

成人言语输入对儿童早期单词获得影响的个案追踪

陈 杰^{1,2} Setoh Peipei^{3,4} 孟祥芝¹ Tardif Twila²

(¹ 北京大学心理系, 北京 100871) (² Department of Psychology, University of Michigan, 美国) (³ Department of Psychology, Nanyang Technological University, 新加坡) (⁴ Department of Psychology, University of Illinois at Urbana-Champaign, 美国)

摘 要 追踪观察一名婴儿(6~20个月), 分析其中看护者的言语输入特征及婴儿早期词汇获得的发展变化。主要对成人言语输入中动、名词比例、单词在句中的位置、具体环境等因素及婴儿早期动、名词理解和产生等方面进行探讨。结果显示, 成人言语输入中动词比例显著高于名词, 更多动词位于句首或句尾使得主语和宾语省略; 同时, 这种动词优势的输入特征促进儿童早期动词获得, 使得婴儿早期语言样本中动、名词理解相对比例与成人言语输入一致。这一结果表明, 婴儿词汇发展初期就已经利用言语或社会线索, 同时这种早期词汇组成中较高的动词比例进一步支持“名词优势”理论存在跨语言差异。

关键词 成人言语输入; 动词优势; CLAN

分类号 B844

1 前言

儿童早期词汇发展是非线性的过程。在语言产生初期(12个月左右), 单词获得非常缓慢; 但是到19个月, 词汇飞速增长, 进入“单词迸发阶段”(word spurt), (Hollich, Hirsh-Pasek, & Golinkoff, 2000; Tardif, Casasola, Choi, & Meng, 2006)。目前关于儿童早期词汇学习存在很多理论, 但这些理论往往只强调词汇发展的某一单一机制, 没有考虑多种因素的共同影响。例如, 原则限制理论(constraints/principles theories)强调词汇获得的先天认知策略, 认为儿童天生具有一些偏向, 帮助缩减单词可能意思的范围, 从而促进单词获得(如: *whole object assumption*, *mutual exclusivity*, *taxonomic assumption* 等)(陈杰, Tardif, 孟祥芝, 2007; Golinkoff, Mervis, & Hirsh-Pasek, 1994; Tardif, Casasola, Choi, & Meng, 2006)。然而这种认知偏向很难解释不同文化或个体间的语言发展差异。与此相反, 社会语用论(social-pragmatic theory)强调社会、语言环境对词汇发展的作用, 提出儿童并不是通过一些偏向去猜测单词的意思, 而是成人通过猜测儿童的注意焦点从而提供适当的单词(Nelson, 1988)。

1988)。虽然这一理论可以解释词汇获得的差异性, 但对不同语言中存在的类似倾向说明不足 (Bornstein, Cote, Maital et al., 2004)。另一种理论是联结主义的观点(associationistic view), 认为单词获得是一种直接的匹配。儿童通过发现单词与参照物同时出现的频率来积累词汇。这一理论虽然没有直接强调成人的言语输入和环境的作用, 但这种词-词(word-to-world)匹配是需要通过观察和丰富的语言环境来实现的。以上三种理论虽然可以从不同角度分析儿童如何获得单词, 但都无法解释不同阶段中词汇发展的差异。最近提出的生成联合模型(Emergentist Coalition Model, ECM)(Hollich, Hirsh-Pasek, & Golinkoff, 2000) 融合了多方面因素来解释词汇获得和迸发现象。这一理论认为词汇获得是这些因素共同作用的生成产物, 其中包括认知限制、社会语用以及一些注意机制。在词汇获得的不同阶段, 每个因素的权重是不同的, 而儿童倾向使用的原则是从不成熟(attentional cues)到成熟(social cues 和 linguistic cues)的过程。虽然一些实证研究已经发现儿童早期词汇发展过程中线索使用的不平衡性(Fernandes, Marcus, Nubila, & Vouloumanos, 2005; Hirsh-Pasek & Golinkoff, 1996; Hollich,

Hirsh-Pasek, & Golinkoff, 2000; Naigles, 1990; Pruden, Hirsh-Pasek, & Golinkoff, 2006), 但是目前 Tardif 等研究发现, 在语言发展的最早阶段(0~10 个单词)仍然存在跨语言差异(Tardif, 2006), 这些差异与言语输入特征相关而并不是知觉线索可以解释的, 因此对 ECM 和认知偏向理论提出挑战。

Tardif 等人分别对 8~16 个月的英语、普通话和广东话婴儿进行 MacArthur Communicative Development Inventory(MCDI)测量, 结果发现早在婴儿 0~10 词阶段已经存在词汇获得的跨语言差异: 名词优势并不是普遍存在的, 汉语儿童早期词汇中动词比例要显著高于英语儿童。可见, 以上各种理论虽然可以从不同角度解释儿童早期的单词学习, 但对于词汇获得的跨语言差异现象却没有一致的看法, 而其中动、名词获得优势仍是跨语言研究的核心内容。

名词优势一直是儿童早期词汇获得的争论焦点(Au, Dapretto, & Song, 1994; Bates, Marchman, Thal et al., 1994; Choi & Gopnik, 1995; Gentner, 1982; Kim, McGregor, & Thompson, 2000; Tardif, 1996; Tardif, 2006; Tardif, Gelman, & Xu, 1999)。例如, Gentner 最早提出一个自然分割(natural partitions)假设, 认为儿童先天就可以在知觉和概念上区分具体概念(如人物、客体)与谓语概念(如行为、状态改变或因果关系)并且具有名词获得优势(Gentner, 1982)。另外一些研究者却认为早期词汇获得与先天的认知偏向关系较小, 更多依赖语言的结构特征和输入线索, 因此成人言语输入特征会很大影响儿童早期词汇获得(Choi & Gopnik, 1995; Tardif, 1996; Tardif, 2006; Tardif, Gelman, & Xu, 1999)。虽然很多研究发现大部分语言中名词获得要早于动词, 对于汉语和韩语的研究仍然发现语言特异性(陈杰, Tardif T., 孟祥芝, 2007)。为什么儿童词汇发展的最早阶段就出现跨语言差异, 是否他们很早就可以利用语言线索而不仅仅依赖知觉和注意呢?

Maguire 等人最近利用 ECM 和 SICI(Shape, Individuation, Concreteness, and Imageability)连续体来解释儿童早期词汇中出现动词的现象(Maguire, Hirsh-Pasek, & Golinkoff, 2006)。按照这一理论, 词汇获得差异并不仅仅存在于词类间(动词或名词), 同一词类内部也有不同。决定某一单词获得是否困难并不取决于单词所属的词类范畴, 而与对应参照物的具体性、突显性等因素有关。因此, 大多数语

言中动词获得之所以困难在于它们处于 SICI 连续轴中更抽象的位置。对于另外一些语言(如汉语), 早期出现的动词往往更为具体, 在知觉上更容易获得。除了知觉线索, 词类获得的差异也很大程度受到语言结构特征和输入方式的影响(Bornstein, Cote, Maital et al., 2004; Tardif, Shatz, & Naigles, 1997), 如汉语动词往往省略主语或宾语, 使得频率和位置具有优势。

一些研究已经对儿童早期动词获得的知觉偏向进行了探讨, 发现这些动词相对后来出现的动词在知觉上更容易获得(如“吃”、“跑”要早于“想”、“愿意”等), 出现的环境更为具体(Snedeker & Gleitman, 2004; Tardif, 2006)。此外, 针对可能利用的语言线索, 部分研究者也很早就关注言语输入与儿童早期词汇发展的关系。例如, Goldfield 曾对 12 对一岁婴儿与母亲的交流活动进行观察, 发现母亲在游戏场景中说出的名词比例高于动词, 而这一差异与孩子说出的单词比例有显著正相关(Goldfield, 1993)。Tardif 对英语、意大利语和汉语儿童与母亲自然交流的样本进行分析, 发现母亲言语输入中的动、名词比例与儿童自发说出的动、名词比例非常一致, 而且不同语言存在很大的差异: 英语母亲更强调名词, 且名词更多出现在句尾, 而汉语母亲动词的频率更高, 出现在句尾的比例也更大(Tardif, Shatz, & Naigles, 1997)。与此类似, Choi 和 Gopik 及 Kim 的研究也发现英语和韩语母亲的言语输入与儿童早期词汇获得的一致性(Choi & Gopnik, 1995; Kim, McGregor, & Thompson, 2000)。

虽然这些研究通过自然观察的方法客观地记录成人言语输入与儿童早期的动、名词比例, 但都无法考察儿童词汇获得的发展趋势。同时, 一次场景中的观察结果存在偶然性, 而长期追踪更能发现其中的规律和稳定性。即使 Choi 和 Gopik 的研究对婴儿进行了纵向追踪(14~20 个月), 但也仅仅涉及儿童本身的词汇发展历程, 没有从纵向发展的视角分析成人言语输入对儿童词汇获得的影响。另外, 韩语特征与汉语完全不同, 即使二者都存在动词优势。韩语的句法结构是主语—宾语—谓语(SOV)形式, 动词更多出现在句末, 因此更加突显(大量研究发现位于句末的单词更容易被孩子注意和学习)(Bornstein, Cote, Maital et al., 2004; Casell et al., 1995; Goldfield, 1993; Snedeker & Gleitman, 2004; Tardif, Shatz, & Naigles, 1997)。但是汉语本身的句法结构是 SVO 形式, 因为输入时常省略主语和宾

语才使得动词的位置更为突出。因此对于汉语成人言语中动词出现的位置进行分析更能看出输入信息对于儿童早期词汇发展的影响。本研究拟采用对个案进行追踪的自然观察法,考察以下问题:

(1) 检验汉语成人言语输入中动词和名词比例是否存在差异,并分析其原因。

(2) 成人语料中的动、名词比例和句法结构是否随时间和场景变化。

(3) 汉语儿童早期词汇中动词和名词比例是否存在差异,这种差异与成人言语输入是否一致。

由于要考察成人言语输入的特征和稳定性以及儿童早期语言的发展变化,本研究采用个案追踪的方法。纵向追踪可以全面、稳定地发现成人言语输入的规律并且可以记录儿童早期语言的发展过程。另外,本研究完全在自然环境中进行,获得的语料更接近真实而不是由实验控制。虽然个案研究样本较少,但能产生更细致真实的观察结果,可以为今后的研究提供基础和切入点。

2 研究方法

2.1 被试

本研究对一名女性婴儿(DD)进行每月一次的追踪,从单词理解初期(6个月)到进入单词迸发阶段(20个月)。由于我们的研究关注成人言语输入对儿童早期词汇发展的影响,被试包括 DD 以及她的抚养者(追踪中成人抚养者为母亲,父亲和祖母)。DD 生于北京,知识分子家庭。出生时体重 3450g,身高 52cm,无任何生理疾病。追踪过程中,DD 在 10 个月说出第一个单词;14 个月可以说出 16 个单词;17 个月开始词汇量已接近 100。整个发展过程处于正常水平。母亲为北京本地人,本科学历;父亲是河南人,研究生学历。父母间以及与孩子交流都使用普通话。

2.2 研究材料

除了自然观察,14 个月开始对 DD 进行词汇测量。测量采用 MacArthur Communicative Development Inventory(MCDI)量表的汉语版本,为 Tardif 等人修订(Hao, Liang, Wang et al., 2004a; Hao, Liang, Wang et al., 2004b; Liang, Hao, Tardif et al., 2002; Liang, Hao, Tardif et al., 2003; Tardif, Fletcher, Zhang, Liang, & Zuo, in press),用以评估语言理解和产生。MCDI 量表包含两种:在 14~17 个月时使用“婴儿沟通发展问卷”,评估婴儿对单词和手势的理解和产生,适用于词汇量较小(<150 word)的孩

子;18 个月开始,使用“幼儿沟通发展问卷”,涉及一些复杂的语法或句子,因此适用于已经可以形成多个单词联结的幼儿(通常 18~30 个月)。婴儿量表涉及 20 类单词,包括象声词、人名、动词、身体部分、动物、形容词、玩具和其它虚词等,分为“不懂”、“听懂”和“能说”三个选项;幼儿量表包含 24 类单词组成,只分为“会说”和“不会说”两项。

2.3 程序

正式拍摄前实验者与被试进行几次熟悉接触,使得被试可以在实验过程中尽量表现自然。被试从 6 个月 11 天开始接受每月一次的观察拍摄,每次历时一小时。实验开始前研究者将研究目的与内容很详细地介绍给 DD 父母,使他们可以在实验阶段尽量保持“自然”。同时,家长填写了一份“活动记录表”,确保每次上门孩子都处于清醒活动状态。整个观察集中在上午 10 点和下午 4 点左右,包括游戏、读书、吃饭等不同场景。从 14 个月开始,父母在拍摄结束后填写 CDI 问卷,耗时 20~30 分钟。每次拍摄付给被试 100 元报酬。本次分析抽取其中七个时间样本,(6 months 11 days、9 months 29 days、10 months 27 days、12 months 27 days、15 months 1days、16 months 26 days、19 months 5days),原因主要包括代表性、样本质量、成人数目、观察地点等因素,选取可以代表婴儿语言发展的关键时间段,并且尽量选择父母都在家中并且与孩子较多交流的样本,使得语料充分并且外界噪音干扰较少。样本 1 为观察的起点,并包含大量的言语交流,因此较适合作为研究的基础样本;10 个月开始,婴儿说出第一个单词,因此作为第二个分析样本;此后,婴儿词汇增长迅速,因此以两个月为间隔,选取其余 5 个时间样本;另外由于最后三个样本包含不同的场景信息,对于我们进一步分析具体场景提供帮助。拍摄器材为 Sony21E DV 摄像机,采用 SP 模式(DV 长度为 60 分钟,较清晰)。拍摄过程中研究者避免与孩子有任何交流,包括言语甚至表情和眼神,距离大约 3~5 米。环境要求尽量没有干扰声音,例如洗衣机、电视机、喧哗噪音等,也较少进行户外拍摄。

2.4 转录

录像材料通过 CLAN 程序(Computerized Language ANalysis)转录成为 CHAT 格式以便后来分析(MacWhinney, 2000; 温志军, 胡瑰玲, 2001)。CLAN 程序包含两个部分:一个是 CLAN 编辑,用于把文件转录成 CHAT 或 CA 格式;另一个部分是

一系列数据分析程序,对转录文件进行加工、分析并自动输出结果。转录每个时间样本需要 10~15 个小时,其中包括初次转写、再次校对、CHECK 命令检验、修改等过程。转写内容包括录像中出现的所有声音、动作,即孩子周围与言语相关的全部输入。平均每个时间样本包含 5% 难以清晰辨别的言语输入。

2.5 编码与分析

虽然转写过程记录了所有出现的声音刺激,本研究只涉及其中成人(父母和阿姨)与孩子进行交互的言语输入,进行以下几方面分析。

2.5.1 词类分析 对通过 CHECK 命令后的转录文件首先进行词类分析,采用 MOR 命令自动从词典中给每一个单词标明词类,通过 POST 命名从多种可能词类中选取适当的一个,正确率达 90%~95%。之后,手动进行检查和修正,进一步确定准确的词类。对于词典中没有包含的单词按照汉语词性自行分类,此词类分析通过三位专家的检验。

2.5.2 频率 经过 MOR 处理后的文件可以进一步进行词频分析和句法分析。本研究通过 Freq 命令计算出每个时间样本以及全部样本中的名词和动词比例(包括 type、token 和 type-token ratio)。Type 指不同单词的数目,Token 是单词出现的总次数,Type-Token Ratio (TTR) 指单词种类与所有出现总数的比例。例如样本中出现两个“跑”,type 就为 1,而 token 为 2, TTR 为 0.5。通常一些研究者报告“名词偏向”结果时使用的都是 type 这一指标,但由于成人言语中的 Type 和 Token 频率都与儿童早期词汇组成有很大相关(Goldfield, 1993; Tardif, Shatz, & Maigles, 1997),二者结合更能给看出输入语言的特点。

2.5.3 单词位置 不同语言研究都发现单词在句中的位置与获得的难易程度有显著相关。例如,意大利语名词相对动词更多出现在句尾,而儿童早期语言中的名词组成也多于动词(Casell, 1995);在汉语中,动词更多出现在句首或句末位置,相应地,汉语儿童的早期语言中动词比例高于名词(Tardif, Shatz, & Naigles, 1997)。本研究希望进一步探索成人言语输入中词类差异的深层原因,因此对于单词位置的检验尤为重要。

句末动词 (final verb): 句末动词结构反映了汉语句法的特殊性,是宾语省略的结果。

句首动词 (initial verb): 句首动词结构也反映了汉语句法的特殊性,是主语省略的结果。

独立动词(one verb; v): 句子只有一个动词组成(“来!”)

与动词相对应,本文也检验名词在句中的位置。

2.5.4 词素结构分析 主要分析成人言语输入的复杂性,计算每句话包含的词素数目,即 the mean length of utterance (MLU)。由于 DD 在研究阶段大多处于语言理解期和单字产生期,MLU 只针对成人言语输入进行分析。

2.5.5 具体场景分析 Tardif 等人指出成人的言语输入很大程度受到观察场景的影响(Tardif, Gelman, & Xu, 1999)。例如, Fu 等人发现英语母亲在与孩子游戏的 15 分钟内说出的动词比例显著高于名词,而 Goldfield 的研究显示母亲在游戏场景说出的名词多于动词,非游戏场景下动词多于名词。对于这一冲突结果, Tardif 等人认为与游戏场景的安排和刺激使用有关,后者场景中提供的刺激材料可能更容易引起对客体的命名。因此, Tardif 等人采用两种语言被试,在观察场景完全控制的条件下测量成人和孩子的言语产生。结果发现,无论英语还是汉语被试,在阅读场景下说出的名词显著多于动词,而游戏场景恰恰相反。鉴于观察场景的重要性和特殊性,本研究试图分析被试在不同场景下的言语交流,希望发现其中的差异。因此,最后三个时间样本中的三个场景(认图、吃饭和游戏)被再次进行言语分析。选择最后三个样本的原因主要有两个:其一是三个样本分别包含不同的三个场景;其二是 DD 在后三个样本中言语输出明显增加,可以更好地分析成人与孩子的交流。

2.6 MCDI 量表分析

MCDI 量表分析包含对动、名词比例的统计。婴儿量表中包含 78 个动词项和 188 个名词项。名词项涉及更具体的范畴,包括食物、身体部分、动物、用品、玩具、衣服、家具、外面的东西和车 9 类。本研究考察被试对全部动、名词理解和能说的比例。其中单词理解包括量表标记的“听懂”和“能说”两项。幼儿量表包含 194 个动词和 322 个名词,研究只考察 DD 能说出的动、名词比例。

3 结果

3.1 成人言语输入分析

时间样本 1

言语样本中共包含动词 126 种(type), 612 个(token), 平均 TTR 为 0.21; 名词共 102 种(type), 264

个(token), 平均 TTR=0.39。因此, 无论 type 还是 token 动词比例都高于名词。进一步卡方检验显示, 动词和名词在 type 上无显著差异($\chi^2=2.53, p>0.05$), token 差异显著, $\chi^2=138.25, p<0.01$ 。

在所有动词中, 句首出现 202 次, 句末出现 80 次, 独立动词 29 个, 因此所有出现在句中突显位置的动词占有所有动词比例的 51%; 所有名词中, 句首出现 23 次, 句末出现 92 次, 独立名词 9 个, 所有处于突显位置的名词占有所有名词的 47%。父母说出的平均句子长度(MLU)分别为 3.53 和 3.24。

时间样本 2

时间样本 2 为 DD9 个月 29 天, 此次成人言语输入中共包含动词 77 种, 418 次, 平均 TTR 为 0.18; 名词 59 种, 148 次, 平均 TTR 为 0.40。卡方检验结果显示, 动词和名词在 type 上无显著差异($\chi^2=2.38, p>0.05$), token 上差异显著($\chi^2=128.80, p<0.01$)。

单词位置和句法结构分析显示, 动词出现在句首 112 次, 句末 55 次, 独立出现 20 次, 所有出现在句中突显位置的动词占有所有动词的 45%; 名词出现在句首 17 次, 句末 43 次, 独立出现 10 次, 所有处于突显位置的名词占有所有名词的 47%。父母说出的平均句子长度(MLU)分别为 3.20 和 3.38。

时间样本 3

此次追踪为 DD10 个月 27 天, 成人言语输入中包含动词 88 种, 303 次, 平均 TTR 为 0.29; 名词 61 种, 123 次, 平均 TTR 为 0.50。进一步卡方检验发现, 动词和名词在 type($\chi^2=4.89, 0.01<p<0.05$)和 token($\chi^2=76.06, p<0.01$)上都达到显著。

单词位置分析显示, 动词出现在句首 81 次, 句末 40 次, 独立出现 8 次, 所有出现在句中突显位置的动词占有所有动词的 43%; 名词出现在句首 13 次, 句末 27 次, 独立出现 11 次, 所有处于突显位置的名词占有所有名词的 41%。父母说出的平均句子长度(MLU)分别为 4.28 和 3.19。

时间样本 4

第四个时间样本为 DD12 个月 27 天。成人言语输入包括母亲和阿姨, 产生的动词为 60 种 218 次, 名词为 41 种 83 次, 平均 TTR 分别为 0.28 和 0.49。卡方检验显示, 动、名词频数在 type 上无显著差异($\chi^2=3.57, p>0.05$), token 上差异显著($\chi^2=60.55, p<0.01$)。

在所有动词中, 句首动词出现 68 次, 句末动词出现 42 次, 独立动词 12 次, 共占 56.0%; 所有名词

中, 句首名词出现 9 次, 句末名词出现 38 次, 独立名词 13 次, 共占 72.3%。母亲说出的平均句子长度为 2.48, 阿姨为 3.87。

时间样本 5

第五个时间样本为 DD15 个月 1 天。本次拍摄只有母亲一人看护, 主要场景是读画报和闲玩。母亲言语输入中共包含 70 种 250 次动词(平均 TTR 为 0.28)和 70 种 216 次名词(平均 TTR 为 0.32), 在 type 和 token 上动、名词都没有显著差异($\chi^2=0.00, p>0.05$; $\chi^2=2.48, p>0.05$)。此样本中母亲平均句子长度(MLU)为 2.76。

所有动词中, 句首动词出现 75 次, 句末 28 次, 独立动词 17 次, 突显位置的动词占全部动词的 48%; 所有名词中, 句首名词 17 次, 句末名词 84 次, 独立名词 48 次, 所有突显位置的名词占全部名词的 69%。

时间样本 6

此次时间样本中(16 个月 26 天)成人言语输入包含动词 121 种 685 次, 平均 TTR 为 0.18; 包含名词 98 种 310 次, 平均 TTR=0.32。卡方检验显示, 动、名词 type 间无显著差异($\chi^2=2.42, p>0.05$), 动词 token 显著高于名词($\chi^2=141.30, p<0.01$)。另外, 父亲的平均句子长度(MLU)为 2.97, 母亲为 3.31。

单词位置分析显示, 句首动词出现 207 次, 句末动词出现 95 次, 独立动词 35 个, 因此所有出现在句中突显位置的动词占有所有动词比例为 49%; 句首名词 32 次, 句末名词 93 次, 独立名词 52 次, 所有突显位置的名词占有所有名词的 57%。

时间样本 7

本研究最后一个时间样本为 19 个月 5 天。成人言语输入中共包含 166 种动词, 共 872 次(平均 TTR=0.19); 150 种名词, 共 364 次(平均 TTR=0.41)。动、名词在 type 上没有显著差异($\chi^2=0.81, p>0.05$), 而在 token 上, 动词显著高于名词($\chi^2=208.80, p<0.01$)。父亲和母亲的平均句子长度(MLU)分别为 3.35 和 3.77。

单词位置和句法结构分析结果显示, 在所有动词当中, 句首动词共 225 个, 句末动词 173 个, 独立动词 43 个, 所有位于突显位置的动词占全部动词比例为 51%; 所有名词中, 句首名词 35 个, 句末名词 106 个, 独立名词 47 个, 所有突显位置名词占全部名词的 52%。

图 1 和图 2 为 7 个时间样本中成人言语输入的动、名词频率统计。

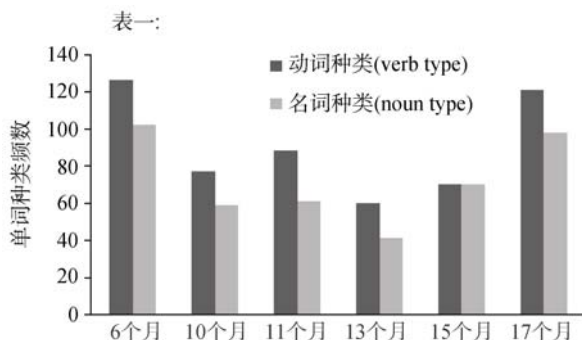


图1 成人言语输入中动词和名词的 type 频率

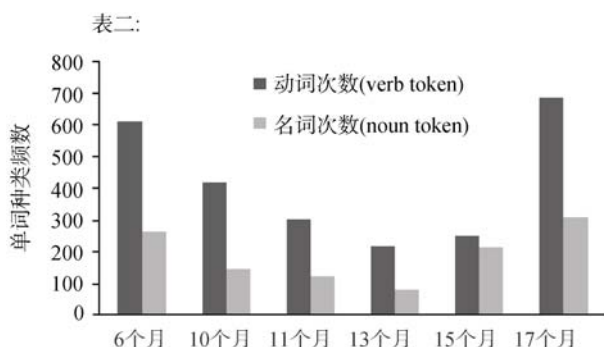


图2 成人言语输入中动词和名词的 token 频率

3.2 婴儿词汇发展结果

整个追踪过程中, DD 说出 18 种动词和 24 种名词, 并且在 15 个月前动词数目要多于名词(见图 3)。DD 在 10 个月(时间样本 2)开始说出第一个单词“打”, 共 56 次。13 个月时说出 6 种不同单词 18 次, 分别为“打”、“来”、“阿姨”、“妈妈”、“这”、“个”。15 个月时, DD 不仅能够说出“妈妈”、“爸爸”、“姨”等人称词, 言语样本中还包括“打”、“拿”、“背”、“走”四个动词和“笔”、“门”、“鸭”、“菜”、“牙”五个名词。在第 17 个月的观察中, DD 共说出单词 60 次, 17 个种类(爸爸、妈妈、姨、娃娃、球、水、冬瓜、瓜瓜、臭臭、拿、尿、没、踢、热、凉、这、个)。其中, 动词 4 个, 名词 6 个, 人称词 3 个, 其它词类 4 个。在 19 个月时, DD 说出 44 个不同单词, 共 142 次。这些单词中, 动词包括 13 类, 28 次, 名词 15 类, 34 次。其中次数最多的 11 个单词依次为: 个、妈妈、这、一、笔、拿、好、爸爸、包包、洗、没。

由于观察样本中 DD 的言语输出受到场景影响, 单词理解能力也难以测量, 因此本研究进一步对 CDI 测量结果进行分析。CDI 量表包含不同年龄段婴儿和幼儿词汇组成中相对容易和困难的不同词类单词, 能够较全面的了解词汇发展的不同层面

(单词理解和单词产生)。自然观察结合 CDI 量表结果, 能够更准确了解被试语言发展的真实情况。

时间样本 5

在 MCDI 量表包含的 78 个动词中, DD 可以听懂 66 个单词, 能说出 6 个(“抱”、“打”、“拿”、“飞”、“喂”和“下”)。动词理解(听懂+能说)占有所有动词的 92.3%。在 188 个名词中, DD 可以听懂 124 个, 能说出 5 个(“牛奶”、“水”、“帽子”、“袜子”、“花”), 名词理解比例占 69%。动、名词相对比例为 1.34:1(动词理解比例/名词理解比例)。

时间样本 6

DD 在此次测量中听懂的动词为 62 个, 能说出 14 个(除上次 6 个外, 还包括“睡觉”、“尿”、“要”、“撕”、“踢”、“玩”、“背”和“掉”), 理解动词总数为 76 个, 占 97%。在 188 个名词中, DD 能说出 43 个, 可以理解其中的 79%, 动、名词相对比例为 1.28:1。

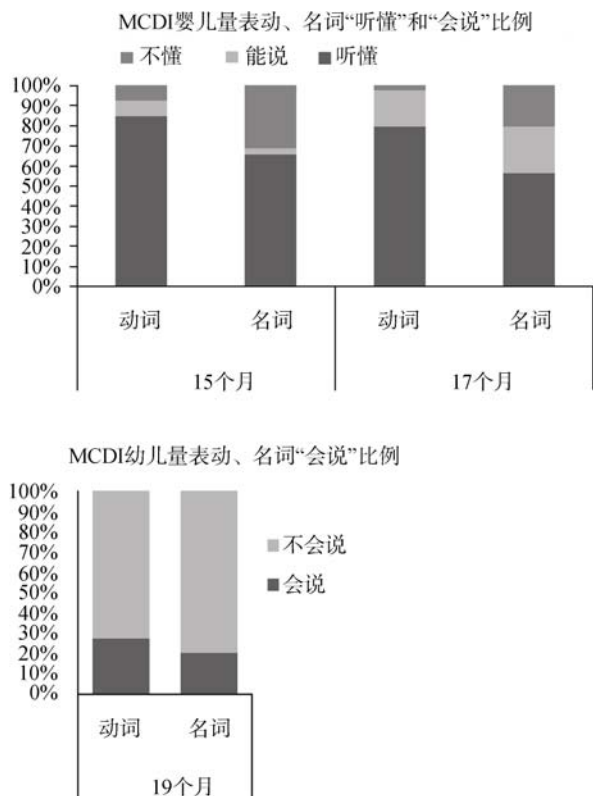


图3 MCDI 词表结果

时间样本 7

之前的理解测量显示 DD 在 17 个月左右几乎可以理解所有的动词, 因此此次测量采用幼儿问卷。幼儿问卷测量显示, DD 在 194 个动词中可以说出其中的 53 个, 占 27%。相比时间样本 6 中能说出

出的动词比例(17%)有很大提高;在 322 个名词中可以说出其中的 65 个,占 20%(样本 5, 样本 6 中能说出的名词比例分别为 4%和 23%)。能说出的动词、名词相对比例为 1.35:1。

因此,与成人言语输入一致,DD 在三次 MCDI 测量中动词与名词获得的相对比例较稳定,且动词数目高于名词。

3.3 具体场景分析

认图场景

对时间样本 5 中认图场景进行分析结果显示,母亲言语输入中动词数目为 28 次,包含 11 个种类(依次为“喝”、“是”、“说”、“想”、“翻”、“给”、“在”、“来”、“带”、“到”和“完”);名词共出现 83 次,包含 29 个类别,其中出现最多次数的 8 个单词依次为“熊”、“娃娃”、“吉普车”、“水”、“火车”、“球”、“兔兔”、“鸭鸭”。无论 type 还是 token 名词频率都显著高于动词($p<0.01$)。DD 在此过程中没有说出单词(语气词除外)。

从母亲的具体语句分析来看,73 个有效语句中(除只包含语气词的句子),47 句为客体命名(例如:“娃娃!”“这是剪刀。”),占 64%,9 句为疑问句(例如“这是什么呀?”“两个小熊一样不一样啊?”),8 句为祈使句(例如:“给娃娃喝水!”“小熊喝水!”)以及 9 句一般陈述句(例如:“兔兔说不喝不喝。”)。这些语句大多数与图片上的客体有关,因此包含名词较多。

吃饭场景

时间样本 3 包含部分吃饭场景,从结果来看,父母的言语输入中共包含 37 类 97 次动词,出现次数最多的 11 个动词依次为“热”、“吃”、“吐”、“来”、“放”、“吹”、“是”、“出来”、“摸”、“给”和“掉”。相对动词,此场景中名词出现的次数较少,共 20 类 42 次,其中出现最多的为“肉肉”和“饭”(除“爸爸”、“妈妈”)。 χ^2 检验显示,动词频率显著高于名词(type: $0.01<p<0.05$, token: $p<0.01$)。另外,DD 在此场景中共说出 4 类 15 个单词,依次是“这”(6 次)、“个”(6 次)、“拿”(2 次)、“妈妈”(1 次)。

父母的具体语句分析发现,在 87 个有效语句中,祈使句共 42 句,占 48%;一般陈述句 32 句,占 37%;疑问句 8 句;命名 5 句。大多数语句都与吃饭行为相关,较少涉及吃的东西,因此语句中动词比例更高。

游戏场景

时间样本 7 包含父母与 DD 玩电脑的游戏场

景。言语输入分析结果显示,此场景中共出现 39 类 100 次动词和 17 类 28 次名词。动词出现最多的为“看”和“敲”;名词次数较分散,除“眼睛”(4 次)、“画笔”(3 次)、“键盘”(2 次)、“点儿”(2 次)其余都只出现一次(除人称词)。动词的 type 和 token 都显著高于名词(type: $\chi^2=8.64$, $p<0.01$; token: $\chi^2=40.50$, $p<0.01$)。DD 在游戏场景中共说出单词 15 次(7 类),分别是“好”、“要”、“一”、“二”、“爸爸”、“敲”、“没”。

游戏场景中父母与 DD 交流的语句共 59 个(除去语气词和父母本身的对话),祈使句 29 句,占 49%,一般陈述句 18 句,占 30%,疑问句 2 句。大多数语句与游戏动作相关,所有语句都包含动词,名词没有单独出现。

4 讨论

本研究通过纵向追踪的方法探讨了汉语成人言语输入的动、名词频率和句法结构特点以及儿童早期词汇发展的动、名词差异。具体结果显示,除时间样本 5(认图场景),成人言语输入中动词的次数(token)和种类(type)都高于名词。这一结果与英语成人言语输入形成对比(例如,Goldfield, 1993)。成人言语输入中较高动词比例的特征影响婴儿早期词汇获得和发展,在本次追踪过程中虽然 DD 词汇量有限,但很早已经可以说出一些行为动词(10 个月),并且动词占据较大比例。相比下,英语婴儿追踪结果发现,被试在 16 个月第一次说出动词,并且名词比例远远高于动词(Tardif, Casasola, Choi, & Meng, 2006)。因此,本次研究结果对名词优势普遍存在的现象提出质疑,因此进一步支持 Tardif 等人提出的语言特异性(Choi & Gopnik, 1995; Tardif, 1996; Tardif, 2006; Tardif, Gelman, & Xu, 1999)。

4.1 成人言语输入特征

汉语成人言语输入中动词比例较高与动词在句中的位置有关。在总体时间样本中处于句首、句末和独立结构的动词占有所有动词的 49%。这种省略主语和宾语的句法结构使得名词比例远远减少。

除了动词优势特征,成人言语输入还具有时间上的稳定性和具体场景的变化性。在七次时间样本中,单词的平均 TTR、动词和名词的相对比例(v/n)、平均句子长度具有稳定性,没有随婴儿的语言发展而变化。另外,在不同的场景中,成人输入的动、名词相对比例具有显著差异:在“阅读”场景下,成人主要指导孩子对图片命名,因此名词在言语输入中占绝大比例(n/n+v=75%);相反,在“吃饭”和“游

戏”场景下,成人更多指导孩子去完成一些动作,因此多以祈使句方式出现动词,使得动词相对比例分别为70%和78%(v/n+v)。这一结果与Tardif(1999)的研究较为一致。

以上分析表明汉语成人言语输入具有明显的动词优势。那么这种语言特异性是否会影响儿童早期词汇获得呢?

4.2 婴儿早期词汇获得

虽然DD在追踪过程中说出的名词种类和数目要多于动词,但动词最早出现的时间和频率要远远早于和高于英语儿童。这一结果说明动词学习并非如此困难,汉语婴儿在词汇发展初期已经可以掌握相当数目的行为动词,因此对“名词优势”普遍理论提出质疑。

在本研究阶段,DD的词汇水平在15个月开始呈现显著增长趋势,观察场景中可以说出12个不同的单词。综合CDI测量结果,DD至少可以说出9种动词和9种名词,并且理解量表上92.3%的动词和69%的名词,因此也进一步证明名词优势并不普遍存在于儿童早期词汇中。17个月后,DD的词汇水平有了进一步发展,除了人称词、主动词和普通名词外,观察过程中也说出几个形容词和量词。在13类说出的单词中,动、名词比例没有很大差异。由于单词数量较少以及受到说话动机的影响,词汇分析需要进一步结合CDI结果。测量结果显示,DD可以说出14个动词和43个名词,占量表中所有动词和名词的18%和23%;理解的动词和名词分别占量表的97%和79%。因此,从单词理解水平来看,DD早期的词汇组成中动词仍占据优势。19个月开始,DD的词汇水平出现明显进步,在观察过程中可以说出37类单词。CDI分析进一步发现,DD能说出53种动词(占27%)和65种名词(占20%),基本进入词汇发展高峰期。此时能说的动词比例高于名词,再次质疑名词优势论(Gentner, 1982)。因此,DD的早期词汇中虽然能说的动词并不多于名词,但动词理解比例仍占据优势,与成人言语输入的动词优势特征一致。陈萍等人早期进行的追踪研究也发现,汉语儿童最初词汇中动词出现的较早,而且也占据一定比例(陈萍,许政缘,1993)。

针对这种现象,我们认为,首先,成人言语输入中绝大多数说动词位于句中突显位置,使得动词在知觉上更容易获得(Bornstein, Cote, Maital et al., 2004; Casell, 1995; Tardif, Shatz, & Maigles, 1997)。其次,儿童早期词汇获得受输入频率的影响

(Bornstein, Cote, Maital et al., 2004; Gentner, 1982; Tardif, Shatz, & Maigles, 1997),汉语的特殊句法结构使得动词频率明显增加。最后,成人言语输入中动、名词的平均TTR也具有一些特点:动词出现的type/token比例远小于名词(动词平均TTR为0.21;名词平均TTR为0.39)。这一特点说明不仅成人言语输入中动词的种类和次数都远远高于名词,每一类动词重复出现的次数也更多。这种多次重复的特征使得儿童获得动词更容易。

本研究中DD早期的词汇发展不仅反对名词偏向论,也对ECM理论提出质疑。ECM认为儿童语言发展早期主要采用知觉线索和天生的认知偏向获得单词,直到19个月后才出现社会线索的作用。但是,本研究发现在婴儿最初获得单词开始,社会和语言线索就对词汇发展产生影响。

4.3 结论与局限

汉语作为表意文字,在结构上与拼音文字有很大差别,这种语言的特殊性会影响到认知加工(陈新葵,张积安,2008)。本个案研究通过纵向追踪和详细分析,发现了汉语成人言语输入的特殊性,无论在单词频率、位置、句法结构上动词都处于优势地位。同时,这种言语输入与儿童早期词汇获得存在很大相关,进一步支持“语言特征”和“社会线索”的作用。

个案研究在儿童语言发展领域相当普遍。近10年来,国际心理学研究相关数据库中已包含200余篇儿童语言发展的个案研究,涉及语音,词汇,句法等多个层面,并且覆盖第一语言,双语,非正常儿童等多重样本。虽然个案研究存在样本数量、分布和代表性等问题的局限,但此研究采用纵向追踪可以在自然环境中真实、全面、稳定地发现成人言语输入的规律并且可以记录儿童早期语言的发展过程。目前,国际上一些关于成人言语输入对婴儿早期词汇获得影响的研究主要采用自然或半自然观察方法,其发现与本研究结果一致(Naigles, 2000; Naigles & Hoff-Ginsberg, 1998; Tardif, Shatz, & Naigles, 1997)。例如:Goldfield对12对英语婴儿与母亲的交流活动进行观察,发现母亲说出的名词比例高于动词,而这一差异与孩子说出的单词比例有显著正相关(Goldfield,1993);Tardif等人(1997)对六对英语、六对意大利语和十对汉语儿童与母亲自然交流的样本进行分析,发现母亲言语输入中的动、名词比例与儿童自发说出的动、名词比例一致,而汉语母亲动词使用频率更高。因此,本研究结果

从成人言语输入影响、婴儿词汇发展趋势以及婴儿词汇获得的跨语言差异等多方面提供了证据, 对大样本研究奠定了经验基础。建议之后的研究方向可以进行大样本追踪观察, 进一步考察成人言语输入在儿童语言发展过程中的重要作用。

参 考 文 献

- Au, T. K. F., Dapretto, M., & Song, Y. K. (1994). Input versus constraints: Early word acquisition in Korean. *Journal of Memory and Language*, 33, 567–582.
- Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S. et al. (1994). Development and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, 21, 85–123.
- Bornstein, M., Cote, L., Maital, S., Painter, K., Park, S. Y., Pascual, L., et al. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean and American English. *Child Development*, 75, 1115–1140.
- Casell, M. C., Bates, E., Casadio, P., Fenson, J., Fenson L., Sanderl, L., et al. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, 10 (2), 159–199.
- Chen, J., Tardif, T., & Meng, X. Z. (2007). Cross-linguistic differences in children's early word acquisition. *Advances in Psychological Science*, 15 (3), 423–428.
- [陈杰, Tardif, T., 孟祥芝. (2007). 儿童早期词汇获得的词类差异, *心理科学进展*, 15 (3), 423–428.]
- Wen, Z. J., & Hu, G. L. (2001). Exploiting the world's largest child language database—CHILDES. *Foreign Language Teaching and Research*, 33 (5), 374–377.
- [温志军, 胡瑰玲. (2001). 开发利用世界上最大的儿童语料库—CHILDES. *外语教学与研究*, 33(5), 374–377.]
- Chen, P., & Xu, Z. Y. (1993). Initial Word Acquisition in Children and its Process. *Acta Psychologica Sinica*, 25(2), 195–202.
- [陈萍, 许政缘. (1993). 儿童最初词汇的获得及其过程. *心理学报*, 25(2), 195–202.]
- Chen, X. K., & Zhang, J. A. (2008). Review and comparison of influencing factors on Chinese nouns and verbs' processing. *Journal of South China Normal University*, 4, 119–123.
- [陈新葵, 张积安. (2008). 汉语动词、名词认知的影响因素比较. *华南师范大学学报*, 4, 119–123.]
- Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language*, 22, 497–529.
- Fernandes, K., Marcus, G., Nubila, J., & Vouloumanos, A. (2006). From semantics to syntax and back again: Argument structure in the third year of life. *Cognition*, 100 (2), B10–B20.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development: Language, cognition and culture*. (2, pp. 301–334). Hillsdale, NJ: Erlbaum.
- Goldfield, B. A. (1993). Noun bias in maternal speech to one-year-olds. *Child Language*, 20, 85–99.
- Golinkoff, R., & Hirsh-Pasek, K. (2006). Baby Wordsmith: From associationist to social sophisticate. *Current Directions in Psychological Science*, 15, 30–33.
- Golinkoff, R., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language*, 21, 125–155.
- Hao B., Liang, W. L., Wang, S., Tardif, T., Fletcher, P. et al. (2004a). Individual and family factors influencing vocabulary development in toddlers. *Chinese Journal of Pediatrics*, 42(12), 908–912.
- Hao, B., Liang, W. L., Wang, S., Tardif, T., Fletcher, P. et al. (2004b). A study on related factors of phrase and sentence development in toddlers. *Chinese Journal of Child Health Care*, 12(5), 379–381.
- Hirsh-Pasek, K., & Golinkoff, R. (1996). *The origins of grammar: Young children's use of syntactic frames to derive meaning*. Chap. 6. Cambridge, MA: MIT Press.
- Hollich, G. J., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). Breaking the language barrier: an emergentist coalition model of the origins of word learning. *Monographs of the Society for Research in Child Development*, 65(3), 1–123.
- Kim, M., McGregor, K. K., & Thompson, C. K. (2000). Early lexical development in English and Korean-speaking children: Language-general and Language-specific patterns. *Journal of Child Language*, 27, 225–254.
- Lee, J. N., & Naigles, L. R. (2005). The input to verb learning in Mandarin Chinese: A role for syntactic Bootstrapping. *Developmental Psychology*, 41(3), 529–540.
- Liang, W., Hao, B., Tardif, T., Fletcher, P., & Wang, Y. (2002). Vocabulary development of toddlers in urban Beijing. *Chinese Journal of Pediatrics*, 40(11), 650–653.
- Liang, W. L., Hao, B., Tardif, T., & Fletcher, P. (2003). Study on early syntax and expression of syntactic length in Chinese children. *Chinese J of Child Health Care*, 12(3), 206–208.
- MacWhinney, B. (2000). *The CHILDES project: tools for analyzing talk*. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates.
- Maguire, M., Hirsh-Pasek, K., & Golinkoff, R. (2006). An unified theory of word learning: putting verb acquisition in context. In K. Hirsh-Pasek, R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. (pp. 364–392). Oxford University Press.
- Naigles, L. (1990). Children use syntax to learn verb meanings. *Child Language*, 17, 357–374.
- Naigles, L. (2000). Manipulating the input: studies in mental verb acquisition. In B. Landau & J. Sabini (Ed.), *Perception, cognition, and language: Essays in honor of Henry and Lila Gleitman*. Cambridge, MA: The MIT Press.
- Naigles, L., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language*, 25, 95–120.
- Nelson, K. (1988). Constraints on word learning. *Cognitive Development*, 3, 221–246.
- Pruden, S. M., Hirsh-Pasek, K., & Golinkoff, R. (2006). The birth of words: ten-month-olds learn words through perceptual salience. *Child Development*, 77, 266–280.
- Snedeker, J., & Gleitman, L. R. (2004). Why is it hard to label our concepts? In D. G. Hall & S. R. Waxman (Eds.) *Weaving a lexicon*. Cambridge, MA: MIT Press.
- Tardif, T. (1996). Nouns are not always learned before verbs:

- Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology*, 32, 492–504.
- Tardif, T. (2006). But are they really verbs? Mandarin words for action. In K. Hirsh-Pasek, R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. Oxford University Press.
- Tardif, T., Casasola, M., Choi, S., & Meng, X. Z. (2006). Cultural bases and biases for word-learning. *Unpublished manuscript*. Ann Arbor, Michigan.
- Tardif, T., Fletcher, P., Zhang, Z. X., Liang, W. L., & Zuo, Q. H. (in press). *The Chinese Communicative Development Inventory (Putonghua and Cantonese versions): Manual, Forms, and Norms*. Peking University Medical Press.
- Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the "noun-bias" in context: a comparison of English and Mandarin. *Child Development*, 70, 620–635.
- Tardif, T., Shatz, M., & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language*, 24, 535–565.

The Influence of Adult Input on Children's Early Word Learning: A Case Study of A Mandarin-Speaking Child

CHEN Jie^{1,2}, SETOH, Pei-pei^{3,4}, MENG Xiang-Zhi¹, TARDIF Twila²

(¹Department of Psychology, Peking University, Beijing 100871, China)

(² Department of Psychology, University of Michigan, USA)

(³ Department of Psychology, Nanyang Technological University, Singapore)

(⁴ Department of Psychology, University of Illinois at Urbana-Champaign, USA)

Abstract

Children's early vocabulary development is not linear. At the outset, word learning is very slow. However, by approximately 19 months of age, children's vocabulary rapidly expands, entering the phase of the "word spurt". The phenomenon of breaking word learning's bottleneck can be interpreted by several theories, such as constraint theories which emphasize innate cognitive biases, social-pragmatic theories emphasizing the role of social and linguistic environments, and associationistic views involving computations of the co-occurrence between words and their referents in naturalistic speech. However, these theories cannot account by themselves for cross-linguistic differences or similarities across children, despite differences in input—didn't really understand this parenthesis. Although the updated theory of Emergentist Coalition Model (ECM) combines early use of attentional cues with later use of social inputs and linguistic cues, it cannot account for the consistent cross-linguistic differences appearing at the very beginning of children's vocabulary which correspond to linguistic features and social inputs. Such differences, particularly the composition of verbs and nouns in children's early vocabularies, challenge the theory of "Noun bias" supported by many researchers. Thus, whether children can use social and linguistic cues in the beginning of word development, noun and verb acquisition in particular, is of interest for this paper.

Although there are many studies investigating the relationship between adult's speech input and children's early language development, they do not reveal the regularity and developmental changes both in caregivers' input and infants' word acquisition. Furthermore, none of these studies focused on Chinese, a language with vastly different linguistic properties from English.

A longitudinal case method with a Chinese female infant was used. Tracking the infant from 6- to 20-months of age, the researcher visited the family monthly for one-hour recordings of naturalistic interaction. The naturalistic data on caregiver-to-child input were transcribed into CHAT format and analyzed with the CLAN program, counting the frequency of nouns and verbs and word position in caregivers' utterances. Then, the development of child's comprehension and production vocabulary was assessed by both observed data and the Putonghua version of the MacArthur Communicative Development Inventory(PCDI). Seven spontaneous

speech samples (6, 10, 11, 13, 15, 17 and 19 months) were analyzed. In addition, the study also focused on the roles of specific contexts of caregivers' input, selecting three contexts in the naturalistic data (booking reading, feeding and playing).

The caregivers were found to produce more verb than noun in tokens and types in most samples, although the difference was significant only for tokens. The position in caregivers' utterances was favorable for verbs, but not nouns, with many verbs dropping subjects and objects. However, the ratio of verbs and nouns varied with specific context. The frequency of nouns was higher relative to verbs in the booking reading context, but verbs were more frequent in both the feeding and playing contexts. In addition, more verbs than nouns were found in the child's early vocabulary, regardless of how they were measured and this corresponded to the caregivers' speech input. Specifically, the child could both comprehend and produce higher ratios of verbs than nouns with the CDI measure, and produced a larger cumulative number of verbs than nouns up to the 15 month.

Children can use linguistic input in their vocabulary acquisition even at the very beginning of language development. The dominance of verbs in Chinese adults' speech influences children's early word composition, with more verbs than nouns in both comprehension and production. This result again supports the view of cross-linguistic differences but not the "noun bias" in children's early vocabularies. In addition, the frequency of verbs and nouns varies with specific context.

Key words adult input; verb dominance; CLAN