湖南大学青年社科学者文库

曾涛 著

词汇飞跃现象研究

——两名长沙儿童的早期词汇与语义发展

An Investigation of Word Spurt Phenomenon

——Early Lexical and Semantic Development of Two Changsha Children

湖南大学出版社

湖南大学青年学者学术著作出版资助项目

词汇飞跃现象研究

——两名长沙儿童的早期词汇与语义发展

An Investigation of Word Spurt Phenomenon

——Early Lexical and Semantic Development of Two Changsha Children

曾 涛 著

湖南大学出版社

内 容 简 介

通过描绘两名儿童在词汇飞跃时期表现出的系列语言发展特征，探究儿童早期词汇和语义的发展规律，同时，探索词汇飞跃背后的习得机制，对命名洞察力的现实性进行了深入考察，以期在一定程度上揭示词汇飞跃现象的本质。

图书在版编目(CIP)数据

词汇飞跃现象研究——两名长沙儿童的早期词汇与语义发展/曾涛著.
—长沙：湖南大学出版社，2012.2
ISBN 978-7-5667-0138-1
Ⅰ.①词… Ⅱ.①曾… Ⅲ.①儿童语言—研究—英文
Ⅳ.①H003
中国版本图书馆 CIP 数据核字(2012)第 022348 号

词汇飞跃现象研究——两名长沙儿童的早期词汇与语义发展
Cihui Feiyue Xianxiang Yanjiu
——Liangming Changsha ertong de Zaoqi Cihui yu Yuyi Fazhan

作　　者：曾　涛　**著**
责任编辑：王桂贞
特约编辑：卢付林　王湘平
出版发行：湖南大学出版社　　**责任印制**：陈　燕
社　　址：湖南·长沙·岳麓山　　**邮　　编**：410082
电　　话：0731-88822559(发行部)，88821343(编辑室)，88821006(出版部)
传　　真：0731-88649312(发行部)，88822264(总编室)
电子邮箱：wanguia@126.com
网　　址：http://www.hnupress.com
印　　装：长沙瑞和印务有限公司
开本：710×1000　16 开　　**印张**：17　　**字数**：306 千
版次：2012 年 3 月第 1 版　　**印次**：2012 年 3 月第 1 次印刷
书号：ISBN 978-7-5667-0138-1/H·169
定价：42.00 元

致　谢

本书的出版获得了国家社科基金一般项目“汉语儿童早期范畴分类能力的发展研究”(11BYY037)、教育部人文社会科学研究青年基金项目“词汇飞跃的本质”(10YJC740008)、湖南省哲学社会科学基金项目一般项目“儿童基本层次词汇的习得”(09YBB088) 以及湖南大学中央高校基本科研业务费专项资金项目“词汇飞跃现象研究——两名长沙儿童的早期词汇与语义发展”、“词汇飞跃的本质”以及“汉语词类的基本层次效应”的资助，在此深表谢意！

2001 年，我有幸参与了湖南大学认知科学研究所李行德教授和宁春岩教授领导下的“中国早期儿童语言发展(CELA)”研究项目，由此踏入了儿童语言习得研究领域。2007 年，我考入广东外语外贸大学语言学及应用语言学研究中心，师从李行德教授和王初明教授，攻读博士学位，由此开始了对汉语儿童早期词汇与语义发展的系统研究。2010 年，我顺利地完成了博士论文的写作和答辩，本书就是在博士论文的基础上整理完善和发展的。

在广东外语外贸大学、香港中文大学和湖南大学的求学经历，不论是在学术上还是在个人情感上，都是我一生极其宝贵的财富。在这些年里，很多人对我学术研究的进步和心智的成熟给予了无私的帮助和扶植。

首先，我要衷心感谢我的导师李行德教授。他渊博的知识、深度的评论以及建设性的意见，对本研究的顺利完成起了至关重要的作用。李行德教授对待学生热忱耐心，对待研究与教学全身心地投入，是我效仿的楷模。我深深地感谢李行德教授对整个研究过程的全程指导，感谢他对我严格的学术训练，感谢他创造许多机会让我得以展现自己的研究成果，感谢他培养我语言学之外的独立的人文思想。不论是在学术研究、教书育人，还是为人处世上，李行德教授永远都是我学习的榜样。

其次，我要特别感谢另外一位联合培养我的导师王初明教授，感谢他给予我宝贵的机会来继续我的研究，也感谢他帮助我重拾信心去面对困难。我很感谢王初明教授对这项研究感兴趣并对其进行深度的点评，也感谢他在这整个研究过程中所给予的指导和支持。我想很多年后我都不会忘记王初明教授主持的读

书报告，通过回答他提出的引人深思但有时令人尴尬的问题，我受益匪浅。他的幽默、睿智及独特的视角使我的博士求学生涯变得难忘、快乐与丰富。

再次，我要感谢李行德教授和黄爱军，感谢他们成为我合作和探讨的伙伴。我从我们的讨论中收获颇多，特别要感谢黄爱军慷慨地与我分享她收集的 AJR 的原始词汇数据以及相关的数据的分析，否则这项研究就只能是 LSY 的个案研究了。

同时，我感谢本研究的开题委员会的教授，Cathy Wong，Samuel Leung 和 Peppina Lee，谢谢他们给我的建设性的反馈意见，让我收获良多。我还要深深感谢预答辩和答辩委员会的老师：吴旭东教授、郑超教授和何晓玮教授，他们给予的宝贵点评和启发性建议拓宽了我的思维，使得论文的数据分析更具严谨性。我特别要感谢吴旭东教授，感谢他对论文各个方面深刻而又详细的点评。我还要感谢所有为此论文的修改提出过建设性意见的匿名审稿人。

此外，我非常感激所有教导过我的老师。宁春岩教授不仅让我欣赏到语言的严谨美，也体会到语言世界的奥妙之处。我也要感谢李兵教授、徐烈炯教授、伍雅清教授、胡建华教授和张文忠教授，感谢他们所讲授的课程以及给予的指导。我还要真心地感谢广东外语外贸大学的老师，特别是桂诗春教授、钱冠连教授、吴旭东教授、温宾利教授、刘建达教授、冉永平教授、霍永寿教授、郑超教授、何晓玮教授和董燕萍教授。在充满压力而又辛酸的求学岁月里，他们那温暖人心的笑容，亲切的鼓励以及大力支持给我留下了很多欢乐与难忘的回忆。

另外，我感谢湖南（CELA）实验室的成员：陈敏、陈艳华和陈菲艳，他们在数据的收集以及处理上付出了很多的努力。还要感谢我的研究伙伴：艾朝阳、蔡欣、李茹亚、廖慧、肖玲、罗琼鹏、杨洁、胡亚娟、李洁琼、庄宇峰和汪浩。我也要感谢北京（CELA）实验室的成员：宋刚、范丽、杨小璐、李晓茜、张庆文等等以及香港（CELA）实验室的成员：Margaret Lee、陈卉、黄爱军、Kitty Szeto 等等，感谢他们的宝贵建议和大力支持。我为我能够成为如此庞大的学术家庭的一员深感骄傲。

本文的部分内容曾在以下研讨会中做过报告：第三届形式语言学国际会议以及第二届岳麓语言习得研讨会（长沙，2005），第十届儿童语言研究国际会议（柏林，2005），第十七届国际中国语言学学会年会（巴黎，2009）以及汉语习得里程碑研讨会（香港，2009）。在此我想感谢那些与会者所给予的建议与反馈。我要感谢 Lise Menn 教授与我分享她对幼儿话语分析的专业见解，也要感谢 Jennifer Ganger 教授对确定词汇飞跃的统计方法的专业建议，还要特别感谢 Shih Chilin 教授与我分享她做长期跟踪研究的经验以及对改善本研究的建设性意见。

我觉得我是如此幸运地拥有朋友给我的爱、智慧和欢乐，这是无法用言语表达的。我要特别感谢 Rain，Zhang Jinghan，汪朋，贺川生和贺文丽给予我极大的帮助和支持，也要感谢广东外语外贸大学的朋友和校友：袁邦照、何嫣、魏行、王敏、张珊珊、杨梅、唐安华、高唯、赖小玉、陈金诗、龚双萍、胡文飞、夏立新、徐福平、王家明等等。尽管他们自己学业紧张，仍然在我需要的时候给予我学术上和精神上的支持。我将永远不会忘记我们曾经在美丽的校园散步，每周定期攀登白云山的时光。我想若干年过后，我仍然会记得我们的乒乓球活动室、北门外的糖水，钱柜的卡拉 OK 以及在广州经历过的激动人心的点点滴滴。

最后，我想要感谢我深爱的父母和丈夫，感谢他们分享我的汗水和泪花；感谢他们对我选择这条漫长的学术道路的理解；感谢他们在我到广州和香港求学时对家里的照料。我还要感谢我亲爱的儿子斯毛，他总是给我带来意想不到的灵感。谨以此书献给我的家人。

作 者

2011 年 12 月

List of Abbreviations

CB	Context Bound
CDI	Communicative Development Inventory
CF	Context Free
CT	Contrast Theory
D (S)	Determinacy (Scope)
E (F)	Expressive (Function)
FCT	Functional Core Theory
FLA	First Language Acquisition
FMT	Fast Mapping Theory
IMIT	Imitation
I (S)	Indeterminacy (Scope)
IT	Information Theory
L1	First Language
LS	Language Score
MLU	Mean Length of Utterance
N3C	Novel Name-Nameless Category
NP	Noun Phrase
NV	Non-verbal
P (F)	Performative (Function)
R (F)	Referential (Function)
SFP	Sentence Final Particle
SFT	Semantic Feature Theory
SPON	Spontaneous
SRN	Simple Recurrent Network
UG	Universal Grammar
V	Verbal
Y/M/D	Year / Month / Day

Contents

Chapter 1 Introduction: Word Spurt in Early Lexical and Semantic Development

Children's abilities to learn new words are full of miracles, and the study of these abilities bears on the most central questions in cognitive science. What kind of vocabulary knowledge does the child have at very early stages of language acquisition? What is the nature of human learning? These questions do not have simple answers, but involve different cognitive capacities working together in an intricate fashion (P. Bloom, 2000). The present study investigates children's early vocabulary learning systematically, both for its own sake and for its potential contributions to such fields as language universality and language assessment. It addresses a number of theoretical and methodological issues in early lexical and semantic development, specifically, the word spurt, naming insight and semantic characteristics of early lexicon, based on a longitudinal study of two Mandarin-speaking children. This chapter introduces the word spurt phenomenon and its cognitive significance, the scope of the present research, the theoretical and methodological considerations, key research questions and the outline of the remaining chapters.

1.1 Word Spurt and Its Cognitive Significance

An intriguing phenomenon that many researchers have reported in literature is word spurt①. It is characterized by a sudden rise of vocabulary growth toward the end of the one-word stage, typically around one and a half years old, when the child has a vocabulary of around 50 words. The existence

① "Word spurt", sometimes is called "lexical spurt" or "lexical explosion" by different researchers.

of a vocabulary spurt in children's early lexical development has been claimed by many researchers on the basis of parental reports (Goldfield and Reznik, 1990; Fenson, Dale, Reznick, Bates, Thal, and Pethick, 1994; Caselli, Casadio, and Bates, 2001) as well as longitudinal case studies based on varying number of observation points (Bloom, 1973; Nelson, 1973; McShane, 1980; Dromi, 1987; Bates, Bretherton, and Snyder, 1988; L. Bloom, 1993).

This lexical spurt carries special cognitive significance. Chomsky (2002) claims that lexical explosion (namely, the ability to handle a large number of lexical items) is a cognitive capacity that is part of the biological endowment of species and is thought to be unique to humans. Other researchers propose that word spurt should be linked to the appearance of the naming insight essential for grasping denotation (Dore, 1976; Kamhi, 1986; McShane, 1980) and the entry into the two-word stage (Goldfield and Reznick, 1990; Fenson et al., 1994; Dromi, 1987/1996). According to Goldfield and Reznick (1996), vocabulary spurt marks a qualitative change in the nature of language. With some critical mass of words in the lexicon, the child can generalize the word-referent relationship to include all manner of entities, qualities, and events, linking words and word meanings into a semantic system.

While word spurt has been documented in a number of longitudinal and cross-sectional studies, and its cognitive significance has also been examined, the reality of word spurt and its links to cognitive and early language development has been critically reassessed in recent years. On methodological grounds, scholars have disputed whether one can establish a lexical spurt quantitatively in terms of the rate of acquisition of new words (Bloom, 2000; Ganger and Brent, 2004). On cognitive grounds, researchers have challenged the reality of the naming insight, arguing for a gradual transition from context-bound prelexical forms to fully denotational words (Carter, 1979; Bates, Camaioni, and Volterra, 1979; Nelson and Lucariello, 1985; Vihman and McCune, 1994). Connecting the word spurt with the onset of syntax can also be questioned if one sees syntax as available to the child in the one-word stage (McNeill, 1970) or that predicate-argument structures are already present in gesture-word combinations (Özçalişkan and Goldin-Meadow, 2005).

In Mandarin Chinese, few studies have been conducted on children's early lexical acquisition, except for one indirect research carried out by Tardif etc.

(2008). They reported on the construction and norming of parent report instruments for typically-developing Mandarin- and Cantonese-speaking children between eight and 30 months of age, based on MacArthur Communicative Development Inventories. In addition, their manual provides the information that users need to apply and interpret these procedures. However, the work by Tardif etc. (2008) does not give enough information to explore the nature of lexical spurt, although word spurt is counted as an important stage in children's language development by many studies.

The present research intends to describe a series of characteristics of early lexical and semantic development around the time of word spurt in an effort to explore the mechanism underlying word spurt, especially the reality of the naming insight essential for grasping denotation. Additionally, this study sets out to clarify some methodological issues concerning the study of early vocabulary development, in particular, those relevant to the study of lexical acquisition of Mandarin-speaking children.

1.2 Scope of the Present Study

How can a person interpret word spurt phenomenon? What is the mechanism behind the rapid increase of children's early vocabulary? To answer these questions, the present study investigates the nature of word spurt in both early lexical and early semantic development.

With regard to word spurt in early lexical development, the present study centralizes on its link to the onset of the two-word stage. Word spurt is considered to be related to early syntactic development or to be a precursor to syntactic development. Rapid increase in naming is quickly followed by the emergence of two-word utterances, and the development of names for objects usually precedes the development of the structured speech (Leopold, 1939; Halliday, 1975; Nelson, 1973; McShane, 1980; Bates et al., 1988). Many studies have confirmed a potential link between word spurt in the second year and the subsequent appearance of multi-word speech. Word spurt is deemed to be a milestone in children's early lexical development, contributing to their later grammatical development.

With respect to word spurt in early semantic development, the present research focuses on its relation to children's categorization capabilities. In early lexical development, children's categorization capabilities help them acquire and expand their vocabulary. In this sense, a study of these abilities can provide a good way for understanding the relationship between lexical acquisition and conceptual development. Categorization abilities can decrease the complexity of human perception, and category systems can provide maximum information with the least cognitive efforts (Rosch, 1978). It is through categorization that concepts are gradually formed and children learn to group things for the purpose of naming (Jackendoff, 1983). Through categorization, children know not only the individual objects, but also the world they live in. Two phenomena are observed around the spurt time, which represent children's categorization capabilities in early semantic development:

> The word *o* 鹅 "goose" was first used by the child LSY at 1;4;26 (Y/M/D) while he was answering adults' identification questions about a goose toy. At 1;6, LSY extended the word to the same referent "duck" twice when responding to an adult's questions about a duck drawing.
>
> The word *bi* 笔 "pen" was first uttered by the child AJR at 1;6;24 (Y/M/D), when she looked at a pen or pencil or expressed a desire for a pen or pencil. At 1;8;7, she also called a mini-recorder she saw as *bi* 笔. At 1;10;23, this word was used when AJR picked up a chalk in reply to adults' request in terms of addressing. ①

In the above examples, the first phenomenon is the use of overextended words. Both LSY and AJR used a word for a broader range of referents than was conventional in adult usage. For example, LSY used *o* "goose" to refer to "duck", and AJR pronounced *bi* "pen" for a mini-recorder. The second is the selection of basic-level terms. Both children selected a basic-level word (Rosch

① These two examples are drawn from the paper presented by Zeng and Lee (2009) at the 17th Annual Meeting of International Association of Chinese Linguistics (IACL-17), Paris, France, July 2-4.

et al. ,1976;Lakoff,1987) that represented the basis of their conceptualization. These basic-level words maximized the number of attributes shared by members of one category and minimized the number of attributes shared with members of other categories (Lakoff, 1987). For instance, AJR uttered *bi* 笔 instead of *gangbi* 钢笔 "pen", LSY produced *o* 鹅 "goose" instead of *dongwu* 动物 "animal".

Whether early words are overextended to a wider referential scope than conventional and whether early vocabulary is composed of basic level terms[①] are two issues important for evaluating children's categorization capabilities and for understanding the link between lexical and semantic acquisition. A lot of research has provided evidence for the overextension of children's early words (Preyer, 1889; Lewis, 1951; Leopold, 1939; Rescorla, 1980; Bowerman, 1978; Barrett, 1986; Clark, 1983, 1993; Dromi, 1996). Scholars argue that overextension tends to occur in high-frequency and early acquired words, covering such domains as vehicle, clothing, numbers and letters (Clark, 1993; Rescorla, 1980; Benedict, 1979). Overextension is an important issue in the study of early lexical and semantic development. The study of overextension can offer evidence to reflect children's early formation of rules, which bears special theoretical significance. Moreover, it can provide crucial evidence for deciding on one of the major theories in semantic development: the Semantic Feature Theory (Clark, 1973), which hypothesizes that children begin with a subset of the semantic features associated with a word, predicting overextension. Additionally, the present study furnishes systematic data on children's use of overextended words in Mandarin Chinese, which has hitherto been unavailable.

The present study examines the extent to which early Mandarin vocabulary is composed of basic-level terms. Basic-level words occupy the majority of children's early words and reflect categories of high psycholinguistic accessibility (Rosch, 1976), although whether the predominance of basic-level words is affected by input is still queried (Mervis and Mervis, 1982; Clark, 1993). The study of basic-level

① "Basic-level terms" and "basic-level words" are used interchangeably in this paper.

words is crucial for understanding the link between word acquisition and conceptual development. ① So far, no longitudinal research has been conducted on Mandarin-speaking children's early basic-level words. Only one study (Jiang, 2000) has reported data on basic-level vocabulary of Mandarin-speaking children. However, in Jiang's study, how basic-level vocabulary evolved before two years of age was not studied. Furthermore, adult criteria were adopted to determine the basic-level vocabulary of children, based on productivity and high frequency of corpus data.

Previous research reveals that few studies have charted the characteristics of early lexical and semantic development around the time of word spurt, and that some methodological considerations have cast doubt on the claims of early vocabulary studies. Furthermore, few studies have explored the development of children's naming capabilities or the reality of naming insight.

The present research probes into children's early language development by investigating the nature of word spurt at a close range through longitudinal audio-visual recordings of two Changsha infants from around nine months old to around two years old. Five aspects of children's cognitive development are surveyed: firstly, the rate of lexical development, especially the reality of word spurt; secondly, the relative chronology of word spurt and the onset of two-word combinations; thirdly, the proportion and characteristics of overextended words in early language production; fourthly, the distribution of basic-level words as well as the development of noun hierarchies; finally, the mechanism underlying the use of early words, especially the reality of naming insight.

1.3 Theoretical and Methodological Considerations

One of the major goals of generative grammar is to account for the logical problem of language acquisition, namely, how humans can acquire a complex internalized language (I-language) in such a short time on the basis of impoverished input. Rich innate domain-specific knowledge in the form of

① The present study only considers those conceptual structures that can be represented by language, not those that are independent of language.

universal grammar must be postulated to explain the basic facts of language learning (Chomsky, 1965; Wexlerand Culicover, 1980). Chomsky (2002) expresses his idea of studying language as a natural object, an ability that is part of the biological gift of our species, physically represented in the human brain. In his critique of Mark Hauser's view on evolution of communication, he claims that the capability to handle a large number of words is sometimes called lexical explosion and is considered to be unique to human beings.

Word learning is an intellectual feat accomplished by two-year-olds. It is obvious that word-world pairing is insufficient to account for word learning, for children know too much, from too little information, about the world. For instance, in verb learning, not every verb uttered by adults goes with its corresponding action. "Conversation, even from mothers to babies is not a running commentary on the objects, events, properties, and relations presently on exhibit in the world. No mother carefully utters 'open' every time she opens the door; worse, 'open' is frequently uttered—even systematically so—when the door is shut." (Gleitman and Fisher, 2005: 125)

This study probes into the nature of word spurt in early lexical and semantic development in light of the logical problem of language acquisition, with the following theoretical and methodological considerations.

1.3.1 Theoretical Considerations

The theoretical considerations of the present study are three-fold. Firstly, the present study tackles the issue of naming insight in the first language (L1) acquisition of early words to see if the development is sudden or gradual. The study of word spurt phenomenon is related to two perspectives on the transition from word-like forms to true words. One view holds that the mastery of naming is a continuous process, whereas the other view sees the developing of such ability as the grasping of a sudden insight. The former has challenged the reality of naming insight, holding that children have a gradual increase in the rate of word learning (Bates and Goodman, 1997; P. Bloom, 2000) and the grasp of denotation involves a slow and gradual transition from pre-lexical forms to fully referential words (Bates, Camaioni and Volterra, 1979; Nelson and Lucariello, 1985; Vihman and McCune, 1994; Tomasello,