lee et al., 1996), 追溯八名操粵語的兒童由 1;05.22 至 3;04.14 歲「畀」與格結構的早期縱向發展 (longitudinal development)。本論文還會用一個香港粵英雙語發展兒童的語料庫 (Yip, Matthews and Huang, 2001)、粵英雙語發展兒童的日記資料 (張氏, 2002, p.c.), 以及由本港言語治療師所提供的臨床資料, 作為以實驗所得的兒童發展研究數據。與此與格結構習得有密切關係的假說: 象似性 (iconicity) 假說 (Cho et al., 2002)、語料輸入頻率假說 (O'Grady, 2000), 與以應用為本的語言習得理論 (usage-based theory) 中知識結構 (schematization) 有關的語料輸入特徵假說 (Tomasello, 2003), 以及標記限制假說 (O'Grady, 2000), 均在本文討論之列。研究結果顯示, 雖然慣常 (canonical) 的雙賓語 [畀-客體-接受者] ([*bei2*-Theme-Recipient]) 句式在成人語料輸入中比非慣常句式的出現頻率較高, 並且得到象似性假說的支持, 但是, 兒童於發展期較後階段才可以學習得到慣常的 [畀-客體-接受者] 雙賓語句式, 而非慣常 (non-canonical) 的 [畀-接受者-客體] ([*bei2*-Recipient-Theme]) 雙賓語句式及 [畀-客體-畀-接受者] 連動(動)詞句式, 則從成人語言的角度看, 早已以非目標 (non-target) 方式出現。然而, 就成人語料輸入的數據分析顯示,

成人用慣常的 [界-客體-接受者] 雙賓語句式的同時，均會使用其他不表達客體或客體被放在其他位置的「界」與格結構，以及其他包含 [界-接受者] ([*bei2*- Recipient]) 序列的相關連動與格結構。從應用為本理論中，兒童知識結構形成的角度來看，在此等語言環境下，兒童要在早段發展期時，形成慣常的 [界-客體-接受者] 知識結構應該比較難。此外，從跨語言比較的角度來看 (Michaelis and Haspelmath, 2003)，[動詞-客體-接受者] ([Verb-Theme-Recipient]) 雙賓語句式較 [動詞-接受者-客體] ([Verb-Recipient-Theme]) 雙賓語句式及 [動詞-客體-與格標記-接受者] ([Verb-Theme-Dative Marker-Recipient]) 句式更有標記限制。是次兒童在較後段發展期才學習到典型 [界-客體-接受者] 雙賓語形式的研究結果，可支持輸入特徵假說，因為粵語的語料輸入特徵有機會使其知識結構形成複雜化；而非慣常句式早於慣常句式出現的另一項研究結果，可支持標記限制假說，因為從跨語言比較的角度來看，非慣常句式比慣常句式更普遍，其標記性程度較低。是次成人粵語語料輸入的研究結果，也帶出兒童如何能在此等語言環境下形成慣常的 [界-客體-接受者] 知識結構。而是次兒童早期粵語發展的研究結果，也顯示語言類型學與兒童語言發展存在對應關係 (typological-developmental parallels)，應就有標記限制的 [動詞-客體-接受者] 雙賓語句式習得進行更多跨語言的研究。



# Chapter One. Introduction

## 1.0 Introduction

This thesis investigates the development of dative constructions with the verb *bei2* ‘give’ in early child Cantonese, focusing on the canonical and non-canonical forms (see section 1.1). Few if any facts are currently known with regard to the acquisition of the Cantonese *bei2* dative construction. One goal of this thesis is to document the developmental facts of this dative construction in child Cantonese. The thesis aims to *relate* the acquisition of these language particular facts to the following hypotheses already formulated in the existing literature for investigating the acquisition of dative constructions: the markedness hypothesis (O’Grady, 2000), the iconicity hypothesis (O’Grady, 2000; Cho et al., 2002), and the input frequency hypothesis (O’Grady, 2000; Campbell and Tomasello, 2001). As we shall see, these existing hypotheses yield different predictions for the acquisition of the canonical and non-canonical forms of the Cantonese *bei2* dative construction. In addition, based on the language specific characteristics of Cantonese, I shall investigate one more hypothesis, called the input properties hypothesis here, in relation to schematization in the usage-based theory of language acquisition (Tomasello, 2003).

## 1.1 The Target Construction

### 1.1.1 The Canonical [*bei2*-T-R] Double Object Form

In this thesis, I limit the target construction under the present study to the ditransitive constructions with the verb *bei2* ‘give’ which encode linguistically only the transfer event (transfer of objects or information) and its participants. (1) shows how native Cantonese speaking adults align the theme (T) and the recipient (R) arguments of a

transfer event with the verb *bei2* ‘give’ under pragmatically neutral situations.<sup>1</sup>

- (1) Ngo5 bei2 jat1 bun2 syu1 lei5 laa1  
I give one CL book you PRT  
‘I give a book to you.’

The theme (*jat1 bun2 syu1* ‘one book’) precedes the recipient (*lei5* ‘you’) as (1) shows. Following Tang (1998), I shall regard examples like (1) as instances of a double object dative construction, because both the theme and the recipient are coded as objects (direct object and indirect object respectively) and both objects are zero-marked. Note that this theme-recipient postverbal ordering of objects differs from the recipient-theme ordering of the double object construction in English and Mandarin.

### 1.1.2 The Non-Canonical [*bei2*-R-T] Double Object Form

Under the following marked situations, adults may place the theme (T) *after* the recipient (R), resulting in the surface form [*bei2*-R-T].

- i) when the theme argumental NP is especially long or heavy, Cantonese adult speakers may postpone the theme after the recipient. Consider the following example (2) with a long theme argument of seven syllables *jat1 bun2 hou2 jau5jung6 ge3 syu1* ‘a very useful book’. In this example (2), the long theme is placed *after* the recipient *keoi5* ‘him/her’ rather than before it as (1)

---

<sup>1</sup> In this thesis, Cantonese examples are transcribed orthographically in the JyutPing romanization system developed by the Linguistic Society of Hong Kong in 1993 (see LSHK, 2002). Tones are marked numerically (1: high level, 2: high rising, 3: mid level, 4: low falling, 5: low rising and 6: low level). See the list of abbreviations for abbreviations used in the glosses.

shows. Matthews and Yip (1994: 137, 407) pointed out that cases like (2) is comparable to heavy noun phrase shift in English, where the recipient-theme<sub>(long)</sub> order is used to avoid the clumsiness of the theme<sub>(long)</sub>-recipient order in (3).

(2) Ngo5 bei2 zo2 keoi5 jat1 bun2 hou2 jau5jung6 ge3 syu1  
 I give PERF 3sg one CL very useful MM book  
 'I have given her/him a very useful book.'

(3) Ngo5 bei2 zo2 jat1 bun2 hou2 jau5jung6 ge3 syu1 keoi5  
 I give PERF one CL very useful MM book 3sg  
 'I have given a very useful book to her/him.'

- ii) when the theme is used contrastively with the following clause, probably with an emphatic stress on the theme to indicate contrast or pragmatic prominence to achieve end-focus (see Tang, 1998), adult Cantonese speakers may also place the theme after the recipient. Consider (4). When the theme *jat1 bun2 syu1* 'one book' is contrasted with another theme *jat1 zi1 bat1* 'one pen' in the following clause, the theme *jat1 bun2 syu1* 'one book' is placed after the recipient *keoi5* 'her/him'. Consider another example (5). When the theme *mei5gam1* 'US dollars' is contrasted with another theme *jat6jyun4* 'Japanese yen' in the following clause, the theme *mei5gam1* 'US dollars' is placed after the recipient *lei5* 'you'.

(4) Ngo5 bei2 zo2 keoi5 jat1 bun2 syu1, m4hai6 jat1 zi1 bat1  
 I give PERF 3sg one CL book not one CL pen  
 'I have given her/him one book, not one pen.'

(5) Ngo5 bei2 lei5 mei5gam1, lei5 bei2 ngo5 jat6jyun4 aal  
 I give you US dollars, you give me Japanese yen PRT  
 'I give you US dollars, you give me Japanese yen.'

### 1.1.3 The Non-Canonical [*bei2-T-bei2-R*] Serial Verb Form

Under certain conditions, adult Cantonese speakers may use the verb *bei2* ‘give’ in the [*bei2-T-bei2-R*] serial verb form.

- i) when the theme argumental NP is long, Cantonese adult speakers may ‘mark’ the recipient NP using a second *bei2* ‘give’ as a dative marker, using a [*bei2-T-bei2-R*] serial verb form in describing a transfer event.<sup>2</sup> In this case, the second verb *bei2* ‘give’ is used as a goal-marking dative marker introducing the entity as the intended recipient. (6) shows an illustrative example actually attested in an adult spoken Cantonese corpus (Law, Leung and Fung, 2002: utterance no. 202).

(6) Zau6 bei2 hou2do1 jau5ceoi3 ge3 min6hung2 bei2 keoi5  
Then give many funny MM face give 3sg  
‘Then give many funny faces to him/her.’

- ii) when adult speakers want to emphasize an NP as playing the recipient role, they may also use a [*bei2-T-bei2-R*] serial verb form. Similar to condition i) mentioned above, the speakers ‘mark’ the recipient NP with a *bei2* dative marker (with even added emphasis as represented by *bei2* below) using a serial verb [*bei2-T-bei2-R*] form, even when the theme NP is not particularly long. (7) shows an example of adult usage from a naturalistic Cantonese corpus (Lee et al., 1996).

---

<sup>2</sup> Tang (1998) also observed that when the distance between the two *bei2s* increases (by increasing the length of the intervening theme NP), the acceptability of [*bei2-T-bei2-R*] increases; otherwise, the closer the two *bei2s*, the greater the unacceptability of [*bei2-T-bei2-R*] and the greater the tendency to use the canonical form [*bei2-T-R*].

(7) From MHZ corpus, File "11023.cha": line 3848.

\*INV: Bei2 bo1bo1 bei2 bin1go3 aa3 ?  
Give ball give who PRT  
'Give ball to whom?'

Although so far we have seen that the [*bei2*-T-R], [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms are possible in adult Cantonese, I shall consider the [*bei2*-T-R] double object form as the canonical form of the *bei2* dative, and the [*bei2*-T-*bei2*-R] and [*bei2*-R-T] forms as the non-canonical forms reserved for marked contexts in this thesis. The reasons are as follows. For the verb *bei2* to occur in the [V-R-T] double object form as in (2), (4) and (5) and the [V-T-*bei2*-R] serial verb form as in (6)-(7), factors such as focus and the length of the theme object come into play. In neutral contexts, the use of the [*bei2*-R-T] form and the serial verb [*bei2*-T-*bei2*-R] form is non-target-like from the adult language perspective, instead the form [V-T-R] is used as a default in pragmatically neutral situations with the verb 'give'. The use of [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms in 'marked' contexts suggests that additional mechanisms are involved in the production of these structures. (8a)-(8c) illustrate the relative grammaticality judgments of these structures for the Cantonese *bei2*-dative.

- (8) a. Ngo5 bei2 zo2 cin2 keoi5 [bei2-T-R] (canonical)  
 I give PERF money 3sg  
 'I have given money to her/him.'
- b. ??Ngo5 bei2 zo2 keoi5 cin2 [bei2-R-T] (non-canonical)  
 I give PERF 3sg money  
 'I have given her/him money.'
- c. ??Ngo5 bei2 zo2 cin2 bei2 keoi5 [bei2-T-bei2-R] (non-canonical)  
 I give PERF money give 3sg  
 'I have given money to her/him.'

#### 1.1.4 The Extended *Bei2*-Dative

In Cantonese (which is a serializing language) and other languages, it is possible to elaborate on the basic 'give' clause with an additional verbal predicate, illustrated by the verb *tai2* 'read' in (9) (see Newman, 1999:124 for more cross-linguistic examples).

- (9) Ngo5 bei2 jat1 bun2 syu1 lei5 tai2 laa1  
 I give one CL book you read PRT  
 'I give a book to you to read.'

The verb *bei2* 'give' can be understood literally, but it also at the same time expresses the means by which a further act of reading can take place with the recipient (Newman, 1999: 124). Here I term examples such as (9) as instances of the *extended bei2* dative construction, where the transfer event expressed by the first verb *bei2* 'give' is *extended*, leading to and enabling the act of reading (Newman, 1999: 124). The extended *bei2*-dative is beyond the present scope of the investigation.

## 1.2 Review of Cantonese Dative Constructions

There are three major dative constructions involving *bei2* and other verbs for expressing the theme and the recipient of a transfer event in Cantonese: the double object [V-T(heme)-R(ecipient)] construction, the [V-R-T] double object construction, and the serial verb [V-T-*bei2*-R] construction. (10) shows an example of each construction.

(10) Examples of the three dative constructions in Cantonese

a. ‘Inverted’ Double Object Construction (IDOC) [V-T-R]:

Ngo5    bei2    bun2    syu1    lei5  
I        give    CL        book    you  
‘I give a book to you.’

b. Serial Verb Construction [V-T-*bei2*-R]:

Ngo5    maai5    bun2    syu1    bei2    lei5  
I        buy     CL        book    give    you  
‘I buy a book for you.’<sup>3</sup>

c. Double Object Construction (DOC) [V-R-T]:

Ngo5    gaau3    lei5    zung1man2  
I        teach    you    Chinese  
‘I teach you Chinese.’

(10a) presents an example of what Tang (1998) called the ‘Inverted’ Double Object construction (IDOC), implying as if the basic form should be the [V-R-T] double object construction, although he proposed that this construction should be

---

<sup>3</sup> The English translations are intended to capture the closest meaning counterparts possible in English for the Cantonese expressions. Here there is a difference between English and Cantonese. For the English prepositional *for*-dative “I buy a book for you”, “you” need not be the intended recipient, it can be the beneficiary (you are too busy to buy a book so I do it on your behalf), some linguists therefore treat the NP2 in the English prepositional dative as having the thematic role RECIPIENT (or GOAL) or BENEFICIARY. In Cantonese, however, *lei5* ‘you’ must be the intended recipient. Another structure is used for the beneficiary reading *ngo5 bong1 lei5 maai5 bun2 syu1* ‘I help you buy a book’.

structurally related to the serial verb construction in (10b) with a null dative marker, i.e. [V-T-(*bei2*)-R] (see section 1.4.4 on the null dative marker hypothesis). It is regarded as a double object construction because both the theme and the recipient are coded as objects and are zero-marked. Here the theme is the first postverbal argument, while the recipient is the second postverbal argument. According to Tang (1998), other Cantonese ‘give’ verbs such as *sung3* ‘give (as a present)’ (see table 1.1 for more examples of the Cantonese ‘give’ verbs classified in Tang (1998)) can occur in this [V-T-R] double object form, although its grammaticality/acceptability judgments vary among native speakers of Cantonese. In contemporary Cantonese, however, it appears to the author that most of these other Cantonese ‘give’ verbs used in the [V-T-R] form are vanishingly rare. The author, a native speaker of Cantonese brought up in Hong Kong, for instance, never encounters people using [*sung3*-T-R] expressions in Hong Kong nowadays. Instead, these other Cantonese ‘give’ verbs are largely used in the [V-T-*bei2*-R] serial verb construction. Moreover, other than *bei2* ‘give’, no other Cantonese ‘give’ verbs appear in the [V-T-R] form in the adult Cantonese utterances of the monolingual and bilingual corpora used in the present study (see sections 4.1.1.1 and 4.1.1.2 for detailed information of the corpora). Positing a [*bei2*-T-R] verb specific construction might therefore be more appropriate in capturing contemporary adult Cantonese usage. Hence I regard the [V-T-R] double object construction as lexically restricted to the verb *bei2* ‘give’ in contemporary Cantonese. Moreover, as mentioned in section 1.1.1, this [V-T-R] double object form is also the canonical form of the Cantonese *bei2*-dative, the target construction under the present study.



(10b) shows an example of what I shall call the serial verb construction. This construction receives various names in the current literature, mainly because people differ in their analysis of *bei2*. Tang (1998) treated the *bei2* dative marker before the recipient object as a goal-marking preposition and thereby termed this type of dative construction the prepositional dative construction (PDC). Xu and Peyraube (1997) also treated the *bei2* dative marker before the recipient as a preposition but termed the same construction as the oblique dative. Liu (2001) and Matthews and Leung (2002), on the other hand, analyzed the construction as a Serial Verb Construction [Verb-Object-Verb-Object], treating the *bei2* before the recipient as a verb. Similar to Liu (2001) and Matthews and Leung (2002), this thesis adopts the serial verb analysis, henceforth I term this construction the serial verb construction, with the verb *bei2* 'give' functioning as a dative marker before the recipient.

(10c) shows an example of what is commonly called the double object construction (DOC). Goldberg (1995) called it the ditransitive construction. The construction has the surface form of [V-R-T], where the recipient precedes the theme.

Tang (1998)'s work is by far the most comprehensive and updated work in describing these three Cantonese dative constructions and their grammaticality with different verbs. He suggests five semantically defined verb classes which can occur in the three dative constructions described in section 1.4.1 above. Table 1.1 below lists all the examples of verbs in these five verb classes mentioned in Tang (1998).

Table 1.1. Tang (1998)'s classification of Cantonese dative verbs

Verb Classes	
The 'give' verbs	<i>bei2</i> 'to give', <i>sing2</i> 'to give (as a present)', <i>sung3</i> 'to give (a present)', <i>zoeng2</i> 'to award'
The 'send' verbs	<i>bun1</i> 'to move', <i>daai3</i> 'to bring', <i>dai6</i> 'to hand to', <i>deng3</i> 'to pelt', <i>gaau1</i> 'to deliver', <i>gaap3</i> 'to lift food with chopsticks', <i>gei3</i> 'to send', <i>lau4</i> 'to reserve', <i>ling1</i> 'to carry with hand', <i>lo2</i> 'to bring', <i>maai6</i> 'to sell', <i>paai3</i> 'to deliver', <i>tek3</i> 'to kick', <i>wui6</i> 'to remit'
The 'fry' verbs (verbs of creation)	<i>caau2</i> 'to fry', <i>jing2</i> 'to photocopy', <i>pai1</i> 'to cut', <i>sai2</i> 'to wash', <i>tong3</i> 'to iron', <i>waak6</i> 'to draw', <i>zam1</i> 'to pour', <i>zik1</i> 'to knit', <i>zing2</i> 'to make', <i>zok3</i> 'to compose', <i>zyu2</i> 'to cook'
The 'pluck' verbs (mainly verbs of obtaining and removing)	<i>coeng2</i> 'to snatch', <i>gaan2</i> 'to choose', <i>maai5</i> 'to buy', <i>ling1</i> 'to take', <i>lo2</i> 'to get', <i>tau1</i> 'to steal', <i>zaak6</i> 'to pluck'
The 'teach' verbs	<i>ceng2gaau3</i> 'to inquire', <i>gaau3</i> 'to teach', <i>haau2</i> 'to test', <i>kau4</i> 'to request', and <i>man6</i> 'to ask'

Table 1.2 shows the three dative constructions and their associated verb classes in Cantonese.

**Table 1.2. Three dative constructions and their associated verb classes in Cantonese (adapted from Tang (1998:40, Table 9)'s classification)**

	'give' verbs	'send', fry', 'pluck' verbs <sup>4</sup>	'teach' verbs
[V-T-R] <sup>5</sup>	OK (but lexically specific to <i>bei2</i> in contemporary Cantonese)	*	*
[V-T- <i>bei2</i> -R]	OK	OK	*
[V-R-T]	in marked contexts <sup>6</sup>	*	OK

From Tables 1.1 and 1.2 above, each verb class seems to be associated with a distinct type of construction, for instance, the 'teach' verbs can only occur in the [V-R-T] double object construction; the 'send', 'fry' and 'pluck' verbs can only occur in the [V-T-*bei2*-R] serial verb construction. Consequently, there is *no* dative alternation between the serial verb construction and its double object counterpart in Cantonese comparable to that in English (the alternation between the *to-/for*-dative and the double object dative). Unlike English, there is not a subset of verbs in Cantonese

<sup>4</sup> Tang (1998) marked the box where the 'send', 'fry' and 'pluck' verbs occur in the [V-T-R] construction using \*/? in his original work. An illustrative example given by Tang (1998: 38) is repeated in (i) below.

- (i)     \*/? Ling4ling2 caau2 zo2   jat1 dip6 coi3       ngo5  
           Lingling   fry   PERF one CL   vegetable   me  
           (= Tang 1998:38, example 5b)

The author, as a native speaker of Hong Kong Cantonese, considers the \*/? examples such (i) above given in Tang (1998) ungrammatical. Moreover, I have not encountered such usage in contemporary Cantonese, hence this box is marked with \* here.

<sup>5</sup> For the ease of exposition, I represent the surface form of the constructions in such a manner throughout this thesis. However, it must also be emphasized that it is not being claimed in constructional grammar that syntactic structures are merely a linear string of elements, but there exists constituency (Goldberg, March 2003, p.c.; see Fillmore and Kay et al. (1988) which preserved constituency in the usual sense (Croft, March 2003, p.c.); but see also Croft, 2001: chapter five which specifically discussed constituency and the problems with arguments for constituency) and hierarchical structure (Croft, 2001: chapter six; Croft, March 2003, p.c.). This aspect of representational issue in adult grammar will not be handled in this thesis.

<sup>6</sup> Tang (1998) in his original work marked this box using the question mark sign ?, because the use of [V-R-T] with the 'give' verbs are limited to marked contexts when the theme is long or stressed.

that can occur in both the serial verb [V-T-*bei2*-R] schema and the double object [V-R-T] construction involving the same thematic roles <agent, theme, recipient/beneficiary>.<sup>7</sup> Moreover, this thesis adopts a constructional view of grammar (see section 2.6). In the case of dative constructions, the constructions are not related by transformational relationships (see for example Larson, 1988) or by lexico-semantic rules (see Pinker, 1989), but are regarded as independent constructions related by constructional links as proposed by Goldberg (1995). Hence, the ensuing learnability issues relevant to the dative alternation in English are tenuous in the case of Cantonese and are beyond the current scope of investigation.

The verb *bei2* is unique in Cantonese because it is the only verb among the other verbs in contemporary Cantonese that i) occurs in *all* these three major dative constructions (see section 1.1), and ii) has the [V-T-R] double object form as its canonical form, a structure which is rarely attested in the languages of the world (see section 2.4.1) and is specific to *bei2* ‘give’ in contemporary Cantonese (see table 1.2).

---

<sup>7</sup> The Cantonese ‘send’ verbs and the ‘fry’ verbs (see table 1.1 for examples of verbs) can occur in the serial verb construction and the double object construction, though the thematic roles involved are different: <agent, theme, recipient> with the dative marker construction and <agent, source, theme> with the double object construction. When these verbs occur in a double object configuration, the resultant utterances are associated with a deprivational meaning. This [Verb-Source-Theme] construction might be termed the Cantonese deprivational double object construction, being associated with the central sense of ‘X causes Y to lose Z.’. This construction in Cantonese is outside the scope of this thesis, but see Tang (1998) for the Cantonese data, and Chung and Gordon (1998) for similar Mandarin data.

## Chapter Two. Theoretical Background

### 2.0 Introduction

This chapter provides the theoretical background relevant to this study. The markedness hypothesis, the iconicity hypothesis and the input frequency hypothesis used in the existing acquisition research on datives in other languages are relevant candidates for predicting the ease of acquisition of the canonical and non-canonical forms of the Cantonese *bei2*-dative construction. Sections 2.1 to 2.3 describe the markedness hypothesis, the iconicity hypothesis, and the input frequency hypothesis in the existing acquisition literature. Section 2.4 investigates how these hypotheses apply to the acquisition of the Cantonese *bei2* dative construction and sets up the empirical predictions accordingly. Section 2.5 discusses the null dative marker hypothesis which is also relevant to the relationship between the canonical [*bei2*-T-R] form and the non-canonical [*bei2*-T-*bei2*-R] form. Since this study adopts the usage-based emergentist view to language acquisition proposed in Tomasello (2003) under a constructional view of grammar (see for example Fillmore et al., 1988; Goldberg, 1995; Croft, 2001), section 2.6 offers a general description of a constructional view of grammar. Section 2.7 gives a brief introduction to the usage-based theory, focusing on the level of abstractness of early linguistic representations.

### 2.1 The Markedness Hypothesis: O'Grady (2000)

The markedness hypothesis is one hypothesis considered by O'Grady (2000) who worked on the acquisition of double object dative constructions (DOC) and prepositional dative constructions (PDC) in developmental English using a

comprehension study. The hypothesis is cited below:

“Children find it easier to acquire unmarked structures- that is, structures that are for one reason or another more common in the world’s languages” (O’Grady 2000: 4)

Cross-linguistic frequency is one of the standard markedness criteria (see Croft, 1990: chapter four). Most linguists believe that cross-linguistic similarities do not come about arbitrarily, although they differ in how to make use of cross-linguistic frequencies in theorizing. Functionally-oriented linguists would interpret the typological patterns as reflecting the fact that some patterns are more functionally motivated than the others. Some child language researchers take the cross-linguistic generalizations as reflections of universal principles that underlie child language development, perhaps in the form of universal human cognitive tendencies:

“While the causes of typological patterns are often unclear..., it seems likely that at least some of them reflect deep-seated properties of human perceptual, cognitive, and communicative activity, which children might also be expected to share.” (Bowerman, 1993: 7).

If the typological and developmental domains share some important common bases, it is possible that the typological patterns noted in the dative constructions find parallels in the acquisition of this domain of grammar. O’Grady (2000) compared the DOC and the PDC along the parameter of cross-linguistic markedness, pointing out that DOC is more marked than PDC across languages because languages like French and Japanese do not have DOC, and in English because only a subset of verbs that occur in PDC can occur in DOC. He (2000: 4) therefore hypothesized that “whatever factors contribute to the markedness of double object constructions across

languages and within languages also impede their acquisition”. DOCs turn out to be more difficult than the PDCs in early English in comprehension studies. Consequently, the markedness hypothesis is *one* among other plausible hypotheses (see also the iconicity hypothesis (section 2.2) and the explicitness hypothesis (section 6.3.2)) in accounting for why comprehending DOCs is difficult in child English.

## 2.2 The Iconicity Hypothesis: O’Grady (2000), Cho et al. (2002)

The iconicity hypothesis<sup>1</sup> is another hypothesis considered by O’Grady (2000) in connection with the acquisition of English dative constructions. The hypothesis is cited below:

“Children prefer iconic structures- that is, structures whose word order reflects the manner in which the corresponding event unfolds.” (O’Grady, 2000: 5)

A sentence is regarded as *iconic* when its word order matches the manner in which the corresponding situation unfolds. O’Grady (2000: 4) described the transfer event as follows: ‘in the case of verbs of transfer like *give, show, tell*, we have an event that originates with an agent who then acts upon a theme and transfers it to a goal.’ (1) shows his description of the structure of a transfer event.

- (1) The structure of a transfer event in O’Grady (2000: 4)  
agent- > theme- > goal

---

<sup>1</sup> The hypothesis was named as ‘the matching hypothesis’ in O’Grady (2000). It was renamed as ‘the iconicity hypothesis’ in Cho et al. (2002), a study conducted by O’Grady and his colleagues. The central idea of the matching hypothesis and the iconicity hypothesis is essentially the same, hence, I adopt the more recent name ‘iconicity hypothesis’ in this thesis.

According to the description in (1), the English prepositional dative construction (PDC) would be regarded as iconic, because the linear sequence of the agent, the theme and the goal in speech is identical (isomorphic) to how the transfer event unfolds sequentially in (1). The isomorphism between the word order of a preposition dative and the corresponding transfer event is spelt out as follows.

“So when I say ‘Mary gave the book to John’, I am describing an event in which the agent (Mary) picks up a book (the theme) and transfers it to John (the goal).” (O’Grady, 2000: 4)

In the case of the English double object dative construction (DOC), it would be regarded as non-iconic, because the recipient precedes the theme. O’Grady (2000) hypothesized that one reason the DOC is difficult for children in comprehension studies- among other possible factors such as its markedness (refer back to 2.1)- might be that the word order of the double object dative does not match with, i.e. is not iconic to, how the transfer event naturally unfolds.

Cho et al. (2002) further pointed out the possible role of iconicity in early developmental word order preferences. They investigated the word order preferences for direct and indirect objects in children acquiring Korean using two comprehension act-out studies: the first investigating the dative pattern, the second the instrumental pattern. In their first study concerning the dative pattern, they compared children’s word order preference of the accusative-dative and the dative-accusative orders in comprehension act-out tasks. Results showed a statistically significant preference for the accusative-dative order over the dative-accusative order, in the sense that the children performed far better on the accusative-dative order than the dative-accusative order, whose interpretation they



tended to reverse to accusative-dative order. (2) shows some sample test sentences in Cho et al. (2002)'s study.<sup>2</sup>

(2) a. Animate direct object; animate indirect object

Dative-accusative order:

So-hanthey kom-ul mile-cwu-llyay-yo?

cow-Dat bear-Ac push-Ben-Fut-SentEnder

'Will you push the bear to the cow?' (=3a) in Cho et al., 2002)

Accusative-dative order:

Kom-ul so-hanthey mile-cwu-llyay-yo?

bear-Ac cow-Dat push-Ben-Fut-SentEnder

'Will you push the bear to the cow?' (=3a) in Cho et al., 2002)

b. Inanimate direct object; inanimate indirect object

Dative-accusative order:

Moca-ey sonswuken-ul tencye-cwu-llyay-yo?

cap-Dat handkerchief-Ac throw-Ben-Fut-SentEnder

'Will you throw the handkerchief at the cap?' (=3b) in Cho et al., 2002)

Accusative-dative order:

Sonswuken-ul moca-ey tencye-cwu-llyay-yo?

handkerchief-Ac cap-Dat throw-Ben-Fut-SentEnder

'Will you throw the handkerchief at the cap?' (=3b) in Cho et al., 2002)

Cho et al. (2002) considered two hypotheses that are compatible with the preference of the accusative-dative order over the dative-accusative order in young Korean speaking children.<sup>3</sup> The two hypotheses are the hierarchy hypothesis and the

---

<sup>2</sup> To ensure that the case markers were relevant to the sentence's interpretation, all test items were semantically reversible with either animate or inanimate referents for both the direct object and the indirect object in Cho et al (2002)'s study. There are two variants of the dative marker found in the test sentences: hanthey, which is used for NPs with animate referents, and -ey, which is used for NPs with inanimate referents in Korean.

<sup>3</sup> According to Cho et al. (2002: 902), this developmental finding could not be attributed to input frequency, because the preferred accusative-dative order was found to be far less frequent than the

iconicity hypothesis. The hierarchy hypothesis hypothesizes that children prefer sentences whose word order reflects the relative prominence of grammatical relations based on the NP accessibility hierarchy (see Keenan and Comrie, 1977). The relational hierarchy cited in Cho et al. (2002: 902) is repeated in (3) as follows.

(3) The relational hierarchy

subject > direct object > indirect object > oblique > ...

The developmental preference of the accusative-dative order (direct object-indirect object) over the dative-accusative order (indirect object-direct object) supports the hierarchy hypothesis, because the accusative-dative order aligns with the structural prominence of the direct object over than indirect object, while the reverse dative-accusative order does not.

The ‘accusative-dative > dative-accusative’ preference findings nevertheless also support the other hypothesis considered, namely the iconicity hypothesis, because the accusative-dative order is iconic to the situation while the reverse dative-accusative order is not. (4) shows an illustrative pair of example from Cho et al. (2002).

(4) The Korean accusative-dative and dative-accusative pattern investigated in Cho et al. (2002)

a. The Korean accusative-dative pattern: (iconic order)

Ai-ka	tol-ul	kay-hanthey	tenci-ess-ta
Child-NOM	stone-Ac	dog-Dat	throw-Pst-Decl

'The child threw a stone to the dog.' (=2b) in Cho et al., 2002)

b. The Korean dative-accusative pattern: (non-iconic order)

Ai-ka	kay-hanthey	tol-ul	tenci-ess-ta
Child-NOM	dog-Dat	stone-Ac	throw-Pst-Decl

'The child threw a stone to the dog.' (=2a) in Cho et al., 2002)

Transfer event: the child acted on the stone causing it to go to the dog (Cho et al., 2002: 903)

In order to tease apart which hypothesis is more tenable in their case, they conducted a follow-up study investigating children's word order preferences for the instrumental pattern in Korean. (5) shows an illustrative pair of example of the two word order patterns investigated in their study.

(5) The Korean instrumental-accusative and accusative-instrumental pattern investigated in Cho et al. (2002)

a. The Korean instrumental-accusative pattern: (iconic order)

Subject	instrument	direct object	
Yenghi-ka	phey-n-ulo	yenphil-ul	kentuli-ess-ta
Yenghi-NOM	pen-Instr	pencil-Ac	touch-Pst-Decl

'Yenghi touched the pencil with the pen.' (=9) in Cho et al., 2002)

b. The Korean accusative-instrumental pattern: (non-iconic order)

Subject	direct object	instrument	
Yenghi-ka	yenphil-ul	phey-n-ulo	kentuli-ess-ta
Yenghi-NOM	pencil-Ac	pen-Instr	touch-Pst-Decl

'Yenghi touched the pencil with the pen.' (=9) in Cho et al., 2002)

Corresponding Event: Yenghi first acted on the pen and then used it carry out an action involving the pencil (Cho et al., 2002: 907)

The instrumental pattern is chosen because unlike the situation in the dative pattern, the two hypotheses would yield *opposite* preference predictions: the hierarchy hypothesis would predict preference of the accusative-instrumental (direct object-indirect object) order, but iconicity would predict the preference of the instrumental-accusative order because only the instrument-accusative linear order aligns with the structure of the corresponding situation. Table 2.1 reports the summary table of predictions given in Cho et al. (2002: 904).

**Table 2.1. Predicted word order preferences in Cho et al. (2002)**  
(=based on table 2 in Cho 2002: 904)

Hypothesis	Dative pattern	Instrumental pattern
Iconicity Hypothesis	accusative-dative	instrumental-accusative
Hierarchy Hypothesis	accusative-dative	accusative-instrumental

Results supported the prediction of the iconicity hypothesis, but not the hierarchy hypothesis, because the children did far better on the instrument-accusative order than the accusative-instrument order, whose interpretation they tended to reverse to the instrument-accusative order.<sup>4</sup>

---

<sup>4</sup> Similar to the Korean dative pattern discussed earlier on, there appears to be nothing in the child-directed speech that could account for this finding of developmental preference. Cho et al. (2002: 906) pointed out that although the developmentally preferred instrumental-accusative order is considered to be basic in adult Korean according to Sohn (1994: 232), this order appears to be infrequent in the adult input. Indeed, their corpus findings on Korean maternal speech showed only two sentences with both the instrument and accusative arguments overtly expressed amidst the prevalence of null arguments, and these two sentences were both in the reverse accusative-instrument order. There is thus no evidence to show that the instrumental-accusative order is significantly more frequent than the accusative-instrumental order in Korean adult input.

### 2.3 The Input Frequency Hypothesis: O'Grady (2000) and Campbell and Tomasello (2001)

Input frequency is another factor considered in the existing literature on the acquisition of dative constructions. O'Grady (2000) also investigated the role of input frequency in his comprehension study on the acquisition of English dative constructions. The hypothesis is cited below:

“Children acquire the more frequent structure first.” (O'Grady, 2000: 3)

The significance of frequency for the emergence of linguistic structures is well attested in current linguistic and psycholinguistic literature (see for example Bybee and Hopper, 2001), but different ways of counting frequency can lead to different predictions, and researchers differ in making use of frequency in accounting for acquisition facts. O'Grady (2000) claimed that if all the prepositional dative constructions (PDC) and double object dative constructions (DOC) attested are counted, regardless of what the verb is, prepositional datives would exceed the double object datives because there are far more verbs that can appear in the prepositional datives in English. However, if only the so-called *dative alternation verbs* appearing in a dative construction are counted, there would be more DOC than PDC. On the other hand, Campbell and Tomasello (2001), working on the early acquisition of English dative constructions in *production*, computed frequency counts at the level of individual verbs.<sup>5</sup> When individual verb frequencies are

---

<sup>5</sup> There is justification in computing verb frequencies in such a manner. There are documented lexical biases in favoring one syntactic frame over the other even when both frames are possible with that particular verb (Wasow, 1997). In English adult input, individual verbs differ with respect to the frequency with which their postverbal constituents appear in one order rather than another (Wasow, 2003).

counted, they claimed that input frequencies suffice to explain the attested developmental order in production.

## **2.4 Relevance to the Cantonese Case**

The acquisition of the *bei2* dative construction in Cantonese exemplifies a case relevant to the acquisition hypotheses stated above. I now attempt to apply the hypotheses to the target Cantonese construction. As we shall see, these hypotheses yield different predictions on the developmental preferences of the canonical and non-canonical forms of the *bei2* dative construction.

### **2.4.1 The Markedness Hypothesis: Empirical Predictions**

I apply O'Grady (2000)'s markedness hypothesis from a cross-linguistic perspective to the present case, while taking into account the cross-linguistic findings specifically related to the verb *give* reported in the current literature. Michaelis and Haspelmath (2003) recently reported on a world-wide survey of ditransitive constructions in about 250 languages from around the world based on the verb 'give'. Their findings are so far the most comprehensive to date that I am aware of in establishing the cross-linguistic markedness of dative constructions related to the target verb 'give'. In order to classify constructions into types, they only looked at the construction used with the verb 'give' because, as they noted, "in many languages, different ditransitive verbs have a different construction". Their cross-linguistic findings would therefore provide a good matching comparison with the present study on the Cantonese verb *bei2* 'give'.

Contrary to what might be generally believed about the markedness of DOC in the world languages, Michaelis and Haspelmath (2003) argued that *both* double object

constructions and what they called indirect object constructions (where the indirect recipient is ‘flagged’- the English PDC would be a case of their indirect object construction, because the indirect object is introduced by a preposition) are cross-linguistically frequent (unmarked) across languages in the world. Their claim is cited below.

“...what seems to appear clearly from this map is that both the IOC and the DOC are major constructions, none of which is in any way unusual or “marked” compared to the other one.”  
(Michaelis and Haspelmath, 2003)

Taking into account Michaelis and Haspelmath (2003)’s recent findings, DOC does not appear to be more marked than PDC on a more global scale, although they referred only to the ditransitive constructions with the ‘give’ verb. Michaelis and Haspelmath (2003) have ignored word order in defining DOCs when reporting their findings, but as I look at all the DOC examples cited in their paper, they all exhibit the [V-R(ecipient)-T(heme)] order (see (9c) in chapter one as an illustrative example). The [V-T-R] double object form (which is also the canonical form of the Cantonese *bei2*-dative (see (1) in chapter one) as described in section 1.1.1), does not appear in any examples cited in their global study. Matthews and Leung (2002) have already pointed out that the [V-T-R] double object construction is a marked order in languages without case. Susanne Michaelis (2003, p.c.) also pointed out that the [V-T-R] double object construction appears to be dispreferred in the world languages. As far as the existing literature on dative constructions is concerned, I am aware that the [V-T-R] double object form is attested only in the following languages: Cantonese (Tang, 1998), Thai and other Tai languages (Matthews and Leung, 2002), some other Chinese dialects (see Liu, 2001) and Ewe in West Africa (see Essegbey, 2002). For other languages attesting DOCs, they are in the [V-R-T] order.

The [V-R-T] word order has also been implicated as the unmarked word order from a typological perspective in Kozinsky and Polinsky (1993). Kozinsky and Polinsky (1993) proposed a tentative universal on coding of the thematic recipient and theme in ditransitive constructions: in *Agent-before-Patient* languages, the recipient precedes the theme; and in *Patient-before-Agent* languages, the theme precedes the recipient. Since *Patient-before-Agent* languages are few, the [V-T-R] order can be regarded as a cross-linguistically marked option by frequency and is predicted to be especially unusual to occur in an *Agent-before-Patient* language. In this regard, Cantonese is a partial exception with respect to the verb-recipient-theme/agent-before-patient correlation, since Cantonese is an *Agent-before-Patient* language but instantiates *both* the [V-R-T] (see (9c)) and [V-T-R] word orders (see (9a)). It is also relevant to point out that the [V-T-R] double object construction is called the ‘inverted’ double object construction in Tang (1998). His use of the word ‘inverted’ appears to reflect the perception that this word order is special, in the sense that the postverbal ordering of the two objects *looks* inverted when compared to the otherwise cross-linguistically frequently attested [V-R-T] order. Taking the cross-linguistic facts into account, the Cantonese *bei2*-dative in its canonical double object form [V-T-R] can be viewed as a marked option cross-linguistically. According to the markedness hypothesis stated in section 2.1, the cross-linguistic markedness of the [V-T-R] construction might imply that, for whatever reasons, this double object construction is supposed to be difficult for acquisition.

On the other hand, the non-canonical [*bei2*-R-T] double object form, being considered as an instance of the [V-R-T] double object construction, and the [*bei2*-T-*bei2*-R] serial verb form, being considered as an instance of the indirect



object construction in Michaelis and Haspelmath (2003)'s terms, in which the indirect recipient object is 'flagged' (here with a dative marker *bei2*, i.e. in the form of [V-T-dative marker-R]) are structurally more unmarked (prevalent) than the canonical form [V-T-R] cross-linguistically. Given their cross-linguistic unmarked status, according to the markedness hypothesis, these non-canonical forms might be easier to acquire than the canonical form from a developmental point of view. (6) presents a predicted preference of acquisition based on the markedness hypothesis.

- (6) The predicted preference of the non-canonical *bei2* forms over the canonical *bei2* form according to the markedness hypothesis:

$$\text{non-canonical [}i2\text{-T-}i2\text{-R]} = \text{non-canonical [}i2\text{-R-T]} > \text{canonical [}i2\text{-T-R]}$$

#### 2.4.2 The Iconicity Hypothesis: Empirical Predictions

Applying the iconicity hypothesis to the present case, this hypothesis would motivate the preference of the canonical [*bei2*-T-R] form, because this word order aligns with the structure of the transfer event described in (1). Moreover, one can also argue that temporal sequence might also be involved (c.f. Tai (1985)'s Principle of Temporal Sequence (PTS), where PTS can be regarded as one sub-case of iconicity), because in this case, one has X(agent) and Y(theme) to start with, and Y ends up in Z(recipient)'s possession at the end of the scene. Therefore, iconicity might be a potential factor in favoring the canonical [*bei2*-T-R] ordering.

As for the non-canonical [*bei2*-R-T] form, like the double object form in English, its word order does not align with the structure of the transfer event stated in (1) because in this case, the recipient precedes the theme. Therefore, based on the structure of a transfer event described in (1), the [*bei2*-R-T] form lacks iconic motivation.

For the non-canonical [*bei2*-T-*bei2*-R] serial verb form, similar to the canonical [*bei2*-T-R] form, the linear ordering of the theme and the recipient arguments is also isomorphic to the structure of the transfer event described in (1) unfolds, in that the causer or the agent directly causing the theme moves to a new location (recipient)), it is therefore motivated by the iconicity hypothesis.

Table 2.2 summarizes whether iconic motivation exists in the canonical and the non-canonical *bei2* forms with reference to the structure of a transfer event described in O'Grady (2000: 4) (see (1)).

**Table 2.2. A summary of iconic motivation for the canonical and the non-canonical *bei2* forms**

Canonical double object form	Non-canonical double object form	Non-canonical serial verb form
[ <i>bei2</i> -T-R]	[ <i>bei2</i> -R-T]	[ <i>bei2</i> -T- <i>bei2</i> -R]
iconic	lack motivation	iconic

Note that the iconicity hypothesis has so far been studied in comprehension studies. How is this hypothesis relevant here in the present context of *production* in Cantonese? It might be relevant in the sense that sentence production involves mapping a conceptual schema onto a constituent structure, and this might be easier to do to the extent that the two are isomorphic. This hypothesis might be especially relevant in the acquisition of pro-drop languages like Korean and Cantonese because of the following reason. Since the target *bei2*-dative is often 'underspecified' on surface syntax because the theme is unexpressed, the child might not be clear about where the implicit theme should be placed. One might wonder whether young children go by the iconicity principle to order the arguments in their *production*

according to how the transfer event naturally unfolds when they have to express both the theme and the recipient.

### **2.4.3 The Input Frequency Hypothesis: Empirical Predictions**

According to the input frequency hypothesis, children acquire the more frequent structure in the input first. If the non-canonical *bei2* forms are used only in marked contexts in adult Cantonese, one might expect that the non-canonical [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms are used relatively less often than the canonical [*bei2*-T-R] form in the adult input. Moreover, given that natural discourse usually does not require speakers to specify a referent with a long NP in most cases, it is likely that the verb *bei2* is used frequently in syntactic contexts with short theme NPs that motivate omitting the dative marker (see section 2.5), hence the frequent use of the canonical [*bei2*-T-R] form. The input frequency hypothesis would therefore predict that the canonical form is preferred over the non-canonical forms. I shall strengthen this empirical prediction based on a corpus analysis of adult Cantonese child directed speech by comparing the frequencies of the canonical and non-canonical *bei2*-datives in the adult input (see section 3.4).

### **2.4.4 An Interim Summary of Empirical Predictions**

Table 2.3 presents an interim summary of the empirical predictions. The markedness hypothesis would predict the developmental preference of non-canonical forms over the canonical form. The iconicity hypothesis would predict a developmental preference for the canonical [*bei2*-T-R] double object form and the non-canonical [*bei2*-T-*bei2*-R] serial verb form, but not the non-canonical [*bei2*-R-T] double object form. The input frequency hypothesis would predict developmental preference of the canonical [*bei2*-T-R] form over the non-canonical forms, based on

the relative frequencies of these forms which are confirmed by a corpus study of adult input findings reported in section 3.4.

**Table 2.3. An interim summary of empirical predictions**

	Canonical [bei2-T-R]	Non-canonical [bei2-R-T]	Non-canonical [bei2-T-bei2-R]
The Markedness Hypothesis: Children find it easier to acquire unmarked structures- that is, structures that are for one reason or another more common in the world's languages	Dispreferred	Preferred	Preferred
The Iconicity Hypothesis: Children prefer iconic structures- that is, structures whose word order reflects the manner in which the corresponding event unfolds	Preferred	Lack motivation	Preferred
The Input Frequency Hypothesis: Children acquire the more frequent structure first	Preferred	Lack motivation	Lack motivation
	See also section 3.4 in chapter three		

There is another hypothesis, called the input properties hypothesis here, which I will investigate in this thesis. Working within the usage-based theory of child language acquisition (Tomasello, 2003), I shall investigate how the Cantonese language specific input properties might influence early schematization in the theory. This issue will be investigated in the next chapter.

**2.5 The Null Dative Marker Hypothesis**

The null dative marker hypothesis was proposed in a generative framework by Tang (1998), Xu and Peyraube (1997) among others. Tang (1998: 42) used the place of pausing, as denoted by # here, noted between the theme and the recipient when Cantonese-speaking adults produce [V-T-R], i.e. [V-T-#-R], as one piece of evidence to support his null dative marker hypothesis. The pause signals where the null

dative marker is. Xu and Peyraube (1997: 121) provided further important diachronic evidence to give credence to the hypothesis that [V-T-R] is derived from [V-T-(dative marker)-R] by omitting the dative marker. In the oracle bone inscriptions (Pre-Archaic Chinese), the instances of [V-T-R] and the [V-T-dative marker-R] usually appeared in the same inscription and it was found that the [V-T-R] and [V-T-dative marker-R] structures attested were always used with the same verbs and most often even with the same theme and recipient NPs as well.

Given that the lexical verb *bei2* and the dative marker *bei2* are homophonous in Cantonese, working within the minimalist framework, Tang (1998) proposed that avoidance of phonological identity might be the motivating factor in omitting the *bei2* dative marker.

‘Suppose that the avoidance of phonological identity is one of those PF interface conditions... If that turns out to be correct, an optimal output would be a structure in which the dative marker is null when it is required for PF convergence.’ (Tang, 1998: 44)

Here I point out that the motivation for the omission of the dative marker *bei2* may not be restricted to the avoidance of phonological identity. According to Tang (1998), the other Cantonese ‘give’ verbs in table 1.1 are also acceptable in the [V-T-R] form, yet these verbs and the dative marker *bei2* are not phonologically identical. Consequently, avoidance of phonological identity should not be the single determining factor responsible for the omission of the dative marker, because it fails to account for cases when the other Cantonese ‘give’ verbs in table 1.1 are also acceptable in the [V-T-R] form. It seems, therefore, that we have to look for what is in common between the Cantonese ‘give’ verbs as a verb class that allows omitting the dative marker.

Liu (2001) recently offered a functional perspective for omitting the dative marker. His idea is that the omission of the dative marker is motivated by the functional principle of economy. Contrary to Tang (1998), he treated the dative marker as having the categorical status of a verb (and more specifically as a ‘give’ verb) rather than as a preposition, hence the [V-T-*bei2*-R] form is regarded as a serial verb structure in his work. The [V-T-R] double object construction, called a ‘pseudo’ double object construction, is derived from omitting the ‘give’ verb (the dative marker) from the serial verb structure. His idea is cited as follows.

“The ‘pseudo’ double object construction, derived from omitting the ‘give’ verb in a serial verb structure, is a product of the functional principle of economy.” (Liu, 2001: 392, my translation)

I shall elaborate his idea as follows. I propose that the motivation for omitting the dative marker is due primarily to the functional principle of economy because of semantic overlap between the ‘give’ verbs and the dative marker rather than due to the avoidance of phonological identity as argued by Tang (1998). As mentioned in Table 2.1 above, the [V-T-R] form in Cantonese is associated only with the ‘give’ verbs that are genuine ditransitive verbs and inherently transactional verbs that encode a specific path of directed motion (motion in ownership-space) pointing in the direction toward the recipient (as the intended possessor) that overlaps semantically with the directional sense encoded by the dative marker *bei2*. With the existence of semantic redundancy in such cases, the *bei2* dative marker can be omitted driven by the functional principle of economy, yielding the surface form [V-T-R].

## 2.6 Conceptualization of End-State Knowledge: Construction

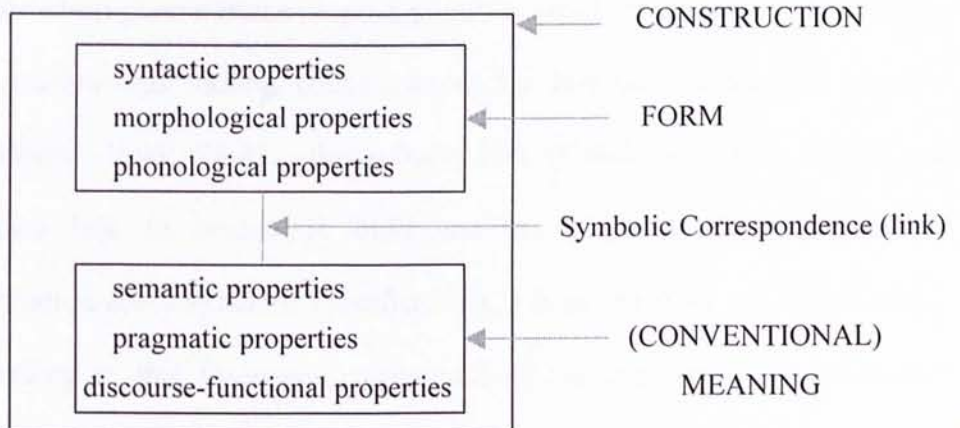
### Grammar

The usage-based approach to child language acquisition (Tomasello, 2003) adopts a constructional view of grammar. Construction grammar (Fillmore et al., 1988; Goldberg, 1995; Croft, 2001) is a proposal of a grammatical framework in cognitive linguistics that aims at a unified representational account of all grammatical knowledge, not just the so-called ‘core grammar’ as in the generative paradigm. The major premises of the construction grammar framework which are adopted by the usage-based theory of child language acquisition are highlighted below:

- i) Construction grammar is proposed as a reaction to a componential view of grammatical organization, and to a reductionist approach to grammar. It explicitly rejects the assumption that grammar is to be neatly divided into separate components - lexicon and syntax (Goldberg, 1995: 23; Croft, 2001), but holds that there is no (and cannot be any) clear division between the lexicon and syntax. The argument stems from the impossibility of making a clear distinction between the core and the periphery of linguistic structure.

- ii) Language is not an autonomous cognitive faculty. The representation of linguistic knowledge is essentially the same as the representation of other cognitive structures. Constructions are viewed as symbolic units, being pairings of a syntactic pattern with a meaning structure, so linguistic constructions themselves are meaningful symbolic units (as lexical items are). In the case of constructions which include phrase level and clause-level phenomena there is further internal structure.

**Figure 2.1. The internal symbolic structure of a construction (=Croft and Cruse, 2003: Figure 10.1)**



- iii) The grammar of a language is represented as a structured inventory of linguistic structures of the language falling along a continuum of abstractness and complexity, consisting of morphemes, words, word classes, phrasal constructions such as noun phrase and prepositional phrase, and sentence-level constructions. Constructions can vary in schematicity and abstractness; some may have some specific words or morphemes associated with them, for instance, the “*let alone*” in the *let alone* construction (Fillmore et al., 1988), the “*by*” in the English full passive construction, the ‘*s*’ in the genitive. As such, grammatical knowledge is represented in constructions, which



encompass all aspects of grammatical units, from substantive lexical items to schematic syntactic constructions all being formalized under the notion of a construction.

- iv) Unlike in transformational grammar, whereby the variable syntactic structures are proposed to be related via syntactic transformations, no transformational relationships are assumed in the construction grammar approach.<sup>6</sup> However, it is not claimed that constructions are organized or stored in the mind of a speaker as an unstructured list of entities. Construction grammarians propose possible non-derivational ways to capture the relationships among constructions, for instance, taxonomic relations (Goldberg, 1995: 74-81), the subpart link (Goldberg, 1995: 78-79), the instance link (a taxonomic link) and the polysemy link whereby all information about syntactic specifications is inherited from the central sense. According to this framework, constructions are organized in a taxonomic network, like the organization of information in other cognitive domains. The linguistic competence of adult mature speakers of a language is characterized as a “structured inventory of symbolic units” in the minds of speakers (Langacker, 1987; Croft, 2001). Linguistic competence is characterized in terms of the mastery of a structured inventory of meaningful linguistic constructions of a particular language. From an acquisition point of view, formalization of the adult language represents the end state of what knowledge is acquired and how it is organized in the adult mental grammar.

---

<sup>6</sup> The need for transformation or derivation is explicitly eschewed in many other contemporary theoretical approaches, see for example, Bresnan (1982); Bresnan (1994); Fillmore and Michaelis (in progress); Lakoff (1987); Langacker (1987); Langacker (1991); Pollard and Sag (1987), all cited in Goldberg (2002a).

- v) Actual expressions are not necessarily stored as entire units unless they occur with some regularity and become collocations or idiomatic expressions, or from a usage-based perspective, they are frequently used and become independently stored in the mental grammar. Otherwise, actual expressions are created on the fly by existing constructions which may be superimposed on one another as long as no conflicts arise.<sup>7</sup>

## **2.7 Introducing the Usage-Based Theory of Child Language Acquisition (Tomasello, 2003)**

The usage-based theory of child language acquisition (Tomasello, 2003) aims at proposing for some usage-based principles of learning and production to complement construction-based accounts of linguistic structure. This theory does not only focus on the core-aspects of the language, but attempts to provide a unifying framework/mechanism for acquisition of all linguistic structures/ items by means of a set of general psycholinguistic processes. One central hypothesis of the theory is that all linguistic structures are acquired by means of a common set of social learning and

---

<sup>7</sup> This point is related to the principle of unification which concerns the conditions under which constructions can be combined. Goldberg (2002b) explained unification as “a formal principle of combination wherein consistent attribute value matrices can be combined unless there is an overt clash of features.” Mirjam Fried (April 2003, p.c.) also provided a detailed explanation of unification as follows:

“Since construction grammar is non-derivational, unification is the system that ensures that constructions do not combine at random, either with other constructions or with lexical items. Anything that is to combine to form larger pieces of linguistic structure must match along specific features. Grammatical information is expressed in the form of feature-value pairs, where the feature is a linguistic category (e.g., role, grammatical function, case, number, animacy, etc.) and the value is either binary (for some of the features), or one out of a list (e.g. SUB, OBJ, etc. for the feature 'grammatical function'), or another feature-value pair, or is left unspecified (in empty square brackets). And the general principle is that two things can unify only if their values do not conflict.”

Sincere thanks to Professor Mirjam Fried for explaining this important concept to me.

cognitive processes. Tomasello (2003: 295) reviewed four basic sets of acquisition processes by means of which children construct a structured inventory of linguistic constructions of the particular language(s) they are acquiring. Table 2.4 shows these processes.

**Table 2.4. Four basic sets of acquisition processes in Tomasello (2003: 295)**

Intention Reading and Cultural Learning	Account for how children learn linguistic symbols in the first place
Schematization and Analogy	Account for how children create abstract syntactic constructions out of the concrete pieces of language they have heard
Entrenchment and Competition	Account for how children constrain their abstractions to those that are conventional in their linguistic community
Functionally based distributional Analysis	Account for how children form paradigmatic categories of various kinds of linguistic constituents

Usage-based approaches to language (Barlow and Kemmer, 2000; Bybee, 1995; Croft, 2001; Langacker, 1987) hold the major fundamental assumption that aspects of use affect grammar. Tomasello (2003) assumes essentially the same central hypothesis, that language structure emerges from language use:

“Beginning at the beginning, for usage-based theorists the fundamental reality of language is people making utterances to one another on particular occasions of use. When people repeatedly use the same particular and concrete linguistic symbols to make utterances to one another in “similar” situations, what may emerge over time is a pattern of language use, schematized in the minds of users as one or another kind of linguistic category or construction.” (Tomasello, 2003: 99)

In this view, performance can affect competence, which is largely different from the classical Chomskian view that sets a clear demarcation between competence and performance. I adopt the usage-based assumption that the properties of use of utterances in children's adult input and children's own usage patterns influence the representation of the grammatical structures in children's mind.

One central concern of the theory is constructional schema formation. Constructional schemas, also called schematic constructions (Langacker, 1987; Bybee, 1995), are symbolic units derived from actual expressions and have the same structure as their instantiations. They emerge as generalizations of concrete expressions by schematization. Their original function is to allow more efficient storage; but once abstracted and emerged, they are also available for the construction of novel expressions. They account for linguistic generalization and contribute to creativity of a language (Tomasello, 1998), which in the generative context, is formulated by grammatical rules. They are schematic symbolic units that provide slots for concrete lexical items.

## Chapter Three. The Input Properties Hypothesis and Adult Cantonese Input

### 3.0 Introduction

This chapter focuses on the input properties hypothesis, which is related to schematization in the usage-based approach to child language acquisition (Tomasello, 2003). I examine how the Cantonese language specific input properties might impact schematization. From a usage-based perspective, I focus on exploring how null arguments and displaced arguments might influence early schematization of the canonical [*bei*2-T-R] schema from the input. My hypothesis is that the Cantonese adult input properties are not structured in a way that facilitates early schematization of the canonical double object [*bei*2-T-R] form. Section 3.1 discusses schematization in the usage-based theory. Section 3.2 discusses the input properties hypothesis in relation to schematization in the theory. Section 3.3 puts forth two empirical hypotheses on the Theme-Recipient asymmetry in argument realization and displacement. Section 3.4 presents findings from a corpus study on Cantonese adult child-directed speech to substantiate the two empirical hypotheses in section 3.3. Section 3.5 discusses the implications the findings bring to early schematization. Section 3.6 offers a chapter summary.

### 3.1 Schematization

The usage-based theory of language acquisition (Tomasello, 2003) holds a discontinuous view of child language acquisition. In this theory, the child is not born with adult-like abstract categories and constructional schemas, but these abstract categories and schemas emerge as result of categorization of linguistic

experience from the input and from the child's own usage patterns. Consequently, constructional schema formation is one major concern in the theory, although there is very little relevant research here (Tomasello, 2003).

Schematization is the psycholinguistic process Tomasello (2003) hypothesized to solve the problem of continuity concerning how children can build up abstract constructional schemas from their early lexically-specific concrete constructions. Empirical studies of how children schematize concrete expressions into constructional schemas actually take place are considered of direct relevance for the usage based theory of language acquisition, although there is so far no study addressing the question of exactly what kinds of linguistic experience children must have in order to form such an abstraction (Tomasello, 2003: 125).

Schematization is a process of categorization and analogy on the basis of both form and function. Although perceptual similarity in form helps, the essence underlying how schematization works lies in analogy based on function. For the formation of sentence-level abstract constructional schemas, the usage-based approach to child language acquisition (Tomasello, 2003) hypothesized that schematization works by making structural alignment across different utterances and verb-specific constructions. Let me illustrate the idea with the following examples in (1) for the present case.

- (1) a. Bei2 go3 caang2 ngo5 aa1  
 Give CL orange me PRT  
 'Give an orange to me.'
- b. Bei2 go3 caang2 lei5 aa1  
 Give CL orange you PRT  
 'Give an orange to you.'
- c. Bei2 bun2 syu1 lei5 aa1  
 Give CL book you PRT  
 'Give a book to you.'
- d. Bei2 jat1 gim6 saam1 lei5 aa1  
 Give one CL suit you PRT  
 'Give one suit to you.'
- e. Bei2 cin2 maa1mi4 aa1  
 Give money mother PRT  
 'Give money to mother.'

The similarity involved between *caang2* 'orange', *syu1* 'book', *saam1* 'suit' and *cin2* 'money' in is that they all refer to entities that are being transferred (hence the notion *theme*) in the event, while the similarity involved in *ngo5* 'me', *lei5* 'you', *maa1mi4* 'mother' is that they all refer to the entities the transferred entities are directed to (the *goal*) or the entities that would receive the *thing* transferred (the *recipient*). In hearing utterances like (1a)-(1e) which describe a transfer event in meaningful context, according to this theory, if we hypothesize a verb-specific [*bei2*-T-R] schema in the adult Cantonese grammar, one way for children to create this schema at this level of abstractness is to abstract it from concrete instances like (1) from the adult input or/and from the children's own utterances. The psycholinguistic process hypothesized in the theory to abstract a constructional schema from concrete instances is called schematization.

Schematization involves making structural alignments and constructional analogies, according to this theory. Function is crucially involved in constructional analogies. The theory hypothesizes that children must have an understanding of the functional interrelation among the elements of the constructions involved. If utterance-level constructions are associated with distinct semantic/ event structures in construction grammar (see Goldberg, 1995), the *arguments* like agent, theme, recipient would be relevant entities, because these NPs play different functional roles in the larger semantic/ event structure.

An important proposal for early schematization concerns the consistency of the items in the slots, that is, a given item occurs in one slot and not in others. Gentner and Medina (1998) suggested that when ‘all kinds of items occur promiscuously in all the slots in two potentially analogous relational structures, structure mapping is made more difficult’ (Tomasello, 2003: 166).

### **3.2 The Input Properties Hypothesis**

In the spirit of the usage-based theory, I shall focus on the above proposal concerning the consistency of functional items.<sup>1</sup> I frame this proposal as the input properties hypothesis, which I define as follows.

---

<sup>1</sup> Type frequency is another crucial factor hypothesized to influence schematization (abstraction), although the type frequencies in the adult input required for creating slots in constructional schemas has also never been studied systematically (Tomasello, 2003: 125). We currently do not know the frequencies required in different cases. I therefore do not attempt to focus on type frequency in this study.



(2) The Input Properties Hypothesis:

Children find it easier to abstract from a linguistic environment structures whose functional items consistently occur in particular positions.

The functional items, however, may not always consistently occur in particular positions in adult Cantonese, because Cantonese allows arguments to be unexpressed and displaced on surface syntax. I shall investigate to what extent the qualitative properties of the Cantonese adult input in terms of null and displaced arguments in the Cantonese input might influence the early abstraction of the canonical [*bei2*-T-R] schema.

### 3.3 Two Empirical Hypotheses on the Theme-Recipient Asymmetry

As stated in section 1.1.1, the canonical form of the Cantonese *bei2* dative construction is [*bei2*-T-R]. However, there are variable surface word orders instantiated in Cantonese discourse that have consequences for the realization and ordering of the theme and recipient arguments of this construction. This section formulates two empirical hypotheses on the Theme-Recipient Asymmetry in argument realization and susceptibility to displacement in Cantonese adult child-directed discourse. The two empirical hypotheses are as follows:

- i) the theme is more often unexpressed than the recipient for the Cantonese *bei2* double object construction (see section 3.3.1)
- ii) the theme is more often displaced than the recipient for the Cantonese *bei2*-double object construction (see section 3.3.2)

### 3.3.1 Unexpressed Arguments: The Theme Versus the Recipient

Cantonese is a pro-drop language allowing null arguments, the *bei2*-double object dative, therefore, much like the other constructions, is bounded by the argument-realization conditions. In certain identifiable contexts, when the referent of an argument is mentioned in previous discourse and is immediately available in the context, the argument may be unexpressed or so-called *null* in Chinese (see Huang, 1984 for Mandarin; and Matthews and Yip, 1994 for Cantonese null arguments). In principle, null subjects and objects are allowed in Cantonese, in other words, the agent, theme or/ and the recipient arguments can be null.

But here the hypothesis that the theme argument is more often unexpressed in natural adult Cantonese discourse than the recipient argument for Cantonese datives. One possible motivation to ‘leave out’ the theme argument more frequently than the recipient argument might have to do with the predictability of the theme argument versus the recipient argument. Intuitively, the theme argument (the entity being transferred) can be more readily presupposed from the immediate context or/and the previous discourse than the recipient argument, as long as the referent of the theme argument is either present in the immediate context or has been mentioned in the previous discourse. (3) shows an illustrative example for the adult use of a null theme *bei2*-dative where the referent of the theme is in the immediate context.

(3) LTF, File 20210, Line 3303

\*CHI: xxx.

Situation: The child uses the puppet's head to press on things.

\*INV: Wa3: , bei2 ngo5 aa1 , hou2 m4 hou2?

PRT, give me PRT, good not good

‘Give me, ok?’

(4) shows an illustrative example for the adult use of a null theme *bei2*-dative because the referent of the theme was mentioned in previous discourse.

(4) MHZ, File 11023, Line 3836

\*CHI: Bei2 bo1bo1 .  
Give ball  
'Give ball'

\*INV: Bei2 bin1go3 a3 ?  
Give who PRT  
'Give who?'

In contrast, the recipient of a dative construction has less predictable referent. Even when the referent of the intended recipient is physically present in the immediate context, unless the destination of the transferred object (the goal/ recipient) can be clearly predicted, as in (5) where the (intended) possession relationship is made clear or as in (6)-(7) where the (intended) possessor can be inferred from the context (the intended possessor of *cin2* 'money' is the speaker in (6) when she is requesting it, and the possessor of *wun6geoi6* 'toys' is the child in (7)), the goal has to be clearly specified, and so the recipient cannot be unexpressed.

(5) CHI: Bei2 ngo5 aa1  
Give me PRT  
'Give me.'

INV: Bei2 me1 aa3?  
Give what PRT  
'Give what (to you)?'

(6) INV: Bei2 cin2 aa3  
give money PRT

Situation: the adult is handing out her hand to the child.

- (7) CHI: Ngo5 jau5 wun6geoi6  
 I have toy  
 'I have toys.'
- INV: Bin1 go3 bei2 gaa3?  
 Who give PRT  
 'Who gave (you) (the toys)?'

Animacy may also play a role here.<sup>2</sup> It has been shown that even in languages where null complementation is fairly free, animate arguments are the least likely to be omitted (see Schwenter and Silva, 2002 on Brazilian Portuguese). A number of authors relate this phenomenon to 'differential object marking', the idea is that less prototypical objects need more marking than prototypical ones. Animate patients are 'marked', as in Spanish, where they receive a special dative preposition. Ruppenhofer's (m.s.) study of null complementation involving ambiguous verbs like 'follow' shows that animate 'followees' are less likely to be null than inanimate ones. Applying the role of animacy in null complementation to the present case, since the vast majority of the recipients are animate, and the vast majority of the theme arguments are inanimate, I hypothesize that the theme is more often unexpressed than the recipient for the Cantonese *bei2*-dative in natural Cantonese discourse.

The consequence of frequent unexpressed themes is that the recipient rather than the theme often surfaces as the first postverbal argument for the Cantonese *bei2*-dative.

### 3.3.2 Susceptibility to Displacement: The Theme Versus The Recipient

The theme, even when overtly expressed, is more likely to be displaced than the recipient in Cantonese discourse. The canonical [*bei2*-T-R] double object form can interact with other constructions such as the topic construction, the right-dislocated

---

<sup>2</sup> Thanks to Professor Laura Michaelis for her input on this point.

construction, the cleft constructions, and the relative clause construction, resulting in the arguments being in ‘displaced’ positions in the surface syntax. There is an asymmetry in ‘displacing’ the theme versus the recipient, the recipient is less susceptible to displacement compared to the theme- it can only be ‘displaced’ in special contexts, such as contrastive contexts. Without these special contexts, displacement of the recipient is not allowed. (8) shows an example of the basic canonical [*bei2*-T-R] double object form of the *bei2*-dative. (9) and (10) show the asymmetry observed when the basic construction interacts with the topic construction. (11) and (12) show the asymmetry observed when the basic construction interacts with the right-displacement construction.

*The basic bei2 double object construction* (canonical form: [V-T-R])

- (8) Ngo5 bei2 lei1 bun2 syu1 lei5 aa1  
 I give DET CL book you SFP  
 ‘I give this book to you.’

*Interaction with the Topic Construction*

The theme, in the ‘topic’ position, surfaces at the utterance-initial position.

- (9) Lei1 bun2 syu1 ngo5 bei2 lei5 aa1  
 DET CL book I give you SFP  
 ‘This book, I give you.’

‘Topicalizing’ the theme sounds more natural than ‘topicalizing’ the recipient (compare (9) and (10)).

- (10) ??Lei5 ngo5 bei2 lei1 bun2 syu1 aal  
 You I give DET CL book SFP  
 ‘You, I give this book.’

*Interaction with the Right-Dislocation Construction*

(11) illustrates the theme argument being in the right-dislocated position.

- (11) Ngo5 bei2 lei5 aal lei1 bun2 syu1  
 I give you SFP DET CL book  
 ‘I give you, this book.’

It is important to point out that the right-dislocated structure exemplified in (11) where the theme argument is *displaced* to the right *after* the sentence final particle is different from the non-canonical [*bei2-R-T*] double object form described in section 1.1.2, where the theme appears to the right of the recipient but *before* the sentence final particle. Again there seems to be an asymmetry between the theme argument and recipient in their ability to occupy the right-dislocated position. Consider (12).

- (12) ??Ngo5 bei2 lei1 bun2 syu1 aal lei5  
 I give DET CL book PRT you  
 ‘I give this book, you’

Right-dislocating the recipient *lei5* in (12) sounds unnatural.

*Interaction with the Relative Clause Construction*

(13) and (14) show another pair of examples illustrating the theme-recipient asymmetry in displacement when the basic construction interacts with the relative clause construction. (13) shows the resultant non-basic surface form when relativizing the theme argument.

- (13) Ngo5 bei2 lei5 go2 bun2 syu1  
I give you DET CL book  
'The book I give you'

Relativizing the recipient argument does not sound very natural. Consider (14).

- (14) ??Ngo5 bei2 go2 bun2 syu1 ge3 jan4  
I give DET CL book MM person  
'The person I give the book'

### *Interaction with the zoengl construction*

As the semantics of dative constructions is associated with transfer of possessions, there is a good semantic match between the dative construction and the *zoengl*-construction in Cantonese<sup>3</sup>, where the *zoengl* construction “retains a sense of displacement, and in colloquial usage at least, is most typically used when the object of the sentence is literally moved from one place to another” (Yip and Matthews, 2001: 61). (15) shows the theme being in a displaced position before the verb *bei2* ‘give’ in a *zoengl* expression.

- (15) Ngo5 zoengl lei1 bun2 syu1 bei2 lei5 laa1  
He cause DET CL book give you PRT  
'I give this book to you.'

Displacing the recipient is ungrammatical, however. Consider (16).

---

<sup>3</sup> *Zoengl* in Cantonese and *bà* in Mandarin have been treated as close counterparts, they are broadly similar in function where they take a direct object and place it before the verb (see (28)), though the use of *zoengl* is more restricted (see Yip and Matthews, 2001: 61 for illustrative examples on this point). *Zoengl* has been treated as a causative verb in Cantonese.

- (16) \*Ngo5 zoeng1 lei5 bei2 lei1 bun2 syu1 laa1  
 I cause you give DET CL book PRT  
 'I give you to this book.'

Displacing the recipient is not totally impossible though.<sup>4</sup> Xu and Peyraube (1997), while also claiming that topicalizing the recipient is difficult to some speakers, nevertheless claimed that it is possible to 'topicalize' a recipient. The example they provided involved the verb *sung3* 'give(as a present)' and is repeated in (17) below.

- (17) Keoi5 ge3 sai3lou6, lei5 sung3 wun6geoi6 m4 hai6 sung3 syu1  
 He MM child you give toy not give book  
 'To his child, you gave toys, not books.'  
 (=Note 4 (i) in Xu and Peyraube 1997: 123)

In (17) above, the theme is situated in a contrastive context. Similarly in such a contrastive context, 'topicalizing' the recipient of a *bei2* dative may be acceptable. Consider (18).

- (18) Lei1 go3 gaau3sau6, ngo5 bei2 lei1 bun2 syu1 gaa3 m4hai6 bei2 go2 bun2  
 DET CL professor I give DET CL book SFP not give DET CL  
 'This professor, I give this book not that one.'

It also sounds natural to 'topicalize' the recipient with a topic marker *lei1* when the recipient is being contrasted with another 'topicalized' recipient in the following clause. Consider (19).

- (19) Lei5 lei1, ngo5 bei2 lei1 bun2 syu1 laa1, keoi3 lei1, ngo5 bei2 go2 bun2 laa1  
 You TOP I give DET CL book PRT, 3sg TOP I give DET CL PRT  
 'You, I give this book, she/he, I give that one.'

---

<sup>4</sup> Similarly in English, Polinsky (1998: 405) argued that the object asymmetries noted in double object constructions should not be treated as 'all-or-nothing effects'.



When the theme is questioned, it also sounds natural to ‘topicalize’ the recipient argument. Consider (20).

- (20) Go2 go3 saai3lou6, lei5 wui5 bei2 mat1je5 aa3?  
DET CL child you will give what SFP  
‘To that child, what will you give?’

Similarly for relativization, Xu and Peyraube (1997: 108), while also noting the difference in acceptability between relativizing the theme argument and the recipient argument, pointed out that it is in fact possible to relativize the recipient. The example they gave is repeated in (21) below.

- (21) Keoi5 bei2 luk6sap6 fan1 ge3 hok6saang1 hou2 gwo3  
He give sixty point MM student good than  
  
ngo5 bei2 baat3sap6 fan1 ge3 hok6saang1  
I give eighty point MM student  
‘Student whom he gave 60 points are better than those I gave 80 points.’  
(= Note 3(i) in Xu and Peyraube 1997: 123)

The authors therefore argued that the constraints on relativizing the objects of a dative construction are semantic and functional rather than structural (c.f. the Noun Phrase Accessibility Hierarchy (DO>IO) in Keenan and Comrie, 1977), though they did not suggest what the other possible motivations are in giving rise to such an asymmetry. Detailing the exact requirements for felicitous relativization and topicalization of the recipient argument is beyond the scope of this thesis, nevertheless it can be stated descriptively here that though relativizing or topicalizing the recipient does not sound very natural, it is acceptable in certain marked contexts. With regard to the asymmetry observed for the susceptibility to displacement of the theme and the recipient, here I point out that the asymmetry might involve functional

bases. Unlike inflectional languages<sup>5</sup>, Cantonese does not have case-markers or other inflectional endings to mark the grammatical relations or the thematic roles of the NPs, therefore, if the recipient argument is displaced, there can be ambiguity in deciding whether the animate NP is the agent or the recipient in these non-basic constructions<sup>6</sup>, especially in cases when one of these two arguments is unexpressed. Consider (22) where the Cantonese *bei2*-dative interacts with the *lin4...dou2* (even...also) construction in Cantonese.<sup>7</sup>

- (22) Lin4 keoi5 dou1 bei2 lei1 bun2 syu1 aa4  
 Even 3sg also give DET CL book PRT  
 ‘Even s/he also gives out this book!’  
 ‘Even to him/her also (the agent) gives this book!’

<sup>5</sup> For instance, the recipient (the indirect object) can be readily topicalized in German, where the subject and objects are all case-marked.

<sup>6</sup> Polinsky (1998: 404) also pointed out that the recipient is ‘an agentive participant other than the agent of transfer.’

<sup>7</sup> The agent, theme or recipient can be focused when the basic construction interacts with the *lin4...dou1* construction (see (i) to (iii) below). In this case, the outcome is that the focused argument (highlighted in capital letters below) comes between *lin4* and *dou1* and in front of the verb *bei2* ‘give’.

*The agent in focus*

- (i) Lin4 lei5 dou1 bei2 lei1 bun2 syu1 keoi5 aa4  
 Even you also give DET CL book 3sg PRT  
 ‘Even YOU give this book to him!’

*The theme in focus*

- (ii) Lei5 lin4 lei1 bun2 syu1 dou1 bei2 keoi5 aa4  
 You even DET CL book also give 3sg PRT  
 ‘You even give THIS BOOK to him!’

*The recipient in focus*

- (iii) Lei5 lin4 keoi5 dou1 bei2 lei1 bun2 syu1 aa4  
 You even 3sg also give DET CL book PRT  
 ‘You even give this book to HIM!’

Notice that in (ii) and (iii) above, when the theme and the recipient are focused, the surface ordering of the arguments differs considerably from that of the *bei2* construction- the theme and the recipient are in ‘displaced’ positions.

Here with an unexpressed argument, the utterance in (22) can be potentially ambiguous without contextual support, with either the third person pronoun being interpreted as the agent or the intended recipient.

On the other hand, the theme is mostly inanimate, so there is less potential ambiguity. I therefore hypothesize that the recipient of the Cantonese ditransitive constructions is not often displaced in the child-directed adult input in order to avoid concomitant semantic ambiguity in the thematic role assignment.<sup>8</sup>

### **3.4 A Corpus Study Of Adult Input**

This section presents empirical findings that support the two hypotheses on the Theme-Recipient asymmetry formulated in section 3.3. I look at the linguistic environment in which the [*bei2*-T-R] double object form is used in child-directed adult Cantonese, with special relevance to the realization and placement of the theme and the recipient. The findings in this chapter serve as the basis for strengthening the input frequency hypothesis (refer back to sections 2.3 and 2.4.3) and the two empirical hypotheses on the theme-recipient asymmetry discussed in section 3.3.

#### **3.4.1 Methodology for Adult Input Analysis**

So far there is no systematic Cantonese corpus data containing interactions between the child and her parent or caretaker, according to the author's knowledge. The closest we can get is corpus data involving interactions between the child and adult RAs. The present empirical findings on adult input were derived from analyses of

---

<sup>8</sup> One might speculate that Cantonese makes use of a number of strategies as cues for identification of the recipient in dative constructions, the possible cues are the *bei2* dative marker, the canonical ordering [*bei2*-T-R], restricted displacement of the recipient, and animacy.

adult child-directed speech in the Hong Kong Cantonese Child Language Corpus (CANCORP) (Lee et al., 1996). Concerning the procedures for data analysis, please refer to section 4.1.4.

For corpus data analysis, the CLAN computer program “kwal” (MacWhinney, 2000) was first used to generate a list of all the utterances (together with contextual data) containing the phonetic string *bei2* for every transcript in CANCORP. There were altogether 7146 lines containing the phonological form *bei2* in CANCORP. The utterances were sorted into speakers (adult, child, and elder siblings). See Table 3.1 below.

**Table 3.1. Utterances containing *bei2* in CANCORP**

	No. of lines containing the phonological form <i>bei2</i>
Adult	5885
Child	1086
Elder Siblings <sup>9</sup>	175
Total	7146

All the extracted *bei2* adult utterances were used for adult input analysis.

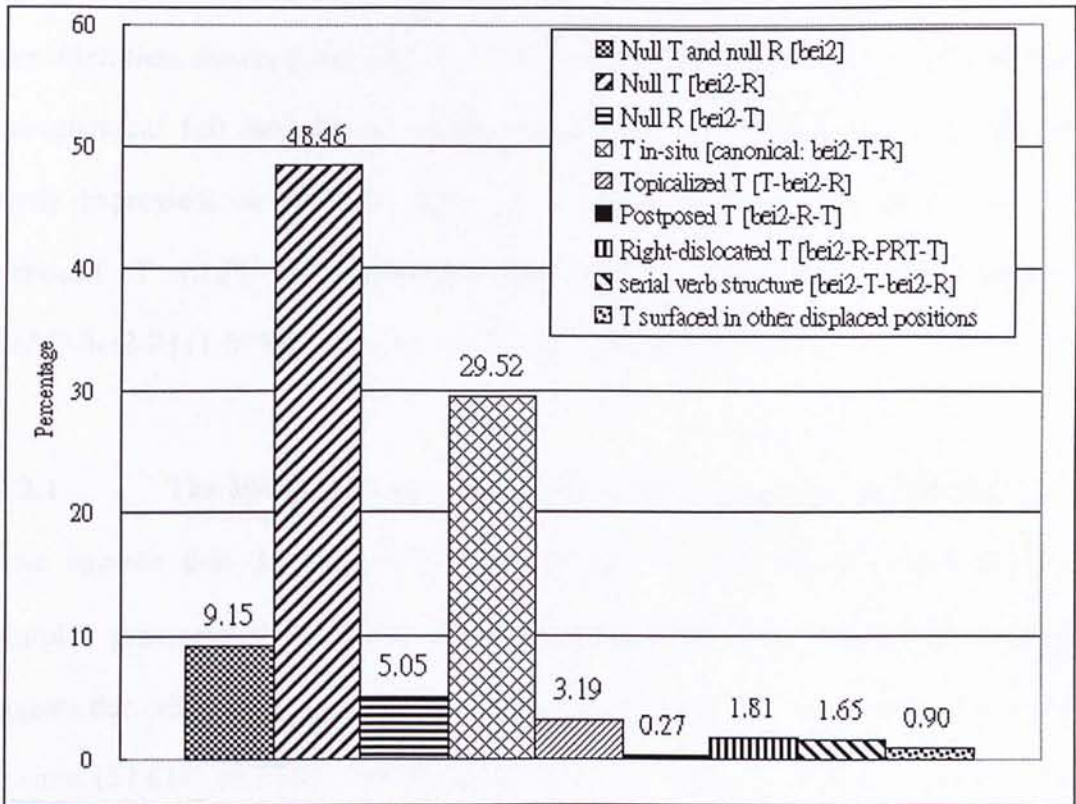
### **3.4.2 Corpus Findings**

There were a total of 1880 tokens of the target *bei2*-dative construction attested in CANCORP, comprising 32% of all the adult utterances containing the phonological form *bei2* in the corpus. Figure 3.1 shows the distribution of the different *bei2*-forms used by adults in child-directed speech in CANCORP.

---

<sup>9</sup> These utterances were from HHC’s elder sister who is seven years older than HHC, LLY’s elder sister who is four years older than LLY, LLY’s elder brother who is ten years older than LLY and LTF’s elder sister whose age was not provided in CANCORP.

**Figure 3.1. Adult use of Cantonese *bei2*-datives in CANCEP in child-directed speech  
(Total tokens: 1880)**



The presentation of the current empirical findings in figure 3.2 is organized with respect to the realization of the theme and the recipient and the placement of the theme. The placement of the recipient turns out to be not a variable here, because all the recipients, when expressed, were not displaced in the adult *bei2*-datives attested in our corpus. As described in section 3.3.2, the recipient is less susceptible to displacement than the theme: the recipient can be ‘displaced’ only in marked discourse contexts such as contrastive contexts in Cantonese.

Recall that the canonical form is [bei2-T-R] for the *bei2*-datives. Of these 1880 tokens, only 29.52% were in the canonical [bei2-T-R] form. Non-full datives<sup>10</sup> constitute 62.66% of all the *bei2*-datives used by adults when speaking to young

<sup>10</sup> They include *bei2*-datives with either T or R unexpressed or *bei2* ‘give’ alone (with or without the agent expressed).

children. The whole paradigm of surface forms is quantitatively headed by *bei2*-datives with null themes (48.46%). The theme (T) was unexpressed four times more often than the recipient (R): T (57.61% unexpressed), R (14.2% unexpressed). Non-canonical full *bei2*-forms, which involve both the theme and the recipient overtly expressed, on the other hand, were rarely used: topicalized T (3.19 %), postposed T (0.27 %), right-dislocated T (1.81%), serial verb structure [*bei2*-T-*bei2*-R] (1.65%), T in other displaced positions (0.90%).

#### 3.4.2.1 The Missing Theme: *bei2*-Datives with Frequent Null Theme

If we assume that the present data constitutes a representative distribution of examples presented to a young child acquiring Cantonese, the present finding suggests that when children hear a *bei2*-dative from the adult input, more than half of the time (57.61% (9.15% + 48.46%)) the theme is absent in surface syntax. For 48.46% of the time, adults tend to background the theme (the entity being transferred) when it can be inferred/ identified from the immediate context or/and the previous discourse. It is thus likely that the abundant use of the *bei2*-dative forms with the null theme is a feature of child-directed speech in Cantonese<sup>11</sup> as adult speech to young children is likely confined to the here-and-now context where the referent of the theme is available in the immediate context.

In addition, we notice that adults sometimes mention T first to establish a shared topic with the child and then predicate something about the topic using a *bei2*-dative. With such a structuring of information, there is no pragmatic motivation to specify the theme argument for the child when using the *bei2*-dative, when it has already

---

<sup>11</sup> Whether this speculation is correct has yet to be verified by doing a systematic comparison with the adult-directed speech.

been mentioned and established in the previous discourse. It is also likely that adults tend to use simplified speech when speaking to young children, so they would not express an argument unless there is a pragmatic motivation to do so. Here adults might be conforming to the application of the Grice's second maxim of Quantity: "do not make your contribution more informative than is required" (Grice, 1975) especially in child-directed speech.

### 3.4.2.2 The Non-Canonical [*bei2-R-T*] Form

The paucity of the use of [*bei2-R-T*] form appears to be a feature of child-directed speech in Cantonese. Suggestive evidence comes from comparing the use of the [*bei2-R-T*] form in CANCELP with that in adult-to-adult speech in an adult Cantonese corpus (Law, Fung and Leung, 2002).<sup>12</sup> In CANCELP, the [*bei2-R-T*] form is used by adults only 0.27% of the time, whereas in the adult spoken corpus, the [*bei2-R-T*] form is used 9.33% of the time (7 tokens out of an overall 75 tokens of basic *bei2*-datives attested). Six out of the seven [*bei2-R-T*] examples attested in the adult-to-adult Cantonese spoken corpus involve a long theme with at least four syllables. The remaining [*bei2-R-T*] example involves an abstract theme argumental NP. (21) and (22) show two representative examples of the [*bei2-R-T*] where T is long in the adult corpus.

---

<sup>12</sup> The Hong Kong Cantonese Adult Corpus (HKCAC) developed by Law, Fung and Leung (2002) contains orthographic and phonetic transcriptions of more than 8 hours of spoken adult-to-adult Cantonese. The recordings come from phone-in programs and forums on the radio in Hong Kong in which the interactions between 69 adult Cantonese-speaking speakers in addition to the program hosts were recorded during November 1998 to February 2000.

(21) Keoi5 bei2 ngo5 hou2 do1 ci3 gei1 wui6  
3sg give me many CL chance  
'S/he gives me many chances.'

(22) Bei2 hoeng1 gong2 din6 seon3 jat1 go3 ge3 hai6 e3 lung5 dyun6 kyun4  
Give HongKong Telecom one CL GEN Vf PRT monopoly  
'Give HongKong Telecom a monopoly.'

It should also be highlighted here that the recordings of the Hong Kong Cantonese Adult Corpus (HKCAC) came from interactions between 69 adult Cantonese-speaking speakers and the program hosts in phone-in programs and forums on the radio in Hong Kong. The radio programs have themes ranging from political issues, economic issues, current affairs, personal issues to ghost stories. With the lack of face-to-face interactions and hence immediate contextual support in such interactions, modifications of the NP referents might be necessary for identification, particularly on specific themes and topics, giving rise to the use of longer NPs. It might be that Cantonese-speaking adults use the *bei2*-dative more often in the marked [V-R-T] form in adult-directed speech.

### 3.4.2.3 The Non-Canonical [*bei2*-T-*bei2*-R] Form

The non-canonical [*bei2*-T-*bei2*-R] serial verb form is rarely used in Cantonese adult child-directed speech. Only 1.65% of all the *bei2*-datives are found to be in this form.

### 3.4.2.4 The Frequent [*bei2*-R] Sequence

The present findings on the adult *bei2*-datives show that in 56.28% of the cases (1058 out of 1880 tokens) the recipient rather than the theme surfaces as the first post-*bei2* argument because of unexpressed or displaced theme. Moreover, there are other distinct but related serial verb constructions with *bei2* as a dative marker



that further instantiate the *bei2*-R sequence. To the extent that these distinct but related *bei2* constructions also form part of the input data to the child, I consider the possibility that these related constructions might also have a potential influence in affecting the early schematization of the [*bei2*-T-R] canonical form. To illustrate the significance of these related serial verb constructions, in the present adult input analysis using CANCORP, apart from the non-canonical form [*bei2*-T-*bei2*-R] which can be regarded as an instance of the [V-T-*bei2*-R] serial verb schema, there were 922 more instances of this serial verb schema with other verbs. Out of these 922 instances, all had the recipient appearing after the dative marker *bei2*, whether the theme was unexpressed or expressed, undisplaced or displaced. 895 instances involved the [*bei2*-R] string placed at the salient utterance final position (see Slobin, 1985).<sup>13</sup> Together with the 1058 tokens which instantiated a [*bei2*-R] sequence in figure 3.1, there were altogether 1980 tokens instantiating a [*bei2*-R] sequence which were more frequently attested than the canonical [*bei2*-T-R] form (555 tokens) in the current adult input findings.

### **3.5 Cantonese Adult Input Properties: Implications for Early Schematization**

In section 3.4, I have highlighted the contrast between how the theme and the recipient of the *bei2*-datives is actually instantiated in adult Cantonese child-directed speech: the theme is frequently unexpressed or sometimes displaced, while the recipient is frequently expressed and undisplaced, and frequently occurs with *bei2*. I would like to emphasize that the quantitative bias in the present adult input findings

---

<sup>13</sup> [*bei2*-R] with or without sentence final particle(s) following it is regarded as appearing at the salient sentence final position here. On the other hand, if there are lexical items following [*bei2*-R], for instance, *hou2-m4-hou2* 'Lit. OK?' or *sin1* 'first', the [*bei2*-R] sequence is not regarded as appearing at the utterance final position.

has support from the characteristics of Cantonese grammar (refer back to section 3.3), so is unlikely an artifact due to sampling bias.

The two empirical hypotheses raised in section 3.3, where the theme tends to be more often unexpressed and displaced than the recipient for the Cantonese *bei2*-datives in adult Cantonese discourse, are strongly corroborated by the quantitative study on the adult input properties reported in this chapter. The present corpus findings on the use of *bei2*-datives in the adult input show that the theme was unexpressed four times more often than the recipient: theme (57.61% unexpressed), recipient (14.2% unexpressed)<sup>14</sup>, and the theme, though displaced only 6.17% of the time (topicalized, postposed, right-dislocated or displaced to other positions), was displaced more often than the recipient, which when expressed, was *always* undisplaced.

The present findings also show that in actual *usage events*, despite the different surface form instantiations, the regularities young children encounter from their adult input are the highly frequent [*bei2*-R], most of which occur at the salient utterance-final position, because of the theme is unexpressed or displaced, or when the verb *bei2* ‘give’ functions as a dative marker in serial verb expressions.

---

<sup>14</sup> Catherine Demuth (2000), working on the acquisition of the Sesotho applicative construction, also found that the themes are far more likely to be elided than the recipients in the use of applicatives in child-directed adult speech. More cross-linguistic support for the idea that the theme tends to be elided more often than the recipient for ditransitive constructions has yet to be further established. Many thanks to Prof Laura Michaelis for pointing out Demuth’s work to me.

The Cantonese input properties do not seem to facilitate early abstraction. Given that schematization involves finding, extracting and generalizing across patterns out of concrete instantiations, one needs (at least) repeated exposure of the instantiating expressions concerned. Frequent exposure to the instantiating expressions is one factor that would facilitate early schematization. However, the canonical [*bei2*-T-R] form, attested only 29.52% of the time in the adult input as far as the basic *bei2*-datives are concerned, is situated in the midst of other ‘competing’ expressions where the recipient surfaces as the first post *bei2* argument when the theme is unexpressed, displaced, or when *bei2* ‘give’ functions as a dative marker in the serial verb expressions. The theme is absent in the speech stream 57.61% of the time. Even when it is overtly expressed, it occurs in more than one position in surface syntax, since it may be displaced (refer back to section 3.3.2 for the discussion on the theme-recipient asymmetry in argument displacement, and section 3.4.2 on corroborative evidence from the adult input corpus findings). Taking these two facts into consideration, it is reasonable to conclude that the Cantonese input properties at least do not seem to be structured in a way that facilitates the early schematization of the canonical [*bei2*-T-R] schema for a young child, based on the input properties hypothesis discussed in (2).

The adult input findings also substantiate the empirical predictions for the input frequency hypothesis formulated in section 2.4.3. The input frequency hypothesis predicts that the canonical form is acquired earlier than the non-canonical forms, which are infrequent in adult input.

### 3.6 Chapter Summary

This chapter pays attention to how the language specific properties of Cantonese might impact the early schematization of the canonical [*bei2*-T-R] double object form from the adult input from a usage-based perspective. I have discussed the input properties hypothesis in section 3.2. In section 3.3, I have also formulated two empirical hypotheses on the asymmetry of the theme and the recipient in argument realization and displacement. First, the theme is more often unexpressed than the recipient in Cantonese adult input data. Second, the theme is more often displaced than the recipient. In section 3.4, I have also done a corpus study of adult child-directed speech to substantiate my two hypotheses. The findings support the two empirical hypotheses mentioned above. The findings are attributed to the fact that Cantonese grammar is structured in a way that gives rise to frequency bias in the input, rather than a matter of sampling bias. Based on the theme-recipient asymmetry in argument realization and displacement for the Cantonese *bei2*-dative as reported above, the adult input properties in Cantonese do not appear to facilitate the early abstraction of the canonical [*bei2*-T-R] form from a usage-based perspective. Building on the two empirical hypotheses related to the asymmetry of the theme versus the recipient in argument realization and susceptibility to displacement substantiated by the adult input corpus findings, I have hypothesized that the input *properties* do not facilitate early schematization of the canonical [*bei2*-T-R] form. The adult input corpus findings also show that the non-canonical *bei2*-forms are much more infrequent than the canonical form, this finding substantiates the empirical prediction I have formulated for the input frequency hypothesis discussed in section 2.4.3, whereby it predicts preference of the canonical [*bei2*-T-R] form over the non-canonical [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms.

## **Chapter Four. Methodology and Early Developmental Findings**

### **4.0 Introduction**

This chapter reports the developmental findings of the study. Section 4.1 first describes the methodology. Sections 4.2 reports the early developmental findings. The main findings are derived from analyses of naturalistic speech of eight children in a monolingual Cantonese child language corpus (Lee et al., 1996). Bilingual and clinical child data would also be included when discussing the early non-target *bei2*-datives (see section 4.2.6). Section 4.3 reports the usage patterns noted in older monolingual children. Section 4.4 summarizes the major findings.

### **4.1 Methodology**

This section provides a detailed description of the nature and source of data, general background and linguistic input of child subjects and procedures for data analysis. The present study involves mainly the analysis of longitudinal developmental corpus data of eight monolingual Cantonese children and three Cantonese-English bilingual children. The naturalistic data used in this study is of two types- longitudinal spoken corpus data of naturalistic adult-to-child interactions in monolingual and bilingual contexts (see section 4.1.1); ii) diary data of a bilingual Cantonese-English child provided by Cheung (see section 4.1.2); and iii) clinical data collected by local speech therapists (see section 4.1.3).

#### **4.1.1 Longitudinal Corpus Data**

The *longitudinal* data came from two released naturalistic child language corpora: The Hong Kong Cantonese Child Language Corpus (henceforth CANCORP) (Lee, Wong and Leung, 1996) and The Hong Kong Bilingual Child Language Corpus (Yip, Matthews and Huang, 2001), deposited at the CHILDES archive.<sup>1</sup> Both corpora consist of transcripts of naturalistic adult-to-child interactions that grew out of different child language research projects in Hong Kong. The monolingual project studies eight monolingual Cantonese children between the ages of approximately 1;05 to 3;08 and the bilingual project studies five Cantonese-English bilingual subjects between the ages of approximately 1;05 to 4;06.

##### **4.1.1.1 Monolingual Child Data: The Hong Kong Cantonese Child Language Corpus (CANCORP)**

The monolingual developmental data came from CANCORP (Lee et al, 1994). There were four boys and four girls, ranging in age from 1;05 to 2;08 when the study began. Each child was observed for around one year, generating a total of 171 longitudinal data files from 1;05 to 3;08 upon completion of the study. Table 4.1 shows the age span and the number of data files created for each child during the longitudinal study.

---

<sup>1</sup> The CHILDES (Child Language Data Exchange System) project (MacWhinney, B. 2000) aims to promote sharing and exchange of cross-linguistic data for language acquisition research. The system provides tools for studying conversational interactions. These tools include a database of transcripts, a standardized format for transcription and methods for linguistic coding (known as CHAT: Codes for the Human Analysis of Transcripts), systems for linking transcripts to digitized audio and video files and programs such as CLAN (Child Language Tools for Analyzing Talk) for analyzing the transcripts.

The CHILDES website is at <http://chilides.psy.cmu.edu> based at the Department of Psychology of Carnegie Mellon University.

**Table 4.1. Age span and the number of data files created in CANCORP**

Child Subject	Sex	Age range of longitudinal developmental data	No. of data files
CCC	M	1;10;08 - 2;10;27	22
CKT	M	1;05;22 - 2;07;22	25
CGK	F	1;11;01 - 2;09;09	19
HHC	M	2;04;08 - 3;04;14	16
LTF	F	2;02;10 - 3;02;18	16
LLY	F	2;08;10 - 3;08;09	20
MHZ	M	1;07;00 - 2;08;06	26
WBH	F	2;03;23 - 3;04;08	27

Each child was visited approximately twice per month mostly at the child's home.<sup>2</sup> The average sampling time was one hour during which the interactions between the child subject, the adult investigator(s), the parent(s), the sibling(s), the other family members and people present at the recording session were audio taped. Activities included reading storybooks, playing with toys and eating snacks. All the recorded utterances were then transcribed and coded according to the CHAT (Codes for the Human Analysis of Transcripts) format used in the CHILDES project and tagged with 33 parts-of-speech labels.

All the monolingual child subjects and their families live in Hong Kong, where Cantonese is the language used in the local community. Table 4.2 summarizes the background information of each monolingual child in CANCORP.

---

<sup>2</sup> Recording was not done at the child's home for the following sampling points:

CCC at 1;10.08 : At the office of the school where the child's mother teaches  
 at 1;11.00, 2;01.17, 2;07.06: At a park  
 at 2;02.13 : At a café  
 at 2;10.27 : At the Chinese University of Hong Kong  
 MHZ at 2;06.18 : At the Chinese University of Hong Kong

**Table 4.2. Background information of the eight monolingual Cantonese children in CANCORP**

Name	Language(s) used at home	Caretaker(s)	Sibling(s)	Parents' occupation
CCC	Cantonese	Maternal grandparents & Mother	--	F: Merchant M: Secondary English teacher
CGK	Cantonese	Mother	--	F: Electronic Technician M: Housewife
CKT	Cantonese	Weekdays: Maternal grandmother Weekend: Parents	--	Unknown
HHC	Cantonese, Thai helper also speaks Cantonese to the children	Thai helper & Mother	1 elder sister (age 7)	F: Engineer M: Typist
LLY	Cantonese, Filipino helper speaks some Cantonese and English to the children	Mother & Filipino helper	1 elder brother (age 10) & 1 elder sister (age 4)	F: Businessman M: Housewife
LTF	Cantonese except when speaking to the Filipino helper	Birth - 1;06: Filipino helper Since 1;06: Mother	1 elder sister	F: Unknown M: School music teacher (until the child reached 1;06)
MHZ	Cantonese, parents occasionally introduce some English terms to the child	Birth - 1;01: Maternal grandmother 1;01- 2;06: Caretaker (weekdays)	--	Unknown
WBH	Cantonese	Daytime: Attends Nursery school After school: Parents and grandmother	1 younger brother (2 years younger)	Unknown

#### 4.1.1.2 Cantonese-English Bilingual Child Data: The Hong Kong Bilingual Child Language Corpus

The bilingual child data came from the Hong Kong Bilingual Child Language Corpus (Yip, Matthews and Huang, 2001). All the five Cantonese-English bilingual subjects were exposed to the two languages from birth. Developmental data from the siblings Timmy and Sophie, and another bilingual child Kathryn were included in the present study because their datafiles had been released at the time of conducting this study. Longitudinal data of Sophie, Timmy and Kathryn's language development were each collected separately when Sophie was 1;06 to 4, and Timmy 1;10 to 3;06.25, and Kathryn 2;09.23 and 4;06.07 at approximately biweekly time



intervals. Spontaneous speech data of each child was recorded at the children's home where the child naturally interacted with two investigators and members of the family in routines such as role-playing, playing with toys and reading storybooks. One language was elicited at a time, for example, in the first half hour of recording, English was spoken by one research assistant (RA) in order to elicit English, while the other RA used Cantonese in the second half hour to elicit Cantonese. The RAs who interacted with Sophie, Timmy and Kathryn are all native speakers of Cantonese except one who is a native speaker of Mandarin and used primarily English in the recording sessions. The two RAs speak English as their second language. Sophie's language development was videotaped whereas Timmy and Kathryn only had audio recording. Recording of the adult-child interaction in each language was then transcribed and coded according to the CHAT format by the research assistants, and divided separately into English and Cantonese data files for each child. Table 4.3 shows the age span and the number of Cantonese and English data files created and released for Sophie, Timmy and Kathryn at the time of conducting the present study. Note that this study only analyzed the Cantonese files.

**Table 4.3. Age span and the number of Cantonese and English data files created in the bilingual corpus**

<b>Child Subject</b>	<b>Sex</b>	<b>Age range of longitudinal developmental data</b>	<b>No. of Cantonese datafiles</b>	<b>No. of English datafiles</b>
Timmy	M	1;05.20 - 3;06.25	34	38
Sophie	F	1;06.00 - 4;00.00	40	40
Kathryn	F	3;06.18 - 4;06.07	17	14

All the bilingual child subjects and their families live in Hong Kong. Table 4.4 summarizes the background information of each bilingual Cantonese-English child.

**Table 4.4. Background information of the three bilingual Cantonese-English children**

Name	Language(s) used at home	Caretaker(s)	Sibling(s)	Parents' occupation
Timmy	Cantonese from mother (native speaker of Hong Kong Cantonese) and maternal grandma English from father	mother, maternal grandma and Filipino maid	two younger sisters (Sophie, two years and nine months younger; Alicia, seven years younger)	F & M: linguistics professors in Hong Kong
Sophie	(native speaker of British English) and Filipino maid Some Chiu Chow from maternal grandma		one elder brother (Timmy, two years and nine months older), one younger sister (Alicia, four years and three months younger)	
Kathryn	Cantonese from father (native speaker of Cantonese), a part-time Cantonese cleaner <sup>3</sup> and school <sup>4</sup> English from mother (native speaker of British English), Filipino domestic helper (until Kathryn was around age 3)	Mother	one elder brother, four years and eight months older; and Alasdair, one year and nine months older	F: neuro-surgeon at a university hospital, M: housewife

#### 4.1.2 Cantonese-English Bilingual Diary Data: Cheung (2002, p.c.)

Sik Lee Cheung, the mother of a bilingual child, Siu Bou, started audio taping her child and taking notes of his vocalization and utterances when the child was 2 months old. The taping was most frequent when the child was between 2;2 and 2;11. There were over 20 tapes from that period. Before the child was two, the mother wrote down her observation of the child's utterances as much as she could. When the child's utterances became more varied and complex, the mother could only jot down the new expressions or complex utterances. In addition, the mother also recorded utterances that were not adult-like in both Cantonese and English.

<sup>3</sup> The part-time Cantonese cleaner can speak fluent English.

<sup>4</sup> Kathryn attended the Cantonese section of an international Kindergarten from age 2;07.

The bilingual child Siu Bou was born on Oct. 15, 1998 in California and has lived in the same house in California since he was born. His mother was born and raised in Hong Kong and is a native Cantonese speaker. His father is of European decent and was born and raised in the US. His mother is a trained child language researcher and his father is an engineer/scientist and has a doctoral degree in engineering. Concerning linguistic input, the child's father understands very little Cantonese and speaks only English to Siu Bou. The mother spoke Cantonese exclusively to Siu Bou till he was about 2;0, but then the conversation between the child and his mother became so complicated that the child's father could not understand what they were saying, so the mother started to switch to English whenever she wanted to involve the father. Siu Bou's dominant language was Cantonese till he started to attend preschool at 2;10. He was at the preschool two afternoons a week in the first year and three afternoons a week in the second year. According to his mother, his English started to become dominant when he started going to preschool three times a week at 3;10 and meeting his friends for play outside school.

#### **4.1.3 Clinical Child Data: Local Speech Therapists in Hong Kong**

Clinical data was investigated to see whether early non-target datives can be found among the clinical developmental population acquiring Cantonese. Some local speech therapists reported that they noted the unconventional usage of word order with the Cantonese datives among the clinical child population they serve, though this phenomenon has never been documented and attended to in the field of acquisition research.

The speech therapists were asked to note down the actual child's utterance (with relevant contextual data if possible) when s/he heard the child producing it and when

s/he noted prior recording of such usage from her/his existing files. The following information was requested: initials of the child's name, medical diagnosis, chronological age, mental age (if available), language age (if available), and the child's actual utterances spoken at a particular chronological age (with relevant contextual data if possible). The utterances relevant to this thesis were presented in section 4.2.6.<sup>5</sup>

The seven children's background information is shown in Table 4.5 below. All of them are monolingual Cantonese children having language delays as indicated by the discrepancies between their language ages obtained at the date of language assessment<sup>6</sup> and the chronological ages at which the assessment was carried out.

---

<sup>5</sup> The scope of clinical data investigation included not only the Cantonese *bei2* dative but also datives involving other Cantonese verbs and ditransitive structures involving verbs such as *baai2* 'put' for future analysis.

<sup>6</sup> The child's language was assessed by the standardized Reynell Developmental Language Scales (Cantonese Version) by the speech therapist. If standardized assessments cannot be conducted, the language ages were provided by the speech therapists based on informal language assessments.

**Table 4.5. Background information of the seven children with language impairments**

Child	Medical Diagnosis	Mental Age	Language Age
LKM	Mild mental retardation	4;00 (assessed at C.A.: 6;06)	3;06 (assessed at C.A.: 8;03)
KTY	Borderline delay in intelligence quotient (IQ)	1;06 (assessed at C.A.: 1;09)	1;00 (assessed at C.A.: 1;09)
LMY	Pervasive Developmental Delay (PDD)	1;06-2;00 (assessed at C.A.: 2;00)	V.C.~2;08, V.E.~2;06 (assessed at C.A.: 4;07)
CMM	Spastic quadriplegic cerebral palsy, Mild grade mental retardation	1;07-1;08 (assessed at C.A.: 3;00)	V.C.~2;11 V.E.~2;07 (assessed at C.A.: ?)
FMH	Borderline delay with autistic features	1;08-1;09 (assessed at C.A.: 1;11.28)	V.C.:1;00 V.E.:1;00 (assessed at C.A.: 1;11.28)
TSW	Spastic cerebral palsy, Moderate grade mental retardation	moderate grade mental retardation	3;04 (assessed at C.A.: 9;07.11) 2;07 (assessed at C.A.: 9;07.11)
CYM	Right cerebral palsy, High moderate to low mild grade mental retardation, Global developmental delay, Dyspraxia, History of head injury in 1998	high moderate to low mild grade mental retardation	V.C.: mild delay (able to handle conversational speech) V.E.: dyspraxia, stuttering, short and simple sentences limited to one main verb for each sentence (cannot produce serial verb sentences)

Keys: C.A.- Chronological Age, V.C.-Verbal Comprehension, V.E.-Verbal Expression

#### 4.1.4 Procedures for Data Analysis

All the extracted *bei2* child utterances in CANCORP (see table 3.1) were examined manually for disambiguation of the *bei2* homophonous forms. Utterances containing unintelligible words (coded as xxx or yyy) and incomplete utterances (marked by +/. at the end of the utterance) were also included at this stage. Instead I paid attention to where the unintelligible strings are identified and how incomplete the utterance was. Direct or partial imitations of the preceding adult utterance were coded separately.

With respect to the present research focus, the realization, placement and ordering of the theme and the recipient arguments (if they were overtly expressed) were only noted. All the child and adult *bei2*-dative utterances were then parsed into argument structure frames. The parsing scheme is as follows. The pre-verbal and post-verbal argumental NPs involved were coded according to their thematic roles in the transfer event: (R)-Recipient, (T)- Theme, though we hold no assumptions that children of this age are operating with these abstract thematic notions (c.f. Lieven and Pine, 1997; Tomasello, 2000a; Tomasello, 2000b for discussions on this issue).<sup>7</sup> These parsing and coding procedures were also applied to the bilingual corpus data, the bilingual diary data and the clinical data.

The coded utterances were then analyzed and tabulated for frequency counts (tokens). In addition, for the child utterances, I have also examined the location of pausing, the substrings that the child repeated within an utterance, the location of interjection particles as fillers, the particular lexical items associated, the relationship between the child's use of a form and the immediately preceding forms, i.e. the prior discourse context contributed by either the child or/and the adult interlocutor(s), and the child's previous usage with *bei2* 'give'. Since the present study works mostly with naturalistic discourse corpus data, the felicity of the child's usage of a particular form in a particular discourse context was also evaluated (see section 4.2.6 on early non-target forms).

---

<sup>7</sup> Tomasello (2000a) argued that children operate with different psycholinguistic units from adults.

## 4.2 Early Developmental Findings

### 4.2.1 Non-Full *bei2*-Datives Before Full *bei2*-Datives

All the monolingual children in CANCORP use non-full *bei2* datives before their first spontaneous use of a full *bei2* dative. At this early stage, they use non-full *bei2*-datives even when the discourse context requires a full *bei2*-dative. Even when the adult investigator make use of questions in the canonical [*bei2*-T<sub>wh</sub>-R] form for clarification or for scaffolding, the children fail to make use of the adult's [*bei2*-T<sub>wh</sub>-R] question to fill the WH-slot and produce a canonical [*bei2*-T-R] utterance. (1) shows an illustrative example from MHZ when he was 2;00.03.

(1) MHZ, at 2;00.03, lines 3270 to 3302

\*CHI: "A" ze4ze1.

'Sister A'

\*INV: Mat1je5 aal ?

'What?'

\*CHI: Bei2 hou6zeon1, bei2 # bei2 # bei2 hou6zeon1.

Give hou6zeon1 give give give hou6zeon1

'Give *hou5zeon1*(the child's name), give # give # give *hou6zeon1* (the child's name).'

\*INV: Bei2 mat1 je5 aal ?

'Give what?'

\*CHI: Bei2.

'Give'

\*INV: Bei2 mat1je5 hou6zeon1 aal ?

Give what hou6zeon1 PRT

'Give what to hou6zeon1?'

\*CHI: Bei2 hou6zeon1.

'Give *hou6zeon1*.'

\*INV: Bei2 mat1je5 hou6zeon1 aal ?

Give what hou6zeon1 PRT

'Give what to hou6zeon1?'

\*CHI: Bei2 # bei2 nei1 di1.

Give give DET CL

'Give these.'

- \*INV: Bei2 matlje5 aa1 ?  
Give what PRT  
'Give what?'
- \*CHI: Bei2 nei1di1.  
'Give these.'

Table 4.6 on the next page shows the first use of [*bei2-T*] versus [*bei2-R*] attested in CANCORP.



Table 4.6. The first use of [*bei2-T*] and [*bei2-R*] non-full *bei2*-datives attested in CANCORP

Child	Before 1;06	1;06-1;09	1;09-2;00	2;00-2;03	2;03-2;06	2;06-2;09	2;09-3;00	3;00-3;03	3;03-3;06	3;06-3;09
CKT	<i>bei2-R</i> (1;05.22)			<i>bei2-T</i> (2;02.15)						
MHZ			<i>bei2-R</i> (1;09.04) <i>bei2-T</i> (1;10.10)							
CCC					<i>bei2-R</i> (2;05.23) <i>bei2-T</i> (2;07.06)					
CGK			<i>bei2-R</i> (1;11.01)	<i>bei2-T</i> (2;00.08)						
LTF					<i>bei2-R</i> (2;03.02)	<i>bei2-T</i> (2;07.20)				
WBH					<i>bei2-R</i> (2;05.06)				<i>bei2-T</i> (3;03.12)	
HHC					<i>bei2-T</i> (2;04.08) <i>bei2-R</i> (2;05.03)					
LLY							<i>bei2-R</i> (2;11.01)			

The shaded area indicates the age span across which the child's developmental data is available in the corpus.

As shown in Table 4.6, [bei2-R] emerges earlier than [bei2-T] for each child with the exception of HHC. However, HHC's first use of [bei2-T] attested in the corpus is a partial imitation of his elder sister's utterance and all his subsequent uses of [bei2-T] as attested in the corpus before his first spontaneous use of [bei2-R] at 2;05.13 are confined to *bei2cin2* 'lit. give money'.

Table 4.7 shows all the uses of [bei2-T] and [bei2-R] attested in CANCORP before children's first spontaneous use of a full *bei2*-dative.

Table 4.7. All the uses of [bei2-T] and [bei2-R] non-full *bei2*-datives attested in CANCORP before children's first spontaneous use of a full *bei2*-dative (numbers in parentheses indicate token measures)

Child	[bei2-T]		[bei2-R]	
	Token	Type	Token	Type
CKT	No spontaneous use of a full <i>bei2</i> -dative is attested.			
MHZ	5	3 <i>cin2</i> 'money' (1) <i>bolbol</i> 'ball' (2) <i>neildil</i> 'these' (2)	48	7 <i>hou6zoen1</i> 'child's name' (25) <i>ze4ze1</i> 'sister' (8) <i>lei5</i> 'you' (2) <i>A ze4ze1</i> 'sister A' (8) <i>maalmaal</i> 'mother' (2) <i>je4je4</i> 'grandpa' (1) <i>ngo5</i> 'me' (2)
CCC	4	1 <i>cin2</i> 'money' (4)	2	1 <i>zeon3zeon3</i> 'child's name' (2)
CGK	2	2 <i>cin2</i> 'money' (1)	10	6 <i>ze4ze1</i> 'sister' (1) <i>ngo5</i> 'me' (2) <i>gaalkei4</i> 'child's name' (4) <i>hung4zai2maau1</i> 'panda' (1) <i>lei5</i> 'you' (1) <i>suk1suk1</i> 'uncle' (1)
LTF	0	0	6	2 <i>maalmi4</i> 'mother' (5) <i>ngo5</i> 'me' (1)
WBH			6	1 <i>ngo5</i> 'me' (6)
HHC	5	2 <i>jat1baa2</i> 'one CL' (1) <sup>8</sup> <i>cin2</i> 'money' (4)	2	1 <i>lei5</i> 'you' (2)
LLY	0	0	1	1 <i>ngo5</i> 'me' (1)

Table 4.7 shows that as far as the present corpus data is concerned, apart from MHZ, the other children's early uses of [bei2-T] are confined to the expression *bei2cin2* 'lit. give money' before they spontaneously use a full *bei2*-dative. MHZ and CGK demonstrate productive use of the [bei2-R] non-full datives with different referents

<sup>8</sup> This is a partial imitation of his sister's prior utterance.

playing the recipient role.

#### 4.2.2 The First Spontaneous Use of Full *bei2*-Datives

Table 4.8 shows the first spontaneous use of a *full bei2-dative* by the eight monolingual children attested in CANCORP. I use neutral terms *preposed* and *postposed* to describe the early placement of the theme relative to its undisplaced position in the canonical [*bei2*-T-R] form.

Table 4.8. Monolingual children's first spontaneous use of a full dative with the verb *bei2* in CANCORP

Child	Age and line number	Utterance	Note
CKT	no instance found		
MHZ	2;03.09: line 1257	Baa1baa1 bei2 hung1sung1beng2 Daddy give muffin  bei2 hou6zeon1. give hou6zeon  'Daddy give a muffin to <i>hou6zeon1</i> (the child's name).'	[ <i>bei2-T-bei2-R</i> ]
CCC	2;08.17: line 49	Ji1sang1 a3 bei2 zo2 jan4 aa3 . Doctor PRT, give ASP people PRT 'Doctor, have given to people.' ( <i>ji1sang1</i> refers to the child's toy doctor kit)	Preposed T
CGK	2;03.04: line 1373	Bei2 cin2 bei2 lei5 aa1 . Give money give you PRT 'Give money to you.'	[ <i>bei2-T-bei2-R</i> ]
LTF	2:03.30: line 2851 <sup>9</sup>	Ngo5 xxx bei2 ngo5 nei1 go3 aa3. Me give me DET CL PRT 'Me xxx give me this one.'	Postposed T
WBH	2;09.19: line 1307 <sup>10</sup>	A zi2 bei2 ngo5 laa1. PRT paper give me PRT 'Paper give me.'	Preposed T
HHC	2;06.10: line 2651	Paa2ce1 bei2 lei5 . Sport car give you 'Sport car give you.'	Preposed T
LLY	2:11.01: line 2117	Bei2 aa3saa1 # sing1sing1 aa1 . Give <i>aa3saa1</i> star PRT 'Give <i>aa3saa1</i> star.'	Postposed T

Recall that the canonical form for a *full bei2*-dative in Cantonese is [*bei2-T-R*].

None of the present monolingual child subjects, however, use this canonical form when they first start to express both the theme and the recipient. From the adult

<sup>9</sup> The child uses this form twice consecutively. The child's next utterance is:  
Bei2 ngo5 nei1 go3 aa3 (2:03.30: line 2854) [*bei2-R-T*]  
'Give me this one.'

<sup>10</sup> The child uses this form three times consecutively. The child's next utterance is:  
A zi2 bei2 ngo5 a zi2 bei2 ngo5 aa1 (2;09.19: line 1310) [*T-bei2-R*]  
'Paper give me paper give me.'

language perspective, these children instead use the non-canonical forms: the [T-*bei2*-R] form with preposed theme in CCC, HHC and WBH; the [*bei2*-R-T] form with postposed T in LTF, LLY; and the serial verb form [*bei2*-T-*bei2*-R] in CGK and MHZ (see sections 1.1.2 and 1.1.3 for discussion of these non-canonical forms in adult Cantonese).

It should also be noted that the adult input findings reported in section 3.4, by contrast, show that the non-canonical *bei2*-forms are rarely used in general and are even used much less frequently than the canonical [*bei2*-T-R] form in Cantonese child-directed speech: the [T-*bei2*-R] form with topic theme constitutes 3.19% of the total, the [*bei2*-R-T] with postposed theme constitutes 0.27% of the total; the serial verb [*bei2*-T-*bei2*-R] form constitutes 1.65% of the total; while the canonical [*bei2*-T-R] form represented 29.52% of the total.

#### **4.2.3 All Full *bei2*-Datives Attested**

Full *bei2*-datives with overt theme and recipient expressed are generally very few before age three among the present monolingual children in CANCORP, relative to the tokens of early non-full *bei2* datives. The two most precocious children CGK and LTF have more tokens (12 and 17 respectively) while other children have tokens ranging from 0 to 6 for each child. Table 4.9 on the next page shows all the full *bei2*-datives attested in CANCORP by age.

Table 4.9. The placement of the theme argument in early *bei2*-full datives in CANCORP

Child	2;03-2;06			2;06-2;09			2;09-3;00		
	Undisplaced T (canonical [ <i>bei2</i> -T-R])	'Displaced' T	[ <i>bei2</i> -T- <i>bei2</i> -R]	Undisplaced T (canonical [ <i>bei2</i> -T-R])	'Displaced' T	[ <i>bei2</i> -T- <i>bei2</i> -R]	Undisplaced T (canonical [ <i>bei2</i> -T-R])	'Displaced' T	[ <i>bei2</i> -T- <i>bei2</i> -R]
CCC					1 (preposed)			1 (preposed)	
CKT									
CGK	4	8 (postposed)	1		1 (preposed)			1 (postposed)	
HHC					1 (preposed)			4 (preposed) 1 (postposed)	
LTF	1	1 (preposed) 3 (postposed)		1	2 (preposed)		1	1 (preposed) 1 (postposed)	
LLY								1 (preposed) 4 (postposed)	1
MHZ		2 (preposed)	1						
WBH								3 (preposed) 1 (postposed) 1(right-dislocated)	

Child	3;00-3;03			3;03-3;06			3;06-3;09		
	Undisplaced T (canonical [ <i>bei2</i> -T-R])	'Displaced' T	[ <i>bei2</i> -T- <i>bei2</i> -R]	Undisplaced T (canonical [ <i>bei2</i> -T-R])	'Displaced' T	[ <i>bei2</i> -T- <i>bei2</i> -R]	Undisplaced T (canonical [ <i>bei2</i> -T-R])	'Displaced' T	[ <i>bei2</i> -T- <i>bei2</i> -R]
LTF	1								
LLY	1	5 (preposed) 1 (postposed)		2	1 (preposed) 1(right-dislocated)	1	1		

The shaded areas indicate the non-canonical *bei2*-forms.

#### 4.2.4 Early Preference for Non-Canonical Forms

Table 4.10 summarizes the overall token measures of the full *bei2*-datives attested for each child in CANCORP. As far as the present data is concerned, the children use more non-canonical forms than the target canonical [*bei2*-T-R] form (see section 4.2.5 for further discussion on the canonical form).

**Table 4.10. Full *bei2*-datives attested in CANCORP (token measures)**

Before Age Three

	Canonical form	Non-canonical forms				Total
	Undisplaced T	Displaced T				
	[ <i>bei2</i> -T-R]	Preposed T [T- <i>bei2</i> -R]	Postposed T [ <i>bei2</i> -R-T]	Right-dislocated T [ <i>bei2</i> -R-PRT-T]		
CCC		2				2
CKT						0
CGK	4	1	9		1	15
HHC		5	1			6
LTF	3	4	4			11
LLY		1	4		1	6
MHZ		2			1	3
WBH		3	1	1		5

After Age Three

	Canonical form	Non-canonical forms				Total
	Undisplaced T	Displaced T				
	[ <i>bei2</i> -T-R]	Preposed T [T- <i>bei2</i> -R]	Postposed T [ <i>bei2</i> -R-T]	Right-dislocated T [ <i>bei2</i> -R-PRT-T]		
LTF	1					1
LLY	4	6	1	1	1	13

A careful look into all these early non-canonical *bei2*-datives in their discourse contexts reveals that the uses of [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms are non-target like for those particular communicative situations from the adult language perspective. In adult Cantonese, the non-canonical [*bei2*-R-T] form is restricted to



formal registers or cases when the theme argumental NP is long or focused, whereas the non-canonical [*bei2-T-bei2-R*] form is restricted to long theme or when the speaker wants to emphasize the second NP as playing the recipient role (refer back to sections 1.1.2 and 1.1.3). In analyzing the developmental cases attested, however, register and the length of the theme argumental NP, do not seem to be the driving motivations for these children to use these non-canonical forms: apart from the serial verb form [*bei2-T-bei2-R*] used at a relatively late age (3;05.20) by LLY, all those postposed themes and themes in the serial verb form [*bei2-T-bei2-R*] attested were not long. Also, the young children clearly were not engaging in formal register at the time of recording. As for information status, it is often hard for the adult researcher to discern whether the theme was stressed or focused at the time of speaking from the child's perspective.

There are qualitative differences associated specifically with some of the early non-canonical full *bei2*-datives. Sections 4.2.4.1 and 4.2.4.2 discuss these differences.

#### **4.2.4.1 Possible Priming Effects**

First, there might be possible lexical and syntactic priming effects in children's use of full *bei2*-datives. The idea is that syntactic and lexical priming may also be one factor in influencing the placement of the theme argument in the early use of full *bei2*-datives. For some early non-canonical *bei2*-datives attested in the corpus, the early placement of the theme seems to instantiate the given (primed)-before-new pattern. Some illustrative examples are shown in (1) to (8) below. The child's non-canonical full *bei2*-dative is highlighted in bold and only English translation is offered for contextual utterances.

- i) Prior mention of the theme leads to subsequent use of a pre-posed theme  
([T-*bei2*-R])

(1) CCC: 2;08.17

CHI: Jau5 ji1sang1 aa3 . (*ji1sang1* refers to the child's toy doctor kit)

'Have doctor.'

INV: Jau5 ji1sang1 aa4 ?

'Have doctor?'

CHI: Hai6 jaa3 .

'Yes.'

INV: Hai2 bin1 dou6 zek1 ji1sang1 ?

'Where is the doctor?'

**CHI: Ji1sang1 aa3 bei2 zo2 jan4 aa3 .**

**Doctor PRT, give PERF people PRT**

**'Doctor, have given people.'**

(2) HHC, 2;06.10

CHI: Ce1 .

'Car.'

CHI: Ce1 aa3 .

'Car.'

**CHI: Paau2ce1 bei2 lei5 .**

**Sport car give you**

**'Sport car give you.'**

(3) WBH, 2;09.19

Situation: The child is looking for some blank paper while the investigator is talking with the child's mother.

CHI: < A zi2 > [/].

'Paper.'

CHI: <A zi2> [/] a zi2.

'Paper, paper.'

CHI: < A zi2 bei2 ngo5 laa1 > [/].

PRT paper give me PRT

'Paper give me PRT.'

CHI: <A zi2 bei2 ngo5> [/] a zi2 bei2 ngo5 aa1.

PRT paper give me PRT paper give me PRT

'Paper give me, paper give me PRT.'

In (1) to (3), these children have first mentioned the theme *ji1sang1* 'doctor' in (1), *ce1* 'car' in (2) and *zi2* 'paper' in (3) in the preceding discourse and then produce a *bei2*-dative in the form of [T-(PRT)-*bei2*-R], suggesting that they might be scaffolding on the theme as the sentential topic to construct these dative constructions.

ii) Prior use of [*bei2*-R] leads to subsequent placement of a postposed T ([*bei2*-R-T])

(4) LTF, 2;03.30

INV: Wa3 , lei5 m4 bei2 ji1ji1 aa4 ## ? [*bei2*-R]

'You don't give auntie?'

CHI: xxx bei2 lei5 aa3 ##. [*bei2*-R]

'xxx give you.'

CHI: Bei2 ngo5 nei1 bun2 aa3 . [*bei2*-R-T]

Give me DET CL PRT

'Give me this one.'

(5) LTF at 2;09.07<sup>11</sup>

CHI: Bei2 ngo5 aa1 . [bei2-R]

'Give me.'

INV: Bei2 mat1je5 nei5 aa3 ? [bei2-T-R]

'Give what to you?'

CHI: Bei2 ngo5 aa3 . [bei2-R]

'Give me.'

INV: Bei2 mat1je5 nei5 aa3 ? [bei2-T-R]

'Give what to you?'

CHI: Bei2 ngo5 go2di1 aa3 . [bei2-R-T]

Give me DET CL PRT

'Give me those.'

(6) LLY at 2;11.08

CHI: Bei2 ngo5. [bei2-R]

'Give me'

SIS: e2 - .

CHI: Bei2 ngo5 "Hello-Kitty". [bei2-R-T]

Give me Hello Kitty

'Give me Hello Kitty.'

(7) LLY at 2;11.29

CHI: Bei2 ngo5. [bei2-R]

Give me.

SIS: < ngo5 wui6 zin2 zo2 nei1 dou6 ge3 > [>].

'I would have cut off here.' (the child's sister is answering the adult investigator's prior question.)

CHI: < Aa3, bei2 ngo5 "Barbie" sin1 > [<]. [bei2-R-T]

PRT give me Barbie first

'Give me Barbie first.'

In four instances of the early non-target [bei2-R-T] forms (two from LTF at 2;09.07 (see (4) and (5)), two from LLY (see (6) and (7))), prior to the child's use of a full *bei2*-dative, the child uses a [bei2-R] non-full dative and then, in a non-target manner,

---

<sup>11</sup> Interestingly, the child is not 'primed' by the adult investigator's use of the canonical [bei2-T-R] form in (5), another piece of evidence that the child has not mastered the canonical [bei2-T-R] form at the time of recording.

uses a [*bei2-R-T*] form, sometimes with the same lexical item as R for the [*bei2-R*] part (as in (4) to (7)), suggesting that the child might be scaffolding on her prior use of the [*bei2-R*] utterance to construct these [*bei2-R-T*] dative forms.

In addition, in one of CGK's early non-target uses of a [*bei2-R-T*] form, the child has produced a [*bei2-R*] form with essentially the same match of lexical items (*jat1 zan6 bei2 suk1suk1*) within the same transcript. (8) shows the child's two utterances.

(8) CGK, 2;03.11: line 1423

Situation: The child has given some imaginary money to the investigator.

CHI:        *Jat1 zan6        bei2   suk1suk1 laa1 .*  
              *a while later   give   uncle        PRT*  
              'Give uncle (this) later.'

CGK, 2;03.11: line 1493.

CHI:        *Jat1 zan6        bei2   suk1suk1 cin2    aa1 .*  
              *A while later   give   uncle    money PRT*

#### 4.2.4.2        **Placement of Pauses**

Second, in five of the early non-target [*bei2-R-T*] forms, there is a significant pause noted between R and T. In the original transcripts, the pause is marked with a # sign. (9) shows all the five child instances, with the # sign highlighted in bold for ease of identification.

- (9) CGK, 2;05.03: Ngo5 bei2 suk1suk1 # di1 cin2 aal .  
 I give uncle # CL money PRT  
 'I give uncle # some money.'
- LLY, 2;11.01: Bei2 aa3saa1 # sing1sing1 aal .  
 Give aa3saa1 # star PRT  
 'Give aa3saa1(a person's name) # star.'
- LLY, 2;11.29 < Wai3 , bei2 faan1 nei5 # e3 &wou1&wou1 aal > [=! talking to her sister] .  
 PRT give PVT you # PRT doggie PRT  
 'Give back you # e3 doggie'
- LLY, 3;00.11 Bei2 ze4ze1 # jam2gun2 .  
 Give sister # straw  
 'Give sister # staw.'
- LLY, 3;03.15 Bei2 ngo5 # xxx nei1 di1 sin1 .  
 Give me # xxx DET CL first  
 'Give me # xxx these first.'

#### 4.2.5 The Late Acquisition of the Canonical [*bei2-T-R*] Form

The canonical [*bei2-T-R*] form is a low frequency structure among the present young child subjects in CANCORP. Apart from the two most precocious children CGK and LTF, no spontaneous use of a canonical form is attested in the other six children before age three. Only one instance of imitation (see (10)) is found when CCC is at 2:10.13, where the adult provides an explicit model for the child to imitate.

(10) CCC, 2;10.13: line 3039

- \*INV: Lei5 waa6 m4goi1 ze2ze2 -: < bei2 go3> [>] zam2tau4 ngo5 aa1 .  
You say please sister give CL pillow me PRT  
'You say "sister please give a pillow to me".'
- \*CHI: < M3 a3> [<].
- \*CHI: M4goi1 ze2ze2 bei2 zam2tau4 ngo5 aa1 .  
Please sister give pillow me PRT  
'Sister please give pillow to me.'

Similarly, only one instance of imitating the adult's preceding canonical [*bei2-T-R*] utterance is found when CKT is 2;07.02: line 3488. See (11).

(11) CKT, 2;07.02: line 3488.

- \*INV: Bei2 jat1 go3 ngo5 la1 ##.  
Give one CL me PRT  
'Give one to me.'
- Situation: The child gives a puppet to the investigator.
- \*INV: Do1ze6 ##.  
'Thank you.'
- \*CHI: Bei2 jat1 go3 ngo5 laa1 .  
Give one CL me PRT  
'Give one to me.'
- \*INV: O4 .
- \*CHI: Jat1 go3 ngo5 laa1 .  
One CL me PRT  
'One me.'
- Situation: The investigator gives a puppet to the child.

In (11), CKT's imitation on his second attempt in the form of *jat1 go3 ngo5 laa1* 'one me' suggests that the child has not parsed the [*bei2-T-R*] string correctly, since [*jat1 go3 ngo5*] is not a possible constituent.

As for MHZ, only one instance of imitating the adult's preceding utterance is noted when he is 1;10.10. See (12).

(12) MHZ, 1;10.10: line 4570

\*MOT: Bei2 faan1 gei1gei1 ze2ze2 aa1 .  
Give PVT machine sister PRT  
'Give back the machine to sister.'

\*CHI: Bei2 gei1 ze2ze2.  
Give machine sister  
'Give machine sister.'

Even imitation of the canonical form [*bei2*-T-R] is not always easy for young children acquiring Cantonese. HHC is not successful in imitating the canonical form in full at 2;04.08 and seems to show reluctance or difficulty in imitating the canonical form at 2;06.10, 2;07.21 and 3;04.14 even though the adult investigator provides frequent explicit models for pragmatic purposes such as politeness. (13) to (16) show the relevant scenarios.

(13) HHC, 2;04.08: lines 1726 to 1734

Situation: The child is playing with some combs.

\*SIS: Bei2 jat1 baa2 ngo5 aa1 .  
Give one CL me PRT  
'Give one to me.'

\*CHI: Bei2 jat1 baa2 aa3 .  
Give one CL PRT  
'Give one.'



(14) HHC, 2;06.10: lines 1621-1650

\*INV: Nei5 waa6 m4goi1 ze2ze2 bei2 di1 nai4gaau1 ngo5 aa1.

You say please sister give CL plasticine me PRT

'You say "sister please give some plasticine to me.'

\*INV: Nei5 gong2 sin1 .

You say first

'You say first.'

\*CHI: Ze2ze2 .

Sister

\*INV: Ze2ze2 mat1je5 aa3 ?

Sister what PRT

'Sister what?'

\*CHI: Ze2ze2 .

Sister

\*INV: Bei2 nai4gaau1 .

Give plasticine

\*CHI: Nai4gaau1 .

Plasticine

\*INV: Bei2 nai4gaau1 .

Give plasticine

\*CHI: Hou2 je3 .

Hurray

(15) HHC, 2;07.21 lines 4339- 4410

\*CHI: M4 m4 a1 [=! wants to take the candies].

M4 m4 a1 [=! wants to take the candies].

\*INV: Lei5 man6 ngo5 aa1 .

You ask me PRT

'You ask me.'

\*INV: Lei5 jung6+...

'You use..'

\*INV: Lei5 gong2.

'You speak.'

\*CHI: Zoeng1zeon3him1 # sik6 tong4.

Zoeng1zeon3him1 # eat candy

'Zoeng1zeon3him1(the child's name) # eat candy.'

\*INV: Lei5 waa6, m4goi1 ze2ze2 bei2 jat1 lap1 tong4 ngo5 aa1 .

You say please sister give one CL candy me PRT

'Say "sister please give one candy to me.".'

- \*CHI: Tong4tong4.  
'Candy.'
- \*CHI: Nil .
- \*CHI: Ak1 .
- \*INV: Lei5 waa6 m4goi1 ze2ze2.  
You say please sister  
'Say "please sister.".'
- \*CHI: Ge1 .
- \*INV: Dak1 m4 dak1 ?  
'Okay?'
- \*CHI: Dak1 ga3.  
'Okay.'
- \*INV: M4goi1 ze2ze2.  
'Please sister.'
- \*CHI: Goi1.
- \*INV: Bei2 jat1 lap1 tong4 ngo5.  
Give one CL candy me  
'Give one candy to me.'
- \*CHI: Tong4 ngo5.  
'Candy me'
- \*INV: Lei5 gan1 zyu6 ngo5 gong2 jat1 ci3 sin1.  
You follow PVT me say one CL first  
'Say it after me once.'
- \*CHI: Lei5 tong4.  
'You candy.'
- \*INV: Lei5 waa6" bei2 jat1 lap1 tong4 ngo5 aa1" .  
You say give one CL candy me PRT  
'Say "give one candy to me.".'
- \*CHI: Hou2 aa3 .  
Good.
- \*INV: Lei5 tung4 ngo5 gong2 aa3 .  
You to me speak PRT  
'Say it after me.'
- \*CHI: Waa1.
- \*CHI: Zoeng1zeon3him1 heoi3 sik6 tong2.  
Zoeng1zeon3him1 go eat candy  
(*Zoeng1zeon3him1* is the child's name)

(16) HHC, 3;04.14, lines 2143 to 2154

- \*INV:     Lei5 waa6 m4goi1 gaa1ze2 bei2 jat1 go3 # gau2 zai2 ngo5 aa1 .  
          You say please sister give one CL doggie me PRT  
          ‘Say “sister please give one doggie to me.”.’
- \*CHI:     M4goi1 bei2 < nei1 go3 gau2 > [>] +/.  
          Please give DET CL dog  
          ‘Please give this dog.’

As for LLY, the first instance of spontaneous use is attested at age 3;01.13. Even though the canonical [*bei2-T-R*] form is noted early (before age three) in the two most precocious children CGK and LTF, non-target [*bei2-R-T*] forms- when the canonical form [*bei2-T-R*] should be the target form for those communicative instances- are observed concurrently. The canonical form is also used less frequently than the non-target forms (target canonical: non-target [*bei2-R-T*] and [*bei2-T-bei2-R*] forms is 4: 9 and 1: 3 (token measures) in CGK and LTF respectively at the early stage between ages 2;03 and 2;06). A detailed lexical analysis of the data reveals that out of the four tokens of the canonical [*bei2-T-R*] used by CGK, three instances involve the theme *cin2* ‘money’ where *bei2cin2* is an idiomatic expression meaning ‘to pay’. As for the other precocious child LTF, there are three tokens of the canonical [*bei2-T-R*] form attested in her corpus before age three. The first spontaneous use however involves significant pauses between constituents while the child is formulating the expression:

(17) LTF, 2;04.27: line 4282

- Bei2 e6     dai6 go3     e6     nei1 go3     ngo5 e3 .  
Give PRT DET CL PRT DET CL me PRT  
‘Give e6 another one e6 this one to me.’

The second instance involves the theme with only the classifier expressed (see (18)).

(18) LTF, 2;07.20: line 4459

Bei2 go3 ngo5.

Give CL me

'Give one me.'

As an interim summary, our developmental data suggests that children before age three- the precocious ones aside- have difficulty in using the canonical [*bei2-T-R*] form spontaneously. Our corpus data even shows that some young children show reluctance or difficulty in imitating the canonical form immediately after the adult's explicit model of the canonical form, despite having the tendency to imitate other structures, as we have seen in the examples from HHC above. Even when two of the most precocious children in CANCORP use the canonical form at their early ages before age three, the few tokens noted in which some tokens are often associated with the idiomatic expression *bei2cin2* meaning 'to pay', and their inconsistent use of the canonical form across early communicative situations which demand the use of a canonical form as the target structure (recall the early concurrent use with non-target [*bei2-R-T*] and [*bei2-T-bei2-R*] forms) does not in any way suggest that CGK and LTF have acquired the canonical [*bei2-T-R*] form. We so far see no strong evidence to suggest that the canonical form is acquired before age three among the present eight monolingual children in CANCORP.

#### 4.2.6 The Non-Target Use of *bei2*-Datives

The non-target [*bei2-R-T*] and [*bei2-T-bei2-R*] forms are observed in children acquiring Cantonese in different acquisition contexts. Tables 4.11 and 4.12 summarize the relevant information.

**Table 4.11. The early developmental non-target [bei2-R-T] and [bei2-T-bei2-R] forms in different acquisition contexts**

Population	Source of Data	
Monolingual Cantonese children	Naturalistic longitudinal corpus data from CANCELP	
Cantonese-English Bilingual children acquiring Cantonese	Naturalistic longitudinal corpus data from the Hong Kong Bilingual Child Corpus (Yip, Matthews & Huang 2001)  Diary Data from a child language researcher, Cheung Sik Lee	
Children with language delays (and disorders) acquiring Cantonese	Anecdotal notes from three local speech therapists in Hong Kong	

**Table 4.12. Prevalence of the non-target [*bei2-R-T*] and [*bei2-T-bei2-R*] forms in early child Cantonese (by token measures)**

	[ <i>bei2-R-T</i> ] Form	[ <i>bei2-T-bei2-R</i> ] Form
Monolingual	21 tokens (21 from 5 children out of a total of 90839 child utterances in CANCORP: 9 from CGK out of her 6300 child utterances in the corpus; 4 from LTF out of her 11099 child utterances in the corpus, 6 from LLY out of her 10976 child utterances in the corpus, 1 from HHC out of his 13782 child utterances in the corpus; 1 from WBH out of her 5558 child utterances in the corpus)	4 tokens (4 from 3 children out of a total of 90839 child utterances in CANCORP: 1 from CGK out of her 6300 child utterances in the corpus; 1 from MHZ out of his 12263 child utterances in the corpus; 2 from LLY out of her 10976 child utterances in the corpus)
Bilingual <sup>12</sup>	42 tokens (32 from 3 children out of a total of 26136 child utterances in the HK Bilingual Child Corpus and 9 from Siu Bou's diary data: 11 from Timmy out of his 10631 child utterances in the corpus; 18 from Sophie out of her 12574 child utterances in the corpus; 4 from Kathryn out of her 2931 child utterances in the corpus)	6 tokens (4 from 3 children out of a total of 26136 child utterances in the HK Bilingual Child Corpus and 2 from Siu Bou's diary data: 1 from Timmy out of his 10631 child utterances in the corpus; 1 from Sophie out of her 12574 child utterances in the corpus; 2 from Kathryn out of her 2931 child utterances in the corpus)
Clinical	7 tokens from 5 children	4 tokens from 3 children

The phenomenon is quite robust. Among the nineteen children<sup>13</sup> whose developmental data is available for current analysis, at least one instance of

<sup>12</sup> Parents of the bilingual Cantonese-English child, Alicia, the younger sister of Timmy and Sophie, reported the productive use of [*bei2-R-T*] by Alicia. Alicia is one of the child subjects in the Hong Kong bilingual children corpus project, although her developmental data has not yet been released at the time of conducting the present study.

<sup>13</sup> They include eight monolingual 'corpus' children, three bilingual 'corpus' children, one bilingual 'diary' children and seven children with language delay (and language disorder).

non-target *bei2*-datives is found in five of the eight monolingual children studied, all the four Cantonese-English bilingual children studied, and seven children with language delay (and disorder). This finding indicates that the non-target [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms is not an idiosyncratic use of structures by only one or two children.

Of the total eighty-five tokens of early non-target *bei2*-datives reported in Table 4.12 above from children acquiring Cantonese in different acquisition contexts, examination of the child's usage in context reveals the following observations.

First, these are all tokens of spontaneous production from young children, that is, they are not direct or partial imitations of the adults' preceding utterances, nor are they structured from or modeled on the adult's use of a preceding question in the form of [*bei2*-R-T<sub>what?</sub>], [*bei2*-T<sub>what</sub>-*bei2*-R?], or [*bei2*-T-*bei2*-R<sub>who?</sub>].

Second, some of the non-target [*bei2*-R-T] forms are even produced immediately after the adult's use of a *canonical* [*bei2*-T-R] form, revealing a marked contrast between the non-target word order use of the child ([*bei2*-R-T]) and the target canonical word order use of the adult interlocutor ([*bei2*-T-R]). (19) shows a representative example.

(19) LTF, 2;09.07

CHI: Bei2 ngo5 aa3 . [bei2-R]  
Give me PRT  
'Give me'

INV: Bei2 mat1je2 lei5 aa3 ? [bei2-T-R]  
Give what you PRT  
'Give what to you?'

CHI: Bei2 ngo5 go2 di1 aa3 . [bei2-R-T]  
Give me DET CL PRT  
'Give me those.'

Third, for some tokens of the non-target [bei2-R-T] form, the consecutive self-repetitions of the [bei2-R-T] form noted in some children suggest that the non-target production is not a "once-in-a-while" performance error. (20) shows a representative example. Moreover, the use of slightly different lexical content (difference in the use of the NP structure) in referring to the same discourse referent within the consecutive self-repetitions suggests that these non-target forms are *not* frozen utterances learnt/ stored as a whole.

(20) CGK, 2;03.11 \*\*\* File "20311.cha": line 1463.

\*CHI: Bei2 ngo5 doi6 aa1 .  
Give me bag PRT  
'Give me bag.'

CGK, 2;03.11 \*\*\* File "20311.cha": line 1466.

\*CHI: Bei2 gaa1kei4 go3 doi6 aa1 .  
Give gaa1kei4 CL bag PRT  
'Give GaaKei (the child's name) the bag.'

CGK, 2;03.11 \*\*\* File "20311.cha": line 1469.

\*CHI: Bei2 gaa1 kei4 doi6 aa1 .  
Give gaa1kei4 bag PRT  
'Give GaaKei (the child's name) a bag.'



#### 4.2.6.1 The Non-Target [*bei2*-R-T] Form

The non-target [*bei2*-R-T] form is prevalently observed in five monolingual children from CANCELP, four Cantonese-English bilingual children (three from the bilingual corpus- Timmy, Sophie and Kathryn; and one from Cheung Sik Lee's diary data- her bilingual son) and five children with language delay (and disorder) acquiring Cantonese. In an informal conversation, two mothers also reported noting the non-target [*bei2*-R-T] form from their monolingual children when the children were young.

Evidence from monolingual children is as follows. Appendix (1) lists all twenty-one examples of the non-target [*bei2*-R-T] form in the order in which they were spontaneously produced during the recordings. The utterances are from the five subjects, CGK, LTF, LLY, HHC and WBH of CANCELP. (21) shows some representative examples.

- (21) CGK, 2;03.11 \*\*\* File "20311.cha": line 1493.  
 \*CHI: Jat1zan6 bei2 suk1suk1 cin2 aa1 .  
 Later give uncle money PRT  
 'Later give uncle money.'
- LTF, 2;03.30 \*\*\* File "20330.cha": line 2854.  
 \*CHI: Bei2 ngo5 nei1go3 aa3 .  
 Give me DET CL PRT  
 'Give me this one.'
- LLY, 2;11.01 \*\*\* File "LLY21101.cha": line 2187.  
 \*CHI: Bei2 aa3saa1 # sing1sing1 aa1 .  
 Give aa3saa1 star PRT  
 'Give aa3saa1 # a star.'
- HHC, 2;10.13 \*\*\* File "21013.cha": line 3322.  
 \*CHI: Baau2 sei2, bei2 nei5 nei1go3.  
 full die give you DET CL  
 '(I'm) too full, give you this one.'
- WBH, 2;09.19 \*\*\* File "20919.cha": line 1342.  
 \*CHI: Bei2 faan1 ngo5 zi2 aa1.  
 'Give PVT me paper PRT.'  
 'Give me back paper.'

Evidence from Cantonese-English bilingual children acquiring Cantonese is as follows. The non-target [*bei2-R-T*] form is noted among four young Cantonese-English bilingual children, Timmy, Sophie, Kathryn and Siu Bou. (22) shows some relevant examples from these four children. Appendix (2) shows all the instances of [*bei2-R-T*] found in the existing corpus data of Timmy, Sophie and Kathryn, and the diary data of Siu Bou, listed in the order in which they were produced in the recordings.

- (22) Timmy, 2;07.14: Bei2 keoi5 zyu1gwu1lik1 laa1 .  
 Give it chocolate PRT  
 'Give it(the mouse) chocolate.'
- Sophie, 2; 03. 24: Bei2 ngo5 jat1 tiu4 aa1 .  
 Give me one CL PRT  
 'Give me one.'
- Kathryn, 3;03.16: Ngo5 dou1 zyu2 di1 je5 aa3 , < ngo5 , ngo5 > [/] ngo5 bei2 nei5 jelly  
 I also cook CL stuff PRT I I I give you jelly  
 laa1 .  
 PRT  
 'I also cook some stuffs, <I, I> [/] I give you jelly.'
- Siu Bou, 2; 04.18 : Ngo5 bei2 baa1baa1 go2-go3 .  
 I give daddy DET CL  
 'I give Dad that one.'

The bilingual siblings Timmy and Sophie show clear indications of dominance in Cantonese in their preschool years- their Cantonese is close to monolingual Cantonese, while their English shows strong Cantonese influence (see Yip and Matthews, 2000; Matthews and Yip, 2003), yet they use the [*bei2-R-T*] form in a non-target manner even till the age of 5;08 for Timmy (as noted in Sophie's corpus data) and 5;03 for Sophie (as noted by her mother), a word order pattern which is more in line with their non-dominant language, English. On the other hand, the language dominance pattern of the other bilingual child, Kathryn, is considered more balanced with respect to her developing Cantonese and English. For Kathryn's corpus data, only seventeen Cantonese files are available for analysis at the time of the present study. Nevertheless, I am still able to find four tokens (three at 3;03.16; one at 4;02.17) of the [*bei2-R-T*] form from the limited developmental data available. Instances of the [*bei2-R-T*] form are also found in the diary data of Siu Bou, whose

Cantonese was stronger than his English before he started preschool at 2;08. Siu Bou used the [bei2-R-T] form in his Cantonese *bei2*-dative constructions even till the age of 4;00.20, as noted and reported by his mother. To what extent we can claim that these young bilingual children's non-target [bei2-R-T] forms, whose word order seems to be more in line with their another developing language- English, is a product of interaction between English influence (if any) and the Cantonese developmental pathway awaits further in-depth study, but for the time being, I note a protracted use of non-target [bei2-R-T] forms in our bilingual subjects.

Evidence from the clinical child population with language delay (and disorder) acquiring Cantonese is as follows.<sup>14</sup> (23) shows all the seven instances of [bei2-R-T] forms recorded by the local speech therapists.<sup>15</sup>

---

<sup>14</sup> Special thanks to Frankie Lui, Suki Lam and Kim Yan, speech therapists in Hong Kong, who provided the clinical data.

<sup>15</sup> The ages shown here are chronological ages. See Table 4.5 for information on mental ages and language ages.

- (23) LKM, 8;02: Bei2 mui4mui2 syut3gou1  
 Give sister ice-cream  
 ‘Give sister ice-cream.’
- CMM, 5;02: Bei2 ngo5 go3 bo1 aa1  
 Give me CL ball PRT  
 ‘Give me CL ball.’
- FMH, 4;03.26: Bei2 jing1ze4ze1 dim2meng2bou2  
 Give Sister *Jing* attendance book  
 ‘Give Sister *Jing* attendance book.’
- TSW, 10;07.28: Bei2 ngo5 ci4gang1  
 Give me spoon  
 ‘Give me spoon.’
- TSW, 10;08.00: Bei2 ngo5 faan6  
 Give me rice  
 ‘Give me rice.’
- TSW, 10;08.07: Bei2 ngo5 tung2  
 Give me bucket  
 ‘Give me bucket.’
- CYM, 15;06.00: Bei2 ngo5 din6waa2  
 Give me telephone  
 ‘Give me telephone.’

#### 4.2.6.2 The Non-Target [*bei2-T-bei2-R*] Form

The non-target serial verb [*bei2-T-bei2-R*] form is also prevalent across acquisition contexts. It has been observed in three monolingual children in CANCORP, three Cantonese-English bilingual children (two from the bilingual corpus- Timmy and Sophie, and one from Cheung Sik Lee’s bilingual diary data- her son), and three

children with language delay (and disorder) acquiring Cantonese.<sup>16</sup>

Evidence from monolingual Cantonese children is as follows. (24) shows all the four tokens attested in CANCELP.

- (24) CGK, 2;03.04: Bei2 cin2 bei2 lei5 aa1  
Give money give you PRT  
'Give money to you.'
- MHZ, 2;03.09 Baa1baa1 bei2 hung1sung1beng2 bei2 hou6zeon1  
Daddy give muffin give *hou6zeon1*  
'Daddy give muffin to *hou6zeon1* (the child's name).'
- LLY, 2;11.08 M2- e3 bei2 jat1 go3 bei2 ngo5  
PRT PRT give one CL give me  
'Give one to me.'
- LLY, 3;05.20 Bei2, bei2 go2 go3 jau5 zi6 go2 go3 bei2 ngo5 aa1, jau5 zi6 go2 go3  
Give give DET CL have word DET CL give me PRT have word DET CL  
'Give, give the one which has words to me, the one which has words.'

Evidence from Cantonese-English bilingual children acquiring Cantonese is as follows. (25) shows all the six tokens attested in the corpus and diary data.

---

<sup>16</sup> The present data shows that the [*bei2-T-bei2-R*] form occurs less frequently than the [*bei2-R-T*] form in early child Cantonese speech, and that the [*bei2-T-bei2-R*] form happens not to show up in CANCELP for some children (see table 4.12). I am not able to tell at this point whether the lower overall frequency observed and the absence in some children are just artifacts resulted from naturalistic sampling of data, because the naturalistic language samples might just fail to catch some of the [*bei2-T-bei2-R*] forms actually produced by the children. One way to improve the situation is to make use of elicited production tasks to set up an obligatory context which requires the children to express both the theme and the recipient overtly for the Cantonese *bei2*-dative, and then we shall see whether the children have a general tendency to use the [*bei2-T-bei2-R*] form less frequently than the [*bei2-R-T*] form. We also need to recruit more child subjects to address the issue of individual differences: whether there are in fact individual differences among children in the use of the [*bei2-T-bei2-R*] form and its frequency of use relative to the [*bei2-R-T*] form.

(25) Timmy, 3;02.26: Li1 go3 tin1 sai3 bei2 sai3 gaan1 uk1 bei2 li1 go3 jan4 laa1 .  
DET CL sky small give small CL house give DET CL person PRT  
'This sky is small give small house to that person.'

Sophie, 2;11.18: Bei2 jat1 go3 bei2 lei5 aa1  
Give one CL give you PRT  
'Give one to you.'

Siu Bou, 3;09.29: Bei2 syu1 bei2 ngo5  
Give book give me  
'Give book to me.'

Bei2 daai6bun6zeong6 bei2 ngo5  
Give big elephant give me  
'Give big elephant to me.'

Kathryn, 4;01.09: < Bei2 naai5 > [<] bei2 keoi5 , bei2 di1  
Give milk give 3sg give CL  
'Give milk to him/her, give some.'

Kathryn, 4;05.10: Ngo5 bei2 saam1man4zi6 bei2 keoi5  
I give sandwich give 3sg  
'I give sandwich to her/him.'

Evidence from the clinical child population with language delay (and disorder) acquiring Cantonese is as follows. (26) shows all the four tokens attested.

- (26) LKM, 8;03: Bei2 sau2biu1 bei2 maa1maa1  
 Give watch give mother  
 'Give watch to mother.'
- LKM, 8;05: Ngo5 bei2 sau2biu1 bei2 maa1maa1  
 I give watch give mother  
 'I give watch to mother.'
- KTY, 3; 11. 27: Bei2 tip3ji2 bei2 ngo5  
 Give sticker give me  
 'Give sticker to me.'
- LMY, 4;07: Bei2 hung5sik1saam1 bei2 go4go1  
 Give red clothes give brother  
 'Give red clothes to brother.'

The early non-target [*bei2-T-bei2-R*] form instead of the target canonical [*bei2-T-R*] form is interesting in terms of the hypothesized relationship between the lexically specific [*bei2-T-R*] double object construction and the productive [*V-T-bei2-R*] serial verb construction in Cantonese (refer back to table 1.2). See section 5.9.2 in chapter five for further discussion.

### 4.3 Usage Patterns in Older Children

Table 4.13 below shows two elder siblings' use of the Cantonese *bei2*-datives as attested in their young siblings' corpus data in CANCEP.<sup>17</sup> LLY's elder sister is four years older than LLY. LLY's speech was recorded when she was 2;08.10 to 3;08.09, hence her elder sister was, at the time of recording, aged 6;08.10 to 7;08.09. HHC's elder sister is seven years older than HHC. HHC's speech was recorded

---

<sup>17</sup> Data from LTF's elder sister is not included in Table 13 because the overall number of *bei2*-datives attested in the present transcripts is too few (n=4) to allow for a detailed analysis of any kind. Also, the age difference between LTF and her elder sister is not provided in CANCEP.



when he was 2;04.08 to 3;04.14, hence his elder sister was, at the time of recording, aged 9;04.08 to 10;04.14. For these two elder siblings, the use of the canonical [bei2-T-R] form does not seem to pose any problem (n=15 for LLY's elder sister; n=7 for HHC's elder sister). If we consider the *percentage* use of the canonical [bei2-T-R] form out of the total tokens of *bei2*-datives attested, it is 41.7% and 31.8% for LLY's elder sister and HHC's elder sister respectively. Their percentage use of the canonical [bei2-T-R] form is either comparable to or no less than the percentage use of the canonical [bei2-T-R] form (29.52%) by the adult interlocutors as noted in the present adult input findings in section 3.4.2.

Based on the current limited data on the older siblings, we only know that a 6;08 child's use of *bei2*-datives resembles the adult usage patterns. To determine at what age children acquiring Cantonese begin to show productive use of the canonical [bei2-T-R] form and at a level comparable to adult use (29.56%), we need at least cross-sectional (but ideally longitudinal) data spanning till at least age seven to determine at what age children acquiring Cantonese show productivity of the canonical [bei2-T-R] form.

**Table 4.13. The elder siblings' use of the Cantonese *bei2*-datives as attested in their young siblings' corpus data (numbers represent token measures)**

	LLY's elder sister	HHC's elder sister
null T and null R	2	5
null T	10	10
null R	3	-
canonical [bei2-T-R]	15	7
preposed T	3	-
postposed T [bei2-R-T]	1	-
relative clause with relativized T	2	-
Overall tokens:	36	22

#### 4.4 Summary of Major Findings

In this chapter, I have established the following important facts:

- i) The monolingual Cantonese children use non-full *bei2*-datives in the form of [*bei2*-R] and [*bei2*-T] before using full *bei2*-datives. Apart from the child HHC, [*bei2*-R] emerges earlier than [*bei2*-T], a sequence that is frequently used in the adult input (see section 3.4.2.4). The primacy of non-full datives in early developmental speech is also observed in other non-pro-drop languages, see for example Kiekhoefer (2002) for similar findings obtained for early child English and German.
- ii) Full *bei2*-datives with overt theme and recipient are generally very few among the present monolingual subjects.
- iii) The Cantonese canonical [*bei2*-T-R] form, in particular, is a low frequency structure among young children. Late use, inconsistent use or even absence of the canonical [*bei2*-T-R] is noted in the present young child subjects before age three. The child HHC even fails to imitate the canonical [*bei2*-T-R] form at an early age. Some precocious children like LTF and CGK are able to spontaneously produce the canonical form but do not do so consistently, as indicated by the concurrent use of the target canonical form and the non-target [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms. We so far see no strong empirical evidence to suggest that the canonical [*bei2*-T-R] form is acquired in these children's early developing grammar before age three.

iv) Non-canonical forms are used instead at the early stage of development when these young children are on the way to acquiring the target canonical [*bei2-T-R*] form but have to express both the theme and recipient overtly. In particular, the [*bei2-R-T*] and [*bei2-T-bei2-R*] forms are used in a non-target manner from the adult language perspective. Among the present monolingual child subjects, the [*bei2-R-T*] form is noted in five out of eight children studied, while the serial verb form [*bei2-T-bei2-R*] is noted in three out of eight children. There are qualitative differences associated specifically with some of these non-target forms: early placement of the theme seems to instantiate the given (primed)-before-new pattern, and there is a significant pause between R and T in [*bei2-R-T*], i.e. [*bei2-R-#-T*]. The non-target [*bei2-R-T*] and [*bei2-T-bei2-R*] forms are also found in the bilingual and clinical child population acquiring Cantonese, as demonstrated by naturalistic bilingual corpus data, bilingual diary data as well as clinical data from some local speech therapists in Hong Kong.

## Chapter Five. Discussion of Findings

### 5.0 Introduction

This chapter discusses the findings of the investigation in relation to the hypotheses which I have established in the previous chapters. Section 5.1 first reviews the empirical predictions based on the markedness hypothesis, the iconicity hypothesis, the input frequency hypothesis and the input properties hypothesis. Sections 5.2 to 5.5 address how the developmental findings reported in chapter four relate to these hypotheses. The remaining sections discuss some *new* issues that come out of the present findings. It is interesting to point out that the occurrence of the non-target [*bei2*-R-T] form in early Cantonese, despite its low frequency in the input, appears to lend support to Bruyn et al. (1999)'s markedness hypothesis from a UG perspective. Section 5.6 addresses this issue and points out some problems with this hypothesis. Section 5.7 considers the preference for the non-canonical forms over the canonical form from a functional perspective. Section 5.8 considers how the non-canonical forms might be generated from a usage-based perspective. Section 5.9 raises some remaining questions. Section 5.10 provides a chapter summary.

### 5.1 A Review of Established Empirical Predictions

In chapter two, I have investigated how the markedness hypothesis, the iconicity hypothesis and the input frequency hypothesis in the existing literature apply to the acquisition of this Cantonese *bei2*-dative construction. In chapter three, I have investigated the input properties hypothesis by examining how the language specific input properties of Cantonese might impact the early schematization of the canonical [*bei2*-T-R] double object form. The empirical predictions are as follows. The

markedness hypothesis predicts the non-canonical forms to be preferred over the canonical [*bei2*-T-R] form, because the non-canonical forms are more unmarked than the canonical form cross-linguistically. The iconicity hypothesis, in O'Grady (2000) and Cho et al. (2002)'s terms, predicts the canonical [*bei2*-T-R] form and the non-canonical [*bei2*-T-*bei2*-R] form to be preferred over non-canonical [*bei2*-R-T] form, because the theme precedes the recipient in the [*bei2*-T-R] and the [*bei2*-T-*bei2*-R] forms, and this theme-recipient ordering is isomorphic to how the transfer event naturally unfolds. The input frequency hypothesis predicts the more frequent canonical [*bei2*-T-R] form to be preferred over the non-canonical forms. The input properties hypothesis formulated in this thesis predicts that the Cantonese input properties are not conducive to early schematization of the canonical [*bei2*-T-R] form, and so disfavor its early acquisition.

## 5.2 The Markedness Hypothesis

The present developmental findings support the markedness hypothesis. Typological-developmental parallels are clearly noted. The non-canonical [*bei2*-R-T] form, being an instance of the [V-R-T] double object construction, and the non-canonical [*bei2*-T-*bei2*-R] form, being an instance of the [V-T-dative marker-R] construction (the indirect object construction in terms of Michaelis and Haspelmath (2003), in which the indirect Recipient object is 'flagged') are commonly attested across languages. The present findings show that these non-canonical *bei2*-forms emerge early in child Cantonese when children are on the way to mastering the canonical form, despite being used in a non-target manner from the adult language perspective. The occurrence of non-target [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms in different acquisition contexts is demonstrated by naturalistic monolingual corpus data and elicited data from a pilot study, naturalistic

bilingual corpus data, bilingual diary data as well as clinical data from some local speech therapists in Hong Kong. To the extent that monolingual, bilingual and clinical child populations acquiring Cantonese are found to use the [bei2-R-T] and [bei2-T-bei2-R] forms in a non-target manner, these non-canonical forms appear to represent an option available to all these language learners. The canonical [bei2-T-R] form, being an instance of the [V-T-R] double object construction, on the other hand, appears to be marked in languages without cases (Matthews and Leung, 2002) and dispreferred in the languages of the world (Susanne Michaelis, 2003, p.c.). The present findings show that the canonical form proves to be *dis*-preferred in early Cantonese production before age three.

### 5.3 The Iconicity Hypothesis

The present developmental findings do not support the iconicity hypothesis. Although the canonical [bei2-T-R] form exhibits a word order iconic (isomorphic) to how the transfer event naturally unfolds as described by O'Grady (2000) (refer back to (1) in chapter two), it is *dis*-preferred in early production before age three as the current findings show. The non-canonical [bei2-R-T] form does not appear to be iconically motivated, yet it is found in the production of most children acquiring Cantonese at an early age. As for the non-canonical [bei2-T-bei2-R] serial verb form, whose ordering of the theme and the recipient is iconic (isomorphic) to how the transfer event naturally unfolds as the canonical [bei2-T-R] form is, iconicity might be a potential factor preferring its early use. To what extent we can invoke the iconicity explanation as a cogent argument in accounting for the present naturalistic production data is debatable, principally because the author cannot tell only from the transcripts and reports of the present naturalistic production data that the children's use of the [bei2-T-bei2-R] non-canonical form is highly correlated

with how the actual situation naturally unfolds, let alone be certain about whether children construe the transfer event in terms of the corresponding caused-motion semantics (Goldberg, 1995) when using the [*bei2-T-bei2-R*] form. To address this issue properly for *production*, we might need to conduct further experimental tasks. One suggestion is to make use of controlled elicited production tasks to present transfer events in two conditions and then elicit children's narration of these events. The first condition is to present the transfer event in a sequential manner, whereby the agent is first highlighted, followed by the theme and then finally the recipient. Sequential highlighting of constituents can be done by scanning the semantic roles of the transfer event in a particular order using a video-cam. The second condition is to present the transfer event with no sequential highlight. The iconicity explanation- particularly in terms of temporal sequence- might be supported if there is a preferred use of the non-canonical [*bei2-T-bei2-R*] form in the first condition (the 'sequentially scanned' theme-recipient order), relative to the second 'control' condition.

Another relevant point to note is that the [*bei2-R-T*] form might arguably "be considered iconic if giving events, especially those involving human recipients, are seen as events in which the agent acts on the recipient causing it to have the theme." (O'Grady, November 2003, p.c.), contrary to the structure of a transfer event described in O'Grady (2000: 4) (refer back to (1) in chapter two). This description of a transfer event is similar to the event structure posited for the English ditransitive construction by Goldberg (1995), in this way, the [*bei2-R-T*] surface form shares the essential constructional semantics with the English ditransitive construction, the central sense being 'X causes Y to receive Z' (Goldberg, 1995: 151). However, some ensuing problems arise when one goes into formulating an iconicity hypothesis

for the [*bei2*-R-T] form. If the alleged schema is something like: [X causes Y to receive Z] for the case of the [*bei2*-R-T] form, and that the notion of iconicity corresponds to a match between syntax and semantics, in the case of the [*bei2*-R-T] surface form, there is nothing in syntax to correspond to "receive".<sup>1</sup>

Putting the above problems aside for the time being, if we do accept that there is an iconic match between syntax and semantics for the [*bei2*-R-T] form, the iconicity hypothesis then becomes untestable in the case of Cantonese *bei2*-datives under the present study, because all three patterns ([*bei2*-T-R], [*bei2*-R-T] and [*bei2*-T-*bei2*-R]) are then iconic, in their own way. Consequently, if any one of the above patterns emerges early (or is preferred in early child language), one can always invoke the presence of iconic motivation as a factor contributing to the early preference, since all of the above patterns can arguably be considered iconically motivated in their own way (although the argument is not without problems, as mentioned above), the iconicity explanation then becomes unfalsifiable, at least for the present set of Cantonese *bei2*-data we are working with in this study. Moreover, if all three patterns are considered iconically motivated in their own way, iconicity in itself cannot explain the following: 1) if one pattern turns out to be preferred over the other, it cannot explain such differential preference; 2) if it turns out that one of the above

---

<sup>1</sup> Putting the iconicity issue aside, if the [*bei2*-R-T] form is associated with the event structure [X causes Y to receive Z], while the canonical [*bei2*-T-R] form is associated with another event structure [X causes Y to go to Z], one might wonder whether the early developmental tendency to use the [*bei2*-R-T] form prior to the [*bei2*-T-R] form reflects children's cognitive development of causative events, which have been claimed by many linguists as the core or primitive semantic structures that motivate syntactic constructions. For instance, if the [*bei2*-R-T] form develops earlier than the canonical [*bei2*-T-R] form, would it serve to indicate that the [X causes Y to receive Z] event is cognitively less complex? Further efforts cannot be devoted to investigate the role of event cognition in children's acquisition of dative constructions.

Many thanks to Professor Gu Yang for her input on this point.



patterns is not preferred in early child language despite being iconically motivated, iconicity in itself cannot explain the lack of early preference in this case.

#### 5.4 The Input Frequency Hypothesis

The present early developmental findings also fail to support the input frequency hypothesis. The monolingual corpus findings indicate that the canonical form of the Cantonese *bei2*-dative [*bei2*-T-R] is acquired later and used less frequently than the non-canonical *bei2*-datives with a topic theme (the [T-*bei2*-R] form), a postposed theme (the [*bei2*-R-T] form), and the serial verb [*bei2*-T-*bei2*-R] form<sup>2</sup>, even though the canonical [*bei2*-T-R] form strongly outnumbers these non-canonical forms in the adult input (canonical [*bei2*-T-R] form (29.52%), non-canonical [*bei2*-R-T] form (0.27%), the non-canonical serial verb [*bei2*-T-*bei2*-R] form (1.65%), see section 3.4.2). On the other hand, it is also relevant to point out that input frequency is able to explain the early emergence of the [*bei2*-R] non-full *bei2*-dative in early child Cantonese. One might therefore suggest that the early use of *bei2*-datives is sensitive to the highly frequent and positionally salient utterance-final [*bei2*-R] sequence in the adult input. Assuming that the sample of adult input is representative in this respect, input frequency appears to play a major role here. However, input frequency cannot explain the early emergence of non-canonical *bei2*-forms: for full *bei2*-datives, non-canonical forms with preposed or postposed theme and the non-canonical serial verb form [*bei2*-T-*bei2*-R] are used earlier and/or more productively in early developmental Cantonese than their frequency in the adult input would predict. The earlier emergence of these non-canonical forms does not seem to be attributed by input frequency.

---

<sup>2</sup> The [*bei2*-R-T] form with postposed theme and the serial verb [*bei2*-T-*bei2*-R] form are used in a non-target manner from the adult language perspective however (see section 4.2.6).

## 5.5 The Input Properties Hypothesis

The present early developmental findings on the late acquisition of the canonical [*bei2*-T-R] form support the input properties hypothesis. Based on the Cantonese adult input properties established in chapter three, the canonical [*bei2*-T-R] double object form is used in a linguistic environment with null and displaced arguments, and with frequent instantiation of the [*bei2*-R] sequence. The canonical form is attested only 29.52% of the time as a result of frequently null or sometimes displaced theme arguments in natural Cantonese child-directed discourse, which often result in the recipient rather than the theme surfacing as the first post-verbal argument. Moreover, there are other cases where the verb *bei2* ‘give’ functions as a dative marker in the serial verb dative [V-T-*bei2*-R] construction in Cantonese, further reinforcing placement of the recipient immediately after the verb *bei2*, resulting in frequent [*bei2*-R] sequences (1980 tokens) which are more frequently attested than the canonical [*bei2*-T-R] form (555 tokens) in the current adult input findings. The Cantonese adult input properties do not appear to be structured in a way that makes schematization of the canonical [*bei2*-T-R] double object form straightforward for a young child from a usage-based perspective. Extraction or abstraction of a full-fledged canonical [*bei2*-T-R] structure might be slower when arguments are unexpressed or displaced and when R instead of T often surfaces after *bei2*.

## 5.6 Markedness From the UG perspective

UG-oriented theorists interpret cross-linguistically frequently attested patterns (also called unmarked patterns) as reflecting the unmarked values specified in UG. In this context, markedness is taken as deviation from the default parameter setting. It is not the goal- and is beyond the present scope- of the present thesis to argue for or against a UG view of language *in general*. This thesis also does not attempt to

compare and contrast the UG theory of language acquisition and the current usage-based theory of language acquisition adopted here. But in so far as the UG oriented hypothesis in accounting for the acquisition of datives is concerned, I will comment on Bruyn et al. (1999)'s UG based hypothesis that is relevant to the present findings. Bruyn et al. (1999) is representative in relating the notion of markedness from the UG perspective to facts attested in the acquisition of dative constructions and in creole languages. The authors cited two phenomena as supporting evidence to hypothesize that the [V-R-T] double object form is an unmarked value of UG. The two phenomena are the cross-linguistically frequent [V-R-T] form attested in Creole languages, and the early emergence of the [V-R-T] double object dative in Dutch and French children's early production. To account for the facts that DOCs *appear* to be a Creole universal and that it is acquired early in children's production, they invoked the preferred form (DOC) as an unmarked value specified in UG: Their hypothesis is cited below:

"The ease of acquisition of DOCs in Dutch and English, as well as their widespread distribution in creole languages suggest that UG provides children with DOCs as an unmarked value."  
(Bruyn et al., 1999: 363).

It is relevant to point out that the occurrence of the non-canonical [*bei2*-R-T] form in early Cantonese, despite its low frequency in the input, appears to lend support to the above universalist hypothesis that this structure is generated from UG specifications.<sup>3</sup> The current concern is whether it makes a sound case to consider these non-canonical [*bei2*-R-T] forms as derivable from the UG specifications for the present case. I do not think that it is a sound case for the following reasons. First, Bruyn et al.

---

<sup>3</sup> On the other hand, to the author's knowledge, the prepositional dative or the serial verb dative has never been posited as a UG unmarked value.

(1999)'s proposal for the universal unmarkedness of the [V-R-T] double object construction in Creole languages is recently challenged by Michaelis and Haspelmath (2003) (refer back to section 2.4.1). A more global look into Creole languages (and languages of the world) finds that DOC is not universally attested in Creole languages. Creoles in India, Indonesia and Melanesia do *not* have DOCs, regardless of their lexifiers, and there is a whole range of Portugese-based Creoles, in particular, the Asian Portuguese Creoles of India, Sri Lanka, Malaysia and Indonesia which do not show DOC.<sup>4</sup> DOC is also not predominant in the world languages, and PDC is no less frequently attested than DOC cross-linguistically as their findings showed. Second, in so far as the developmental findings in the existing literature are considered, the fact that DOC is used in children's production is subject to different explanations. For instance, Campbell and Tomasello (2001) suggested the fact that the [V-R-T] double object dative is acquired early in child English production can also be attributed to its high frequency in the input and therefore invocation of the UG theory is superfluous (refer back to section 2.3). Likewise, input frequency can also be used to explain early child French findings. Bruyn et al. (1999)<sup>5</sup> found that a French-speaking child does not use any [V-R-T] double object

---

<sup>4</sup> Michaelis and Haspelmath (2003) on the other hand argued that, as a more plausible explanation, it is rather the substrates that influenced the creation of the DOC construction in the creoles.

<sup>5</sup> Bruyn et al. (1999) however put forward a markedness hypothesis from a UG perspective. They (1999: 363) pointed out that one possible interpretation- in fact the extreme prediction- of the notion "unmarkedness" from a UG perspective is that children choose that unmarked value of the relevant parameter, *regardless of the input they receive.*", that is, children need no positive evidence to instantiate an unmarked value. If the prediction is correct, we would expect to find DOC attested in the early language of children acquiring languages in the absence of positive evidence in the input. French is an example because adult French does not have DOC, so a child acquiring French would not hear DOCs in her input. In a search of dative constructions in a French child's corpus, they however found no DOC attested. In other words, the child did not produce DOC when he did not receive DOCs in the input. In order to save their theoretical proposal, they formulated a complicated (ad-hoc) proposal putting forward their hypothesis that a theory of UG might contain some notion of markedness. They did not, however, go into details on the markedness hierarchy and the

datives in his production when the [V-R-T] double object dative is not available in the French adult input (French does not have double object construction). Hence one does not need to invoke UG in order to account for the early production of the [V-R-T] double object construction in child language.

### **5.7 The Early Preference for Non-Canonical Forms: A Functional Perspective**

The situation is somewhat different for the Cantonese case, however. The present findings on the early development of the *bei2* dative construction call for explanation(s) that should be different from the case of early English and French in the present acquisition literature. In the case of child Cantonese, we find evidence that children favor the [*bei2*-R-T] form and the [*bei2*-T-*bei2*-R] form despite their low input frequency (refer back to section 3.4.2 for the adult input findings).

The question is: why then do children acquiring Cantonese prefer these non-canonical forms which are low frequency structures in the input? The cross-linguistic unmarkedness of the non-canonical forms is *one* potential explanatory factor in accounting for the early use of these non-canonical *bei2* forms, as I have already considered in section 5.2. This would then leave open the question of why the [V-R-T] form and [V-T-dative marker-R] form are cross-linguistically frequent across the world languages. Functionally oriented theorists would hypothesize that this is because there are various functional motivations which conspire to favor certain options than others. On this view, one might consider the plausibility that the early preference of non-canonical forms in early Cantonese might reflect universal dispositions of functional motivations.

---

hypothesized relevant parameter.

I shall now consider some possible reasons for favoring the non-canonical forms over the canonical form in the early developmental context from a functional perspective. In evaluating these possibilities, it is necessary to consider them in the context of early developmental constraints on language processing capacities, such as short-term memory, and cognition load.

One possibility is that the non-canonical forms are motivated by computational parsimony. As noted in section 4.2.4.1, there are qualitative differences associated specifically with these non-canonical forms. Some early non-canonical [T-*bei2*-R] and [*bei2*-T-R] forms instantiate the [primed(given)-before-new] pattern: there is a possible relationship between the earlier utterances used by the child and how the child subsequently begins her *bei2*-dative and structures her placement of the theme. This finding suggests that the structure of early *bei2*-datives might (at least partly) be influenced by the use of lexical items and syntactic frames in the prior discourse.

Syntactic priming has been related to processing constraints in the existing literature on adult language. Wasow and Arnold (2003) also suggested that syntactic priming might be one factor influencing the ordering of constituents in adult language: items or syntactic frames that have been used in the prior discourse are more accessible and therefore easier to produce early in the utterance. Moreover, the authors suggested that the desire to produce given, accessible information earlier than new inaccessible information is one underlying mechanism in influencing ordering preferences in adult language. In addition, the authors (2003: 12-13) related the order of constituents to processing constraints. They suggested that “constituent ordering is influenced by constraints on planning and production: speakers tend to begin their utterances with constituents that are easier to produce, and save the more difficult


constituents for later in the utterance.”. If these factors also apply similarly to early child language, young children with limited cognitive resources on planning and production as well as a limited repertoire of constructions would prefer the given(primed)-before-new pattern: the theme denoting a given referent primed in the prior discourse tends to be more accessible and facilitates planning (production) of the [T-*bei2*-R] forms, and the [*bei2*-R] primed in the prior discourse tends to be more accessible and facilitates planning (production) of [*bei2*-R-T] form. All these structures are more easily produced when coping with on-line exigent communicative demands than planning with entirely new referents and syntactic frames for young children subject to developmental cognitive/ processing constraints of various kinds and a limited repertoire of constructions in their mental grammar. The observed [primed-before-new] pattern also suggests possible scaffolding relationships between the non-canonical *bei2*-forms and the child’s utterance in preceding immediate discourse. One might hypothesize that when young children lack the target canonical [*bei2*-T-R] form in their developmental repertoire and have developmental constraints on planning and production, they prefer the non-canonical structures which are computationally parsimonious and efficient, and can be scaffolded in natural discourse (given the frequent use of [*bei2*-R]). This proposal seems plausible, in view of computational parsimony, because planning with primed expressions requires less resources than planning with entirely new ones.

Another related perspective is the performance theory of Hawkins (1994). Here the notion of his constituent recognition domain (CRD), which is related to processing constraints/ efficiency, is relevant.

Constituent Recognition Domain (CRD): The CRD for a phrasal mother node M consists of the set of terminal and non-terminal nodes that must be parsed in order to recognize M and all immediate constituents of M. (Hawkins, 1994: 58)

In the case of the double object form, the recognition domain for the VP is the distance between the verb *bei2* and the N of the second NP, crossing over the first NP.

The double object [*bei2*-NP1-NP2] form:  $[_{VP} \textit{bei2} [_{NP} N] [_{NP} N]]$



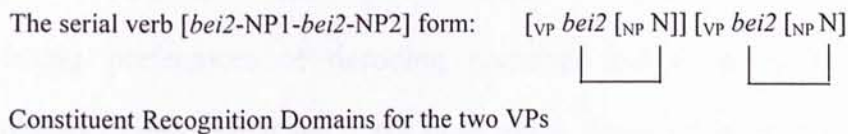
Constituent Recognition Domain for VP

Within the double object construction, I now turn to compare the recognition domain between the theme-recipient postverbal ordering and the recipient-theme ordering, i.e. the recognition domain for  $[V-NP_{\text{Theme}}-NP_{\text{Recipient}}]$  and  $[V-NP_{\text{Recipient}}-NP_{\text{Theme}}]$ . For double object construction, since NP1 is an immediate constituent, the shorter the NP1, the shorter delay its length incurs in the parsing of the second NP. As mentioned in section 1.1.2, adult speakers use the non-canonical [*bei2*-R-T] form when T is long. From a functional perspective, the motivation to postpone the theme in this case is similar to that in heavy NP shift (see Arnold et al., 2000): by postponing the ‘heavy’ theme argumental NP to the right after the recipient, the constituent recognition domain (Hawkins, 1994) for VP is shortened for efficient parsing, because the shorter recipient NP now precedes the longer theme NP.

In the case of the non-canonical [*bei2*-T-*bei2*-R] form in adult Cantonese, as described in section 1.1.3 in chapter one, its uses are identified under the following two situations. One situation is that the theme is long. Another situation is that the adult speaker wants to emphasize the second NP as playing the recipient role. In this case, the speaker would ‘mark’ the recipient NP with a *bei2* dative marker



(sometimes even with an added emphasis on the *bei2* dative marker) using a serial verb [*bei2*-T-*bei2*-R] form, even when the theme NP is not particularly long. Similar to the first condition mentioned above, the verb *bei2* ‘give’ is used as a goal-marking dative marker introducing the entity as the intended recipient. Applying the notion of CRD to the case of the non-canonical [*bei2*-T-*bei2*-R] serial verb form, there are two VPs, VP1 and VP2. The recognition domain for VP1 is the distance between the verb *bei2* and the N of the first NP, while the recognition domain for VP2 is the distance between the verb (V) *bei2* and the N of the second NP:



The recognition domain is thus shorter relative to the case of using the canonical [*bei2*-T-R] double object form, because the recognition domain for VP in the double object form would involve two postverbal object NPs, instead of one object NP for each VP in the case of using the [*bei2*-T-*bei2*-R] serial verb form. From a processing perspective, using the [*bei2*-T-*bei2*-R] serial verb form therefore has the function of facilitating parsing, in this case, shorten the constituent recognition domain (CRD) (see also the Principle of Domain Minimization proposed in Hawkins, 2001).

We have seen that from an adult language perspective, the use of the non-canonical [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms in marked contexts can shorten the recognition domain for VP, in Hawkins’ (1994) terms. Hawkins (1994) would claim that there are processing constraints or efficiency considerations to shorten the recognition time for the phrasal constituents in young children in these marked contexts when adults

are engaging in on-line language use. From an early developmental perspective, the characteristic that the non-canonical forms are associated with a shorter recognition domain is compatible with the view on early processing constraints.

Recall also the proposal that the canonical [*bei2*-T-R] double object form may be motivated by the functional principle of economy, as pointed out in section 2.5. From a functional perspective, the canonical and the non-canonical forms of the Cantonese *bei2*-dative have distinct functional motivations. From a developmental point of view, one might ask how these functional motivations are relevant in an early developmental context. We do not know to what extent concerns about the processing preferences of decoding (parsing) and economy of encoding are manifested in child language. We need much future research to properly address this question. But it is reasonable to conceive that at an early developmental stage, concerns over how to live with *processing constraints* and how to put a restricted set of constructions into *optimal use*, might mean more to the child than to *make economical use* of her grammatical resources, which might come later in a developmental context. The late use of the canonical [*bei2*-T-R] double object form, which is motivated primarily because of economy (see section 2.5, see also Liu, 2001), is consistent with the idea that the functional force to omit the *bei2* dative marker might come into operation at a later stage, while the early preference for the non-canonical forms is compatible with a role for early processing constraints, such as shortening the recognition time for the phrasal constituents.

## 5.8 The Source of the Early Non-Canonical *bei2*-datives: A Usage-Based Perspective

I have pointed out that the non-canonical forms are both motivated on typological and other functional grounds. But where do these structures come from? How are they generated? In section 5.6, I have pointed out the problems with Bruyn et al. (1999)'s UG proposal. The next question at issue is, if we work within the scope of the usage-based approach to child language acquisition, which assumes that storage is determined solely by aspects of use, not by the existence of pre-wired innately given linguistic structures (knowledge), where then do these non-canonical *bei2* forms come from? How are they generated? I shall now consider the possible sources for the generation of the early non-canonical *bei2*-datives from a usage-based perspective.

### 5.8.1 The Early [*bei2*-R-T] Form

Earlier on in sections 4.2.4.1 and 5.7, I have pointed out that some non-target [*bei2*-R-T] forms attested in the present data might be scaffolded from a prior use of a non-full [*bei2*-R] form in discourse, although we have yet to spell out exactly how the scaffolding is done. From the usage-based perspective, there are other possible ways to generate a non-canonical [*bei2*-R-T] expression. In what follows I shall also point out and consider these possibilities.

#### 5.8.1.1 Against Learning Directly From The Adult Speech Models

The first possibility is that the children imitatively learn the non-canonical [*bei2*-R-T] expressions from the adult speech models. However, all the early non-canonical [*bei2*-R-T] expressions attested in our developmental findings were spontaneously produced by our child subjects without prior use of an adult [*bei2*-R-T] utterance,

suggesting that the early non-canonical [*bei2-R-T*] expressions attested in the corpus were not imitatively learnt (at least immediately) from the adult speech models.

#### **5.8.1.2 Against Generating Directly From The [*bei2-R-T*] Verb Specific Schema**

The second possibility is that the children generated the [*bei2-R-T*] expressions out of a verb specific [*bei2-R-T*] schema. This possibility presumes that the schema has emerged in the mental grammar. In principle, the [*bei2-R-T*] schema could be abstracted out of repeated exposures to the [*bei2-R-T*] expressions directly in the adult input. Now recall our adult input findings show that adults used this non-canonical form only 0.27% of the time in their speech to young children acquiring Cantonese. Assuming that our corpus finding on the adult input is representative of the kind of adult input our young child subjects were exposed to, the present findings indicate that the [*bei2-R-T*] form is barely available to young children in their primary linguistic data. Young children, however, used far more [*bei2-R-T*] forms than the adults in terms of percentage use by token measures than they actually heard in their primary linguistic input. Schematization of a [*bei2-R-T*] verb specific schema directly from the input seems unlikely, given the minimal use of [*bei2-R-T*] expressions instantiated in child-directed speech, not to mention the fact that schematization takes time and is itself a gradual and emergent process (Tomasello, 2003).

#### **5.8.1.3 Against Overgeneralizing the Abstract [V-R-T] Schema**

The third possibility is that the children might overgeneralize the abstract [V-R-T] schema (which occurs with the ‘teach’ verbs in adult Cantonese) to the Cantonese verb *bei2* ‘give’, thereby yielding the non-target [*bei2-R-T*] forms in their early

language. I will show that this possibility is unlikely. In order for this hypothesis to be tenable, we need evidence to suggest that the verb-general [V-R-T] schema has emerged in those young children's repertoires in order for them to apply it to the verb *bei2* 'give'. In a search of the adult input data in CANCELP, I found no empirical evidence from the adult input data that the Cantonese child directed speech contains frequent exemplars in the [V-R-T] form to facilitate early schematization of the verb-general abstract schema (see section 5.8.1.2 for similar argumentation). Even if, for the time being, we put aside the questionable issue of whether the verb-general abstract [V-R-T] schema can emerge at such an early age, overgeneralization assumes a certain degree of productivity with the [V-R-T] schema- or else it cannot be productively extended to the other verbs. Therefore, *if* what those children were really doing at the time of producing the non-canonical [*bei2*-R-T] expressions was overgeneralizing a verb general abstract [V-R-T] schema to the verb *bei2*, one should expect to find evidence of productive use of the [V-R-T] forms attested in those children's early speech with other Cantonese verbs prior to or together with the age of first use of the non-target [*bei2*-R-T] form. This is a testable prediction. I therefore checked all the uses of the Cantonese verbs *gaau3* 'teach' and *man6* 'ask'<sup>6</sup> in each child's early production from CANCELP, and found no evidence in our current developmental data that the children who used the non-target [*bei2*-R-T] forms showed any signs of productively using the [V-R-T] forms- not even a single instance of [*gaau3*(teach)-R-T] or [*man6*(ask)-R-T] expression was found from all the child utterances in CANCELP. So far there was no supporting empirical evidence to suggest that the young children who used the non-target [*bei2*-R-T] forms in their early speech were also productively using [V-R-T] expressions with

---

<sup>6</sup> These two verbs were chosen among the Cantonese 'teach' verbs (see table 1.1) because they are considered to be the most commonly used verbs of this type in colloquial speech.

other verbs.

On the other hand, I also acknowledge the fact that we cannot rely on corpus data alone to ascertain the question of productivity. Whether young children who use the non-target [*bei2*-R-T] forms also show productivity in using [V-R-T] expressions with other verbs is yet to be confirmed with further experimentation like elicited production tasks (see chapter seven for suggestions on further research).

Furthermore, the overgeneralization hypothesis considered here would not gain support from the current acquisition literature. Tomasello (2000) has pointed out that the overgeneralization patterns so far reported in the current literature for child English are evident at only age three or beyond – but never before age three. Tomasello (2003) suggested that early child language before age three or three-and-a-half is better characterized in terms of locally structured lexically-specific constructions with the absence of abstract verb-general schemas.

## **5.8.2 The Early [*bei2*-T-*bei2*-R] Form**

The considerations here are essentially similar to the case of [*bei2*-R-T]. I consider two possibilities of generating a [*bei2*-T-*bei2*-R] expression from a usage-based perspective.

### **5.8.2.1 Against Learning Directly From The Adult Speech Models**

As in the case of early non-target [*bei2*-R-T] expressions, all the early non-target [*bei2*-T-*bei2*-R] expressions attested in our developmental findings were spontaneously produced by our child subjects, suggesting that they were not the result of imitating the adult speech models.

### 5.8.2.2 On Overgeneralizing The [V-T-*bei2*-R] Schema

Another possible source of generating a [*bei2*-T-*bei2*-R] expression is from a more abstract [V-T-*bei2*-R] schema. That is, I consider whether the early non-target [*bei2*-T-*bei2*-R] forms were instances produced by overgeneralization of the [V-T-*bei2*-R] schema (which is conventionally associated with the Cantonese ‘send’, ‘fry’ and ‘pluck’ verbs in adult Cantonese (see table 1.2)) to the verb *bei2* ‘give’ in a non-target manner. In order for this speculation to be tenable, we need evidence to suggest that the [V-T-*bei2*-R] schema has emerged in those young children’s repertoires for them to overgeneralize the [V-T-*bei2*-R] schema to the verb *bei2*. In a search of the adult input data in CANCORP, I found many more tokens of [V-T-*bei2*-R] expressions attested in child directed speech than in the case of [V-R-T] expressions discussed earlier in Section 5.8.1.3. In this respect, the finding suggests that at least the Cantonese child-directed speech contains frequent exemplars instantiating the [V-T-*bei2*-R] form to *facilitate* schematization of the [V-T-*bei2*-R] schema directly from the adult input, although schematization itself is a gradual and emergent process and schematization of the complex [V-T-*bei2*-R] schema itself might be a complex task for young children with developmental constraints. For the time being, I put aside the issue of whether the abstract [V-T-*bei2*-R] schema *can* in fact emerge at such an early age even with support from the input properties, and proceed to explore whether there is any evidence for a [V-T-*bei2*-R] schema in those children who used the non-target [*bei2*-T-*bei2*-R] forms. If those children were generating the [*bei2*-T-*bei2*-R] expressions by overgeneralizing the [V-T-*bei2*-R] schema, they should show early productive use of [V-T-*bei2*-R] expressions. I therefore looked for such evidence from the existing corpus data available. I checked all the uses of datives with other Cantonese verbs by MHZ, CGK and LLY, who showed the use of at least one instance of [*bei2*-T-*bei2*-R] expression in their

speech, before and during their first use of a [*bei2-T-bei2-R*] expression as attested in the corpus. For MHZ, there was not a single use of [*V-T-bei2-R*] expression attested before and during his first use of a [*bei2-T-bei2-R*] expression in the corpus. For CGK, there was one token of [*lo2(get)-T-bei2-R*] attested at 2:02.28 just a few days before her first use of a [*bei2-T-bei2-R*] expression at 2:03.04, and only one token of [*maai5(buy)-T-bei2-R*] expression attested within the same transcript at 2:03.04. For LLY, there was only one instance of [*maai5(buy)-T-bei2-R*] expression attested within the same transcript in which she first used the non-canonical [*bei2-T-bei2-R*] form. Two out of the three children who had used at least one non-target [*bei2-T-bei2-R*] expression showed evidence of using [*V-T-bei2-R*] expressions with other verbs at the same time of producing non-target [*bei2-T-bei2-R*]. The findings seem to be consistent with the hypothesis at issue, although there is so far no strong evidence from these children's naturalistic corpus data to show signs of productively using the [*V-T-bei2-R*] expressions with different verbs.

The occurrence of the [*bei2-T-bei2-R*] form is also interesting in terms of the relationship between the lexically specific [*bei2-T-R*] double object construction and the productive [*V-T-bei2-R*] serial verb construction. Under the null dative marker hypothesis in the current literature (refer back to section 2.5), the target [*bei2-T-R*] canonical double object form can be seen as closely related to the productive [*V-T-bei2-R*] serial verb dative construction, where *bei2* is used as V and the dative marker *bei2* in [*V-T-bei2-R*] is omitted. The omission can be attributed to phonological identity as proposed by Tang (1998), or functional economy as proposed by Liu (2001) and in the present account (see section 2.5). Note that we do find evidence that children are producing [*V-T-bei2-R*] with other verbs at the



time of producing the non-target [*bei2-T-bei2-R*] form, though we do not know from the naturalistic corpus analysis how productive the use of the [*V-T-bei2-R*] construction is (see section 6.3.1 for suggestions on further study). The non-canonical [*bei2-T-bei2-R*] form found in early developmental Cantonese might therefore be plausibly seen as an instance of the [*V-T-bei2-R*] serial verb construction with *bei2* used as V. In this case, from the adult language perspective, for reason of functional economy or for whatever reason, the child has failed to omit (or has not yet been driven by the same functional motivation to omit) the dative marker *bei2*. This interpretation is consistent with the null dative marker hypothesis.

## 5.9 Remaining Questions

There are still other questions to be resolved. Investigations into the input properties suggest that the Cantonese language specific input properties are not conducive to early schematization/ extraction. If schematization of the canonical [*V-T-R*] from the input is delayed, what drives its subsequent successful schematization?? The input properties also raise questions of learnability for the schematization hypothesis in the usage-based theory of language acquisition. To what extent can the child schematize the canonical [*V-T-R*] structure from the input given the prevalence of unexpressed or displaced arguments?

At a later point of development, the target canonical [*bei2-T-R*] form is then acquired by these children, as observed in the developmental data of their older siblings. Exactly what would happen to these transient early non-target forms, or whether these non-target forms would be dropped once the canonical form is acquired, is not known at present- we only have longitudinal data available to at most around age

3;08 at present for monolingual children (see section 4.2.6.1, on the other hand, for the protracted use of the non-target [*bei2*-R-T] form in bilingual children).

This study also raises the question of how we should evaluate the role of input frequency in early acquisition. If we evaluate input frequency based only the relative frequencies of the canonical [*bei2*-T-R] form versus the non-canonical [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms, input frequency fails to explain the earlier emergence of the non-canonical forms than the canonical form, because the non-canonical forms are much more infrequent than the canonical form. However, if we consider the frequency of the [*bei2*-R] sequence in the adult input, input frequency is able to explain the early emergence of the [*bei2*-R] non-full dative as section 4.2.1 shows. In this respect, the importance of frequency effects cannot be discounted in accounting for the early emergence of the [*bei2*-R] form. In addition, one can reasonably argue that the frequent recurrence of the [*bei2*-R] sequence resulting from null or displaced theme in the context of a set of related serial verb constructions *bei2* used in the child-directed adult speech might actually motivate the [*bei2*-R-T] ordering. From a usage-based perspective, the high co-occurrence of *bei2* and R on surface syntax in naturalistic child-directed Cantonese speech gives rise to a stronger sequential connection (or a tight sequential link, see Bybee, 2002) between *bei2* and R than between *bei2* and T. These frequently occurring [*bei2*-R] sequences are likely to be extracted, schematized, stored independently and accessed as a single unit in early developmental mental grammar. As I pointed out in section 5.7, some early non-canonical [*bei2*-R-T] forms might be scaffolded by the prior use of the [*bei2*-R] non-full datives in natural discourse.<sup>7</sup> One might therefore argue

---

<sup>7</sup> Although we also need to spell out in more sophisticated terms how the scaffolding of more complex structures is done by young children. I leave this issue for further research to accomplish.

whether input frequency plays a facilitative role in motivating the early use of the [bei2-R-T] form based on the sequential strength between *bei2* and R from a usage-based perspective.

The following questions are also worth considering when one works within the usage-based theory: which part of the input data children are initially sensitive to given their limited developmental working memory? What forms the ‘effective input’ for early schematization? And how can we make use of independent evidence on children’s working memory?

## 5.10 Chapter Summary

In this chapter, I have done the following things. I have first reviewed in section 5.1 the empirical predictions for the four hypotheses (the markedness hypothesis, the iconicity hypothesis, the input frequency hypothesis and the input properties hypothesis) I have established earlier in this thesis. In sections 5.2 to 5.5, I have related the present findings to these four hypotheses. Counter to the predictions of the iconicity hypothesis and the input frequency hypothesis, the canonical form [bei2-T-R] proves to be dispreferred in early developmental Cantonese *production*, as the markedness and the input properties hypotheses predict. This is consistent with the idea that acquisition of a construction can be facilitated in two ways: i) if it conforms to universally shared functional or cognitive principles motivated on typological grounds; and ii) if it is facilitated by language-specific input properties.<sup>8</sup> The developmental preference for the non-canonical forms over the canonical form

---

<sup>8</sup> There is one related question: if the universally shared principles motivated on typological grounds reflect innate (general) tendencies young children might share, how do these general principles interact with children’s incomplete but growing language-specific knowledge of Cantonese?

also supports the markedness hypothesis from a cross-linguistic perspective.

I have also attempted to discuss a number of interesting issues that come out of the present findings, although they were not set out among the original goals for undertaking this study. In section 5.6, I have pointed out the potential relevance of Bruyn et al. (1999)'s markedness hypothesis from the UG perspective to the present developmental findings obtained. In section 5.7, I have pointed out that the early preference for the non-canonical forms might be functionally motivated, notably for reasons of processing constraints. Working within the usage-based theory, I have considered in section 5.8 the possible source(s) for generating the non-canonical [*bei2*-R-T] and the [*bei2*-T-*bei2*-R] forms. One possibility considered is that the early non-canonical *bei2*-datives might be generated by overgeneralizing a more abstract verb general constructional schema. This possibility is unlikely for the case of the [*bei2*-R-T] form, but remains plausible for the case of the [*bei2*-T-*bei2*-R] form. Finally, in section 5.9, I have raised some remaining questions.

## Chapter Six. Conclusions and Further Research

### 6.0 Introduction

This chapter concludes the present study and suggests further research work. Section 6.1 states the principal conclusions. Section 6.2 highlights the contributions of this study. Section 6.3 discusses how the current findings can inspire future research.

### 6.1 Principal Conclusions

The following principal conclusions conclude the present study. At an empirical level, we so far see no strong evidence to suggest that young children before age three have acquired the canonical [*bei2-T-R*] form based on the present findings. The present monolingual child findings show either non-use or only a few tokens of early inconsistent use of this form across communicative situations that demand its use as the target structure among children before age three. Moreover, children acquiring Cantonese in different acquisition contexts (monolingual, bilingual and clinical) exhibit early non-target use of the [*bei2-R-T*] double object form and the [*bei2-T-bei2-R*] serial verb form. At a theoretical level, the markedness hypothesis predicts correctly the earlier emergence of the non-canonical forms than the canonical [*bei2-T-R*] form, while the input frequency hypothesis and the iconicity hypothesis do not. In addition, both the markedness hypothesis discussed in O'Grady (2000) and the input properties hypothesis investigated in this thesis predict correctly the late acquisition of the canonical [*bei2-T-R*] double object form. If the present finding on the late acquisition of the canonical [*bei2-T-R*] form is not an artifact resulting from the lack of opportunities for using this form by our young

child subjects in the present naturalistic data, we need to account for this developmental finding. The cross-linguistic markedness of the [V-T-R] double object form and the Cantonese input properties considered in this thesis might both contribute to the late acquisition of the canonical double object [*bei2*-T-R] form in Cantonese.

## 6.2 Contributions

Double object constructions (DOC) form an area that figures in much acquisition research, where the theoretical interests and empirical focus have been skewed towards the [V-Indirect Object(IO)-Direct Object(DO)] double object construction in English (see for example Pinker, 1989 and Gropen et al., 1989 for child L1 acquisition; Mazurkewich, 1984 and Wolfe-Quintero, 1994 for adult L2 acquisition), but research on the early acquisition of the [V-DO-IO] double object construction, to the author's knowledge, has not been documented in the acquisition literature. The canonical double object form with the verb *bei2* 'give' in Cantonese exemplifies a case of this [V-DO-IO] double object construction. This thesis adds knowledge to the acquisition of this type of double object construction, a missing piece in the literature. I have also related the acquisition of language particular facts in Cantonese to cross-linguistic tendencies. The canonical [*bei2*-T-R] double object form, called inverted double object construction in Tang (1998) and pseudo double object construction in Liu (2001), is marked both cross-linguistically and within Cantonese, and so is predicted to be acquired late by the markedness hypothesis. The present findings show that the canonical [*bei2*-T-R] double object form is indeed *dis*-preferred in early Cantonese production. In addition, the early use of non-canonical forms [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms in early child Cantonese provide strong evidence for parallels between typological distribution and

developmental preferences. This thesis documents new facts that call for more cross-linguistic investigations of the acquisition of the marked [V-T-R] double object construction (see section 2.4.1). This thesis also establishes an initial set of empirical facts about the acquisition of this Cantonese *bei2* 'give' dative construction in the context of many unknown facts about this area. Findings from the normal developing monolingual children serve as important baseline for drawing comparisons with the bilingual and clinical population. Speech therapists should be aware that even normal developing children before age three have early non-target forms with this dative construction.

The theoretical significance of the present findings is as follows. The early occurrence of the non-canonical [*bei2*-T-*bei2*-R] serial verb form is consistent with the analysis of the [V-T-R] double object dative with a null dative marker. The findings on adult input properties raise important questions for schematization in the usage-based theory of language acquisition (Tomasello, 2003). Given that schematization is one major concern of the usage-based theory (Tomasello, 2003), if the theory is to be of universal impact and to be viable for the study of early child Cantonese and Chinese languages in general, it has to factor in and account for how children acquiring these languages are able to abstract concrete expressions into constructional schemas from the input with the language-specific properties discussed in chapter three (null arguments, displaced arguments). I leave this issue for further discussion at this point.

### **6.3                    Suggestions for Further Research**

The present findings based largely on production data from naturalistic corpus analyses also create various openings for further research. Building on the present

study, future research can go along the following lines:

### 6.3.1 Elicited Production Studies

As pointed out in section 5.8.1.3, the possibility that the early non-canonical [*bei2*-R-T] forms attested in the present data were generated by over-generalizing a [V-R-T] schema seems unlikely. We need elicited production studies to confirm whether there is really no correlation between the non-target production of [*bei2*-R-T] and the productive use of the [V-R-T] expressions with other verbs. This can be done by first using elicitation production tasks probing the uses of full *bei2*-datives, when [*bei2*-R-T] forms are noted, we then look for the child's ability to use the [V-R-T] double object construction with the Cantonese 'teach' verbs (see table 1.1 for examples of Cantonese 'teach' verbs). In order to address properly whether a verb general [V-T-*bei2*-R] schema exists or not, we can conduct experiments testing novel verbs.

On the other hand, as pointed out in section 5.8.2.2, there is a possibility that the early non-canonical [*bei2*-T-*bei2*-R] forms attested in the present data were generated by overgeneralizing a [V-T-*bei2*-R] schema. I currently cannot obtain strong evidence from the naturalistic corpus data for this hypothesis, however. We need elicited production studies to further evaluate this hypothesis. If this hypothesis is right, one would expect to see a correlation between the non-target production of [*bei2*-T-*bei2*-R] form and a productive use of the [V-T-*bei2*-R] expressions with other verbs. This can be done by first using elicitation production tasks probing the uses of full *bei2*-datives<sup>1</sup>, when [*bei2*-T-*bei2*-R] forms are noted,

---

<sup>1</sup> A pilot study has been done on a 2;06 child to probe for more full *bei2*-datives. The results obtained from the pilot elicited production task conform to the results obtained from the corpus study



we can then look for the child's ability to use the [V-T-*bei2*-R] serial verb construction with a range of other Cantonese verbs like the 'send', 'fry' and 'pluck' verbs (refer back to table 1.1). Again, in order to address properly whether a verb general [V-T-*bei2*-R] schema exists or not, we can conduct experiments testing novel verbs.

If we can gather such evidence, the next step is to ask what motivates the overgeneralization to occur, and whether there are constraints applying to the semantic class of verbs that have not yet been acquired.

Cantonese-English bilingual children are also found to produce non-target [*bei2*-R-T] forms in their early speech. One area to evaluate the possible influence from English is to check whether bilingual children are productive in using the [V-R-T] double object construction in their English by looking at their English corpus data and by elicited production tasks when they are using the non-target [*bei2*-R-T] forms in their developing Cantonese.

### 6.3.2 Comprehension Studies

We do not have comprehension data so far. It would be interesting to do comprehension studies to see whether there is any discrepancy between comprehension and production as observed in the English double object construction (see O'Grady, 1997: 208-213). As in the comprehension studies on the English

---

in terms of the range of *bei2* structures obtained. The child used the [*bei2*-T] and [*bei2*-R] non-full datives, and also non-target [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms, but failed to use the canonical [*bei2*-T-R] form spontaneously despite various kinds of discourse encouragement, for instance, using modeling of the target canonical structure and probing questions in the form of the canonical structure. Due to time constraints, the elicitation procedures have not been standardized and extended to more children.

double object construction, we have to control for the animacy of the theme and the recipient argument. There are three hypotheses which are relevant in the context of comprehension studies. The first hypothesis concerns how the Cantonese input properties might impact on children's word order learning of the two postverbal objects. Would the frequent [*bei2*-R] sequence attested in the input affect children's early hypothesis on the postverbal ordering of the theme and recipient of the Cantonese *bei2*-dative? Recent research has demonstrated that infants as young as eight months old are sensitive to the distributional properties of the input (Saffran, et al., 1996). If a child acquiring Cantonese is sensitive to the distributional properties of the input and is engaging in some sort of statistical analysis of the input data in learning the word order of the Cantonese *bei2*-dative, with the child-directed Cantonese speech being heavily skewed in instantiating the recipient rather than the theme as the first post-*bei2* argument, there might be some anticipatory effects of predictability (Jurafsky et al., 2001). One might wonder to what extent such frequency bias in the adult input might influence the early acquisition of the postverbal ordering of the theme and recipient of the Cantonese [*bei2*-T-R] canonical double object form. If children were sensitive to the [*bei2*-R] sequence frequently attested in the input, they would prefer the [V-R-T] interpretation, assigning the first post-verbal NP as the recipient role.

The second hypothesis to consider is that children might make use of the so-called extended canonical sentence strategy proposed in the English dative acquisition literature (see O'Grady, 1997: 212 and Pinker, 1989: 401). This hypothesis was proposed to account for why the English-speaking children tended to reverse the recipient-theme order in the English double object construction to the theme-recipient order in comprehension act-out tasks. The idea of the extended

canonical sentence strategy is as follows. The English transitive construction is in the canonical SVO word order, in which the subject is associated with the agent, the verb the action, and the object the theme. One might consider children might extend this association between thematic roles and grammatical relations found in many simple transitive patterns to interpret the dative test sentences, assigning the first NP as the theme, the second NP as the recipient, and thereby reversing the recipient-theme sequence in the double object construction (e.g. show the giraffe the bear) to the theme-recipient order. Since the basic canonical word order for the transitive construction is also SVO in Cantonese, it is relevant to consider whether children acquiring Cantonese might make use of the extended canonical sentence strategy to interpret the test sentences, when they are being assessed in comprehension act-out tasks. If the child subjects make use of the extended canonical sentence strategy, they would prefer the [V-T-R] interpretation in comprehension tasks.<sup>2</sup>

The third relevant hypothesis is called the explicitness hypothesis (see O'Grady, 2000). The explicitness hypothesis states that "children prefer structures in which semantic roles are explicitly indicated" (O'Grady, 2000: 4). Similar to the iconicity hypothesis (refer back to section 2.2), the explicitness hypothesis is tested by means of comprehension studies in which both the DO and IO are full NPs and animacy is controlled, so that animacy and pronominal versus full NP structure cannot be used as cues to aid comprehension/ parsing of the structure. This hypothesis predicts that the comprehension of the preposition-less double object dative construction, in the

---

<sup>2</sup> We also need adult data acting as a control group to see which interpretation they prefer, since according to Tang (1998), even native adult Cantonese speakers differ in their interpretations when both NPs are animate (although we do not know whether there is any statistical preference).

absence of other cues like animacy and pronouns, would be more difficult to process because the semantic roles are not explicitly indicated as in the case of the prepositional datives. Applying the explicitness hypothesis to the Cantonese *bei2*-dative, this hypothesis would predict that the [*bei2*-T-R] canonical double object form would be difficult for young children in comprehension tasks when we control the animacy of the theme and the recipient objects, because the semantic roles are not explicitly indicated by any overt marking in the present case. I do not investigate the explicitness hypothesis in the present study because I focus on naturalistic production data, in which it is very likely that young children would make use of other cues like animacy in aiding their comprehension when encountering this construction in their naturalistic input.

The early developmental findings also call for investigating whether children know that the early [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms they use prior to their mastery of the target canonical form are non-target-like. In situations that require the canonical [*bei2*-T-R] form as the target structure from the adult language perspective, can young children before age three acquiring Cantonese display the knowledge of grammaticality judgment that the canonical form is the target form and the [*bei2*-R-T] and [*bei2*-T-*bei2*-R] forms are non-target? One might also wonder whether the input properties would actually pose ambiguity to young children acquiring Cantonese with respect to the proper placement of the theme, given that the theme is either often missing or sometimes displaced in the adult input. Whether such input properties would give rise to early developmentally misguided word order hypotheses and whether young children have the knowledge that the canonical form [*bei2*-T-R] is the felicitous form in a particular discourse context requires further experimental investigation at the comprehension level such as the grammaticality

judgment tasks in the future.

### 6.3.3 Cross-Linguistic Investigations

This study calls for cross-linguistic research on the early acquisition of the [V-T-R] double object constructions in other languages. Languages like Thai, Ewe and other Chinese dialects are possible candidates. Thai and Ewe exhibit a more productive [V-T-R] pattern across verbs. One might wonder whether children acquiring Thai and Ewe have less problem in acquiring [V-T-R] than the Cantonese-speaking children. Moreover, the null dative marker hypothesis works well in Thai (Matthews and Leung, 2002). One might wonder whether children acquiring Thai use the dative marker form before the [V-T-R] form.

In this study, I hypothesize that the input properties might play a role in delaying the acquisition of the canonical [*bei*2-T-R] double object form (see sections 3.5 and 5.5). To evaluate the role of input properties, we have yet to find languages which have [V-T-R] attested with input properties that facilitate its early acquisition, for instance, the language is not a pro-drop language so arguments are often overtly expressed, the language has relatively fixed word order so arguments are often undisplaced, and the language has [V-T-R] structures with different verbs attested, to see whether early acquisition of [V-T-R] is attested before age three. If it turns out that young children acquiring the language still disprefer [V-T-R] despite possible encouragement from the input properties and produce other non-canonical forms like [V-R-T], [V-T-dative marker-R] prior to the acquisition of [V-T-R], cross-linguistic markedness (and other factors) *must* be at work to *dis*-prefer the early production of the [V-T-R] double object construction.

## Appendix One. The non-target use of the [bei2-R-T] form in monolingual Cantonese children from CANCORP (Lee et al 1996)

1. CGK, 2;03.11 \*\*\* File "20311.cha": line 1463.  
CHI: Bei2 ngo5 doi6 aa1 .  
Give me bag PRT  
'Give me bag.'
2. CGK, 2;03.11 \*\*\* File "20311.cha": line 1466.  
CHI: Bei2 gaalkei4 go3 doi6 aa1 .  
Give *gaalkei4* CL bag PRT  
'Give *GaaKei* (the child's name) the bag.'
3. CGK, 2;03.11 \*\*\* File "20311.cha": line 1469.  
CHI: Bei2 gaalkei4 doi6 aa1 .  
Give *gaalkei4* bag PRT  
'Give *GaaKei* (the child's name) bag.'
4. CGK, 2;03.11 \*\*\* File "20311.cha": line 1493.  
CHI: Jat1zan6 bei2 suk1suk1 cin2 aa1 .  
Later give uncle money PRT  
'Later give uncle money.'
5. CGK, 2;03.11 \*\*\* File "20311.cha": line 1928.  
CHI: Ngo5 jiu3 aa1 bei2 suk1suk1 cin2 aa1 .  
I want PRT give uncle money PRT  
'I want PRT give uncle money.'
6. CGK, 2;04.08 \*\*\* File "20408.cha": line 2072.  
CHI: Ngo5 jat1zan6 bei2 lei5 je5sik6 aa3 .  
I later give you something edible PRT  
'I later give you something edible.'
7. CGK, 2;04.30 \*\*\* File "20430.cha": line 3262.  
CHI: Hai6 aa3, ngo5 bei2 suk1suk1 di1 cin2.  
Yes PRT I give uncle CL money  
'Yes, I give uncle some money.'

8. CGK, 2;05.03 \*\*\* File "CGK20503.cha": line 520.  
 CHI: Ngo5 bei2 suk1suk1 # di1 cin2 aa1 .  
 I give uncle CL money PRT  
 'I give uncle some money.'
9. CGK, 2;09.09 \*\*\* File "20909.cha": line 3272.  
 CHI: La4 , ngo5 bei2 lei5 go3 sing1sing1 .  
 PRT I give you CL star  
 'I give you the star.'
10. LTF, 2;03.30 \*\*\* File "20330.cha": line 2851.  
 CHI: Ngo5 xxx bei2 ngo5 nei1 go3 aa3 .  
 I give me DET CL PRT  
 'I xxx give me this one.'
11. LTF, 2;03.30 \*\*\* File "20330.cha": line 2854.  
 CHI: Bei2 ngo5 nei1 go3 aa3 .  
 Give me DET CL PRT  
 'Give me this one.'
12. LTF, 2;03.30 \*\*\* File "20330.cha": line 4536.  
 CHI: Bei2 ngo5 nei1 bun2 aa3 .  
 Give me DET CL PRT  
 'Give me this one.'
13. LTF, 2;09.07 \*\*\* File "20907.cha": line 4777.  
 CHI: Bei2 ngo5 go2 di1 aa3 .  
 Give me DET CL PRT  
 'Give me these.'
14. LLY, 2;11.01 \*\*\* File "LLY21101.cha": line 2187.  
 CHI: Bei2 aa3saa1 # sing1sing1 aa1 .  
 Give aa3saa1 star PRT .  
 'Give AaSaa star.'
15. LLY, 2;11.08 \*\*\* File "21108.cha": line 4206.  
 CHI: Bei2 ngo5 "Hello-Kitty".  
 Give me Hello-Kitty  
 'Give me Hello-Kitty.'

16. LLY, 2;11.29 \*\*\* File "LLY21129.cha": line 1835.  
 CHI: < Wai3 , bei2 faan1 lei5 # e3 wou1wou1 aa1 > [=! talking to her sister] .  
 PRT give PVT you PRT doggy PRT  
 'Give you back doggy.'
17. LLY, 2;11.29 \*\*\* File "21129.cha": line 4733.  
 CHI: < A3, bei2 ngo5 "Barbie" sin1 > [<] .  
 PRT give me Barbie first  
 'Give me Barbie first.'
18. LLY, 3;00.11 \*\*\* File "LLY30011.cha": line 4783.  
 CHI: Bei2 ze4ze1 # jam2gun2 .  
 Give sister straw  
 'Give sister straw.'
19. LLY, 3;03.15 \*\*\* File "LLY30315.cha": line 470.  
 CHI: Bei2 ngo5 # xxx nei1 di1 sin1 .  
 Give me DET CL first  
 'Give me these first.'
20. HHC, 2;10.13 \*\*\* File "21013.cha": line 3322.  
 CHI: Baau2 sei2, bei2 lei5 nei1 go3.  
 full die give you DET CL  
 '(I'm) too full, give you this one.'
21. WBH, 2;09.19 \*\*\* File "20919.cha": line 1342.  
 CHI: Bei2 faan1 ngo5 zi2 aa1.  
 Give PVT me paper PRT  
 'Give me back paper.'



## Appendix Two. The non-target use of the [bei2-R-T] form in Cantonese-English bilingual children

### a) From the Hong Kong Bilingual Child Language Corpus (Yip, Matthews and Huang, 2001)

#### i) Timmy

1. 2;07.14, line 1524

CHI: Bei2 keoi5 zyu1gwullik1 laa1 .  
Give 3sg chocolate PRT  
'Give it (the mouse) chocolate.'

2. 2;11.12: line 1027

CHI: Jam2 jyun4 tong1 bei2 lei5 sei2 aa1 .  
Drink PVT soup give you water PRT  
'After drinking the soup, give you water.'

3. 2;11.12: line 1261

CHI: Bei2 lei5 hou2 do1 +...  
Give you many  
'Give you many +... .'

4. 2;11.12: line 1301

CHI: Bei2 lei5 hou2 do1 +/.  
Give you many  
'Give me many +/.'

5. 3;00.09: line 1570

CHI: Keoi5 hou2 hou2jan4 gaa3 bei2 ngo5 tong2tong2 gaa3 .  
3sg very nice PRT give me candy PRT  
'It (Bunny) was very nice, (it) gave me candy.'

Situation: Timmy was telling the adult investigator what presents (candy, gun) he got from a hotel at Easter.

6. 3;00.09, line 1600

CHI: Tou3zai2 bei2 ngo5 # # tong2tong2 gaa3 .  
Rabbit give me candy PRT  
'Bunny gave me candy.'

7. 3;00.09. line 1682

CHI: Bei2 ngo5 zi1 < coeng1 > [//] se6 coeng1 aa3 .  
Give me CL gun shooting gun PRT  
'Give me the gun, the shooting gun.'

8. 3;00.09, line 1888

CHI: Ngo5 zung3ji1 go3 < sin1saang1 > [//] lou5si1 # le1 be6 +...  
I like CL teacher teacher PRT PRT  
  
seng4jat6 bei2 ngo5 tong2tong2 ji1 zek3 aa3 .  
always give me candy this type PRT  
  
'I like the teacher, (s/he) always gives me this type of candy.'

9. 4;06.06, from Sophie's corpus file Sc971127: line 1754

BRO<sup>1</sup>: Sophie , ngo5 bei2 lei5 li1 go3 .  
Sophie I give you DET CL  
'Sophie, I give you this one.'

10. 5; 03.06, from Sophie's corpus file Sc980827: line 2678

BRO: Bei2 ngo5 siu2siu2 .  
Give me little  
'Give me a little.'

11. 5; 08.00, from Sophie's corpus file Sc990121: line 1655

BRO: Ngo5 m4 bei2 lei5 tong2tong2 .  
I NEG give you candy  
'I don't give you candy.'

ii) Sophie

1. 2; 03.24, line 534

CHI: Bei2 ngo5 jat1 tiu4 aa1 .  
Give me one CL PRT  
'Give me one.'

---

<sup>1</sup> Timmy's utterances in Sophie's corpus were coded as BRO where BRO stands for brother.

2. 2;05.02, line 190

CHI: Je4sou1 bei2 ngo5 cin2 aa3 .  
Jesus give me money PRT  
'Jesus give me money.'

3. 2;05.02, line 350

CHI: Hai6 je4sou1 bei2 ngo5 cin2 .  
Vf Jesus give me money  
'It's Jesus give me money.'

4. 2;05.02, line 2021

CHI: Je4sou1 bei2 ngo5 cin2 aa3 .  
Jesus give me money PRT  
'Jesus give me money.'

5. 2;05.02, line 2027

CHI: Je4sou1 bei2 ngo5 cin2 aa3 .  
Jesus give me money PRT  
'Jesus give me money.'

6. 2;05.16, line 2211

CHI: Ngo5 m4 bei2 lei5 aam1gim1 gaa3 .  
I NEG give you ice-cream PRT  
'I don't give you ice-cream.'

7. 2;05.16, line 2219

CHI: Ngo5 m4 bei2 lei5 ice-cream aa3 .  
I NEG give you ice-cream PRT  
'I don't give you ice-cream.'

8. 2;05.16, line 2370

CHI: Lei5 bei2 ngo5 min6baau1 aa3 .  
You give me bread PRT  
'You give me bread.'

9. 2;05.30, line 1447

CHI: Bei2 ngo5 gaan1 uk1 .  
Give me CL house  
'Give me the house.'

10. 2;05.30, line 2833

CHI:        Lei5 bei2 ngo5 meow , bei2 ngo5 aa3 .  
              You give me    cat        give me    PRT  
              'You give me cat, give me.'

11. 2;08.08, line 2838

CHI:        Bei2 ngo5 cin2    aa1 .  
              Give me    money PRT  
              'Give me money.'

12. 2;08.08, line 2844

CHI:        Bei2 ngo5 cin2    aa1 .  
              Give me    money PRT  
              'Give me money.'

13. 2;08.08, line 2850

CHI:        Bei2 ngo5 cin2    aa1 .  
              Give me    money PRT  
              'Give me money.'

14. 2;10.24, line 1778

CHI:        Bei2 ngo5 tong2 aa3 .  
              Give me    candy PRT  
              'Give me candy.'

15. 2;10.24: line 1811

CHI:        Timmy < dou1 > [/] dou1 m4    bei2    lei5    tong2    aa3 .  
              Timmy    also        also    NEG give    you    candy    PRT  
              'Timmy also doesn't give you candy.'

16. 2;10.24, line 2075

CHI:        Aa2 , bei2 ngo5 tong2    aa3 .  
              PRT    give me    candy    PRT  
              'Give me candy.'

17. 2;10.24, line 2078

CHI:        Bei2 ngo5 tong2    aa1 .  
              Give me    candy    PRT  
              'Give me candy.'

18. 2;11.00, line 3389

CHI: Bei2 ngo5 ji1 go3 bui1 aal .  
Give me DET CL cup PRT  
'Give me this cup.'

iii) Kathryn

1. 3;03.16, line 339

CHI: Ngo5 dou1 zyu2 di1 je5 aa3 , <ngo5 , ngo5 > [/] ngo5 bei2 lei5 jelly laa1 .  
I also cook CL thing PRT I I I give you jelly PRT  
'I also cook something, I give you jelly.'

2. 3;03.16, line 2010

CHI: Bei2 [/] bei2 ngo5 go2 go3 zaat3 bin1 go2 go3  
Give give me DET CL tie braid DET CL  
'Give me that one, that one for tying a braid.'

3. 3;03.16, line 2010

CHI: Bei2 ngo5 go3 bin1 laa1 .  
Give me CL braid PRT  
'Give me the braid.'

4. 4;02.17, line 1199

CHI: Jan4dei6 maai5 je5 lei5 jiu3 bei2 jan4 ## cin2 go3wo3 .  
Others buy thing you need give people money PRT PRT  
'Others buy things, you have to give people money.'

ii) **From Cheung's diary data (p.c.)**

iv) Siu Bou

1. 2; 04.02

Situation: Mother just gave the child a new toy.

CHI: Maa1maa1 bei2 ngo5 je5waan2  
Mother give me something for play  
'mother give me something (for play)'

2. 2;04.18

CHI: Ngo5 bei2 baa1baa1 go2 go3  
I give daddy DET CL  
"I give Daddy that one."

3. 2;04.19

CHI: Ngo5 bei2 suk1suk1 go2 go3 syun4syun4  
I give uncle DET CL boat  
'I give Uncle that boat.'

4. 2;06.13

CHI: Bei2 go3 baa1baa1 popsickle.  
Give CL daddy popsickle  
'Give daddy popsickle.'

Situation: Child wanted mum to give dad a popsickle.

At 2;06.26, mother recorded in her diary that "[T-R] is always expressed as [R-T].", but she did not record any specific examples in her diary.

5. 3;11.23

CHI: Captain Hook, ngo5 bei2 lei5 ni1 go3 gim3  
Captain Hook I give you DET CL sword  
'Captain Hook, I give you this sword.'

Situation: The child is playing Peter Pan. He is handing mum his toy sword, and he wanted his mum to be Captain Hook.

6. 3;11.23

CHI: Bei2 ngo5 go3 gim3.  
Give me CL sword  
'Give me the sword.'

7. 3;11.23

CHI: Bei2 Mickey Mouse go3 gim3.  
Give Mickey Mouse CL sword  
'Give Mickey Mouse the sword.'

8. 4;00.20

CHI: Maa1maa1, ngo5 bei2 lei5 hung4zai2

Mother I give you bear

'Mother, I give you bear.'

Situation: the child is handing his mum his panda bear.

9. 4;00.20

CHI: Maa1maa1, ngo5 bei2 lei5 me1je5 aa3?

Mother I give you what PRT

'Mother, I give you what?'

## References

- Arnold, Jennifer, Thomas Wasow, Anthony Losongco and Ryan Ginstrom. (2000). Heaviness versus newness: The effects of complexity and information structure on constituent ordering, *Language*, 76, 28-55.
- Barlow, Michael and Suzanne Kemmer. (eds). (2000). *Usage-Based Models of Language*. CSLI Publications.
- Bowerman, Melissa. (1993). Typological perspectives on language acquisition: do crosslinguistic patterns predict development? In Eve Clark (ed.) *The Proceedings of the Twenty-Fifth Annual Child Language Research Forum*. Center for the Study of Language and Information, 25, 7-15.
- Bruyn, Adrienne, Pieter Muysken and Maaïke Verrips. (1999). Double-object constructions in the Creole languages: development and acquisition. In Michel DeGraff (ed.) *Language Creation and Language Change: Creolization, Diachrony, and Development*. The MIT Press.
- Bybee, Joan. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10, 425-455.
- Bybee, Joan and Paul Hopper. (eds). (2001). *Frequency and the Emergence of Linguistic Structure*. John Benjamins Publishing Company.
- Bybee, Joan. (2002). Sequentiality as the basis of constituent structure. In T. Givón and B. Malle. (eds). *The Evolution of Language out of Pre-Language*. Amsterdam: John Benjamins. 107-134.
- Campbell, Aimee. L. and Michael Tomasello. (2001). The acquisition of English dative constructions. *Applied Psycholinguistics*, 22, 253-267.
- Cho Sookeun, Miseon Lee, William O'Grady, Minsun Song, Takaaki Suzuki and Naoko Yoshinaga. (2002). Word order preferences for direct object and indirect objects in children learning Korean. *Journal of Child Language*, 29, 897-909.



- Chung, Rachel T.-T. and Peter Gordon. (1998). The acquisition of Chinese dative constructions. *Proceedings of the 22<sup>nd</sup> Annual Boston University Conference on Language Development*, Volume 1, 109-120.
- Croft, William. (1990). *Typology and Universals*. Cambridge; New York: Cambridge University Press.
- Croft, William. (2001). *Radical Construction Grammar. Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- Croft, William and D. Alan Cruse. (2003). *Cognitive Linguistics*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Demuth, Katherine, Malillo Machobane and Francina Moloji. (2000). Learning word-order constraints under conditions of object ellipsis. *Linguistics*, 38, 3, 545-568.
- Elman, Jeff. L. (1996). Learning and development in neural networks: The importance of starting small. *Cognition*, 48, 71-99.
- Essegbey, James. (2002). *Inherent Complement Verbs Revisited: Towards an Understanding of Argument Structure in Ewe*. Doctoral thesis. Max Planck Institute for Psycholinguistics.
- Fillmore, Charles J., Paul Kay and Mary Catherine O'Connor. (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language*, 64, 501-538.
- Gentner, Dedre and Arthur. B. Markman. (1995). Similarity is like analogy: structural alignment in comparison. In Cristina Cacciari (ed.) *Similarity in Language, Thought and Perception*. Brussels: BREPOLs.
- Goldberg, Adele E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. The University of Chicago Press.
- Goldberg, Adele E. (2002a). Surface generalizations: an alternative to alternations. *Cognitive Linguistics*. 13-3: 327-356.

- Goldberg, Adele E. (2002b). Construction Grammar. *Encyclopedia of Cognitive Science*. Macmillan Reference Limited Nature Publishing Group.
- Gropen, Jess, Steven Pinker, Michelle Hollander, Richard Goldberg, Ronald Wilson. (1989). The learnability and acquisition of the dative alternation in English. *Language*, Vol. 65, Number 1, 203- 255.
- Hawkins, John A. (1994). *A Performance Theory of Order and Consistency*. Cambridge: Cambridge University Press.
- Hawkins, John A. (2001). Why are categories adjacent? *Journal of Linguistics*, 37, 1-34.
- Huang, James C-T. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry*, 15, 531-574.
- Jurafsky, Daniel, Alan Bell, Michelle Gregory, and William D. Raymond. (2001). Probabilistic relations between words: evidence from reduction in lexical production. In Bybee, Joan and Paul Hopper. (eds.). *Frequency and The Emergence of Linguistic Structure*. Amsterdam: John Benjamins. 229-254.
- Kiekhoefer, Kai. (2002). *The Acquisition of the Ditransitive Construction*. Paper presented at the Ninth International Congress for the Study of Child Language and the Symposium on Research in Child Language Disorders (IASCL/SRCLD), Madison, Wisconsin, July, 2002.
- Keenan, Edward and Bernard Comrie. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*. 8: 63-99.
- Kozinsky, Isaac and Maria Polinsky. (1993). Causee and patient in the causative of transitive: coding conflict or doubling of grammatical relations. In Bernard Comrie and Maria Polinsky. (eds). *Causatives and Transitivity*. John Benjamins Publishing Company, Amsterdam/ Philadelphia. 177-240.
- Langacker, Ronald. (1987). *Foundations of Cognitive Grammar*. Volume 1. Stanford, CA: Stanford University Press.
- Langacker, Ronald. (1995). Raising and transparency. *Language*, 71, 1-62.

- Larson, Richard. K. (1988). On the double object construction. *Linguistic Inquiry*, 19: 335-391.
- Law, Sam-Po, Leung Man-Tak, Fung Suk-Yee. (2002). *The Hong Kong Cantonese Adult Language Corpus (HKCAC)*.
- Lee, Thomas Han-Tak, Colleen Wong and Samuel Leung. (1996). *The Hong Kong Cantonese Child Language Corpus (CANCORP)*.
- Linguistic Society of Hong Kong. (2002). *Guide to LSHK Cantonese Romanization of Chinese Characters*. 2nd edition. Linguistic Society of Hong Kong: Hong Kong.
- Liu, Danqing (2001). Hanyu 'gei' zi lei shuang jiwu jiegou de leixingxue kaocha. *Zhongguo Yuwen*, 284, 387-97.
- MacWhinney, Brian. (2000). *The CHILDES Project: Tools for Analyzing Talk*. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.
- Matthews, Stephen and Virginia Yip. (1994). *Cantonese: A Comprehensive Grammar*. London: Routledge.
- Matthews, Stephen and Virginia Yip. (2003). Relative clauses in early bilingual development: transfer and universals," in A. Giacalone Ramat. (ed.), *Typology and Second Language Acquisition*. Mouton de Gruyter.
- Matthews, Stephen and Leung Tsz-Cheung. (2002). *On Dative Constructions in Cantonese and Thai*. Paper presented at the 35th International Conference on Sino-Tibetan Languages and Linguistics, Arizona State University, November, 7-10, 2002.
- Mazurkewich, Irene. (1984). The acquisition of the dative alteration by second language learners and linguistic theory. *Language Learning*, 34, 91-109.
- Michaelis Susanne and Martin Haspelmath. (2003). Ditransitive constructions: Creole languages in a cross-linguistic perspective. *Creolica*. Downloadable at <http://www.creolica.net/articles.php3>.

- Newman, John. (1999). Figurative Giving. In Leon de Stadler and Christoph Eyrich. (eds). *Issues in Cognitive Linguistics: 1993 Proceedings of the International Cognitive Linguistics Conference*. 113-140.
- O'Grady, William. (1997). *Syntactic Development*. Chicago: University of Chicago Press.
- O'Grady, William. (2000). *A Linguistic Approach to the Study of Language Acquisition*. Invited keynote talk given to the Annual Meeting of the Pan-Pacific Association of Applied Linguistics, Honolulu, Hawaii, July 2000.
- Pinker, Steven. (1989). *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, Mass. London: The MIT Press.
- Polinsky, Maria. (1998). A non-syntactic account of some asymmetries in the double object construction. In Jean-Pierre Koenig. (ed). *Conceptual Structure and Language: Bridging the Gap*. Stanford: CSLI, 403-423.
- Ruppenhofer, Josef. (m.s.). *Null Complementation*. Unpublished manuscript.
- Saffran, Jenny R., Richard N. Aslin, and Elissa L. Newport. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926-1928.
- Schwenter, Scott and Gláucia Silva. (2002). Overt versus null direct objects in spoken Brazilian Portuguese: a semantic/pragmatic account." *Hispania*, 85.3. 901-909.
- Slobin, Dan. I. (ed.). (1985). *The Crosslinguistic Study of Language Acquisition: Vol. 1. The Data. Vol.2. Theoretical issues*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Tai, James (1985). Temporal sequence and word order. In John Haiman (ed.), *Iconicity in Syntax*, Amsterdam: John Benjamins, pp.49-72.
- Tang, Sze-Wing. (1998). On the 'inverted' double object construction. In Matthews, Stephen. (ed.). *Studies in Cantonese Linguistics*. Linguistic Society of Hong Kong.

- Tomasello, Michael. (1998). The return of constructions. *Journal of Child Language*, Volume 25, Issue 02. 431-442.
- Tomasello, Michael. (2000a). Do young children have adult syntactic competence? *Cognition*, 74, 209-253.
- Tomasello, Michael. (2000b). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*. 11, 1/2. p. 61-82.
- Tomasello, Michael. (2003). *Constructing a Language: A Usage-Based Theory of Child Language Acquisition*. Harvard University Press.
- Wasow, Thomas. (1997). Remarks on grammatical weight. *Language Variation and Change*, 9, 81-105.
- Wasow, Thomas and Jennifer Arnold. (2003). Post-verbal constituent ordering in English. In Günter Rohdenburg and Britta Mohndorf (eds.) *Determinants of Grammatical Variation in English*, 119-154. Mouton.
- Wolfe-Quintero, Kathryn E. (1994). *The Representation and Acquisition of the Lexical Structure of English Dative Verbs: Experimental Studies of Native English Speakers and Japanese and Chinese Adult Learners of English*. Doctoral Thesis. University of Hawaii. Ann Arbor, Mich. UMI.
- Xu, Liejiong and Alan Peyraube. (1997). Double object and oblique constructions in Cantonese. *Studies in Language*, 21:1, pp. 105-127.
- Yip, Virginia and Stephen Matthews. (2000). Syntactic transfer in a Cantonese-English bilingual child. *Bilingualism: Language and Cognition*, 3.3.: 193-208.
- Yip, Virginia and Stephen Matthews. (2001). *Intermediate Cantonese: a Grammar and Workbook*. London: Routledge.
- Yip, Virginia and Stephen Matthews and Huang Yue-Yuan. (2001). *The Hong Kong Bilingual Child Language Corpus*.



CUHK Libraries



004144367