

# 关于有声思维法口译研究<sup>\*</sup>

香港理工大学 李德超 广东外语外贸大学 王巍巍

**提要:** 本文回顾过去二十余年来口译研究中运用有声思维法的发展历程,总结并介绍有声思维口译研究在口译过程、口译策略及口译质量方面所取得的进展。随后,在参考 Lincoln & Guba (1985)提出的自然主义研究信度衡量标准的基础上,制订出了有声思维口译研究的量尺,并以此来考察 9 篇相关文献的总体研究设计,发现这些研究在信度与效度方面都存在不少不足。本文亦对将来的有声思维口译研究如何确保信度和效度方面提出了建议。

**关键词:** 有声思维、口译研究、自然主义研究

[中图分类号] H059 [文献标识码] A [文章编号] 1000-0429(2011)06-0900-11

## 1. 引言

有声思维法(think-aloud protocols, 简称 TAPs)源于认知心理学的内省法。近三十年来,它作为翻译过程的实证研究方法之一进入译学视野。研究者尝试通过口头报告的方式,揭示笔译思维过程,探索当中涉及的翻译策略、翻译步骤、翻译单位、注意力分配、情感态度等因素(Dancette 1994; Fraser 1996)及其对翻译教学的启示(Wakabayashi 2003; Li 2011)。但这一方法在口译研究中却较为少见,主要原因是口译独特的工作模式造成的困难,即译员通常无法在口译的同时描述或报告自己的思维过程。

对于有声思维法在口译研究中的适用性,学界有些质疑。一是受试会记住

---

\* 本研究得到香港理工大学项目“The application of think-aloud protocols in consecutive interpreting teaching at undergraduate level: Insights from professional interpreters”(LTG09-12/SS/CBS)的资助,特此鸣谢。

和报告“有意识”的口译决策及当中涉及的认知过程,而忽略那些已经“自动化”的口译过程(Ericsson & Simon 1980);二是就算“有意识”的口译过程,受试也不一定都能记得,或者因为某些原因不进行口头报告;三是口头报告数据的可靠性受到受试个人价值判断的影响,例如一些非策略性行为可能被误报为口译策略等(Gile 1997)。但支持者认为,首先,运用该方法的实证研究结果能够反映出口译过程中的重要思维信息,例如译员注意力分配与工作记忆的关系、口译表现与意外问题处理、源语转换对口译自动化过程的影响等等;其次,在对受试进行适当培训,以及研究者合理介入等前提之下,有声数据本身及实验结果的可靠性可以提高(Kalina 1998; Shlesinger 2000)。本文拟对过去二十年来西方使用有声思维法进行的口译研究情况作一回顾,检讨其得失,并讨论改进的方法。

## 2. 有声思维法口译研究概述

因口译工作模式的特殊性,有声思维法口译研究无法采用笔译研究中常用的同步性的“内省法”来观测译员的思维过程,取而代之的是在时间上相对滞后的“一般反省法”或“即时反省法(immediate retrospection)”。前者是指译员在口译任务完全结束之后进行口头报告;而后者是指译员在口译任务过程中以句子、段落或意群为单位稍作停顿后的即时口头报告。这两种“反省法”中,后者兼具“同时性”内省法和“追溯性”反省法两者的特征,这与人类学研究中的“时间凝固法”(time freezing)相似(Shlesinger 2000)。

如依照其研究内容划分,有声思维口译研究主要集中于口译过程、口译策略及口译质量这三类研究,受试涵盖大学不同年级的口译学生(Mead 2002; Bartłomiejczyk 2006)、专业译员(Kohn & Kalina 1996; Pio 2003; Monacelli 2009)、口译专业毕业生及兼职译员(Ng & Obana 1991)等。研究涉及的语言组合均为西方语言,包括“英语-保加利亚语”(Ivanova 2000)、“英语-意大利语”(Mead 2002)、“芬兰语-瑞典语”(Vik-Tuovinen 2002)等。

### 2.1 口译过程研究

有声思维法作为一种过程导向的研究方法,可以在口译教学中用于探究技能习得过程中“黑匣子”的运作机制。例如,Ng & Obana (1991)采用内省法让学生译员描述自己在口译练习中大脑所思所想,以探究英日交传中译员遇到的难点及问题。她发现不同语系的语言组合(如“日语-英语”)的口译任务比相同语系(例如“法语-英语”)难度更大。Vik-Tuovinen (2002)尝试运用反省法来研究同声传译的过程。在参照 Gile (1992)精力分配模式(Effort Model)中提到的口译过程及其细节要素的基础上,他对比了14名学生译员与7名资深译员自

我评价的有声思维资料。他发现,学生译员的“口译资历和表现”与“针对口译技巧的自我评价”成反比,而资深译员则更多地从“转换”及“产品”的层面进行评论。

有声思维方法亦用于口译测试评估之中,以识别学生口译技能的发展阶段。Choi(2006)回顾了目前交传教学中常用的学期测试及课堂评估方法,认为其在信效度方面存在局限。她提出“元认知”的教学评估方式,采用有声思维方法中的“反省法”来鼓励学生自我评价、检讨口译表现,而教师则给予辅助意见。学生通过自我衡量课堂及练习当中的口译表现,来发现自身的不足,并自行总结出个人需提升的口译技能的优先次序。教师则从形成性评价及诊断性评价的视角,根据学生个性化的学习曲线进行监督和评估。该研究打破了有声思维作为收集实证资料的研究方法的定式思维,将其作为一种评估手段引入口译教学,以发展式的评估框架为教学测试提供了新维度,值得同行借鉴。

## 2.2 口译策略研究

“翻译策略”是有声思维法笔译研究的重点之一,尤以 Lörcher (1991) 的论述影响为大。他在 Föerch & Kasper (1983) 交际策略观点的基础上,把“翻译策略”定义为“译者将文本片段从一种语言翻译到另一种语言过程中,解决遇到问题的有意识的过程”(Lörcher 1991: 76)。有声思维法口译研究者亦借鉴这一策略定义,对口译策略种类和特点作了深入研究。

Kohn & Kalina (1996) 通过采用内省法收集了同传译员的有声思维数据来对比单语交际及口译双语交际涉及的心理语言因素,发现口译过程是策略性的话语处理过程,译员采用口译策略时均是有意识和有明确目的。在此基础上, Kalina (1998) 在口译员当中进行了一项小规模即时反省法研究,让译员回顾其口译策略的使用,以识别口译策略及其分类。她发现,口译策略在一定程度上与单语交际策略相关联。如果口译员能成功地在实践中重复使用特定口译策略,并使之自动化,这可以减轻译员口译时的认知负担。

然而, Bartłomiejczyk (2006) 却认为,以上口笔译研究中所提到的“翻译策略”定义大都是“问题导向”策略,类似于二语习得理论中的“交际策略”。例如,在同声传译过程中,译员会在注意力不足等情况下,采用一定口译策略(如总结、概括)以避免潜在问题。他将口译策略定义为“一种有助于解决特定口译问题的方法,它能够辅助译员完成口译任务及预防潜在问题发生”(同上: 152)。他认为,口译策略包括“问题导向策略”及“整体策略”两种。前者指“可能有助于解决特定口译问题的策略”,后者指“辅助译员完成口译任务的一般性策略”(同上: 152)。他继而运用反省法来识别同声传译中译员所用的不同策略,并统计各种

口译策略的使用频率。

### 2.3 口译质量研究

口译产品从产出到完成的过程转瞬即逝,这为相关研究带来了困难。有声思维法中的语料录制和口头报告可以将这一过程“凝固”、“回放”,让译员有时间运用(即时)反省法回顾自己的思维过程,研究者也可以对口译语料及转写的有声思维数据进行条分缕析的对比及评价。比如,流畅度是口译质量的重要因素之一,而迟疑和停顿则是影响流畅度的关键现象。Mead(2002)发现,口译质量很大程度上取决于译语和源语的非自动化处理过程。因此语言与非语言层面的知识和技能对达成口译过程的流畅度同等重要。除了对译员内在因素影响的探索,也有研究者运用有声思维法探究外在环境条件变化对口译质量的影响。例如,源语语速变化是译员常常遇到的外在环境不可控因素之一,会给译员带来额外的认知压力和工作疲乏,甚至可能会影响译员表现。Pio(2003)运用反省法和实验法探究源语语速快慢与同传译员表现的关系,将“源语-目的语对应”与“译语表达流畅度”分别细化为四项衡量指标来评价译员表现,接着再将其与译员口头报告内容对比。Pio发现,在德-意同传中,当源语达到每分钟145字(即过快语速)时,译员的表现八项评价指标中均受到不同程度的负面影响。

## 3. 有声思维法口译研究的信度和效度考察

对于有声思维法在笔译研究中的适用性已有不少学者从方法论上作出探讨(Toury 1991; Bernardini 2001; Li 2004),从信度与效度的角度来考察目前有声思维口译研究的却不多。笔者在综合考虑实证研究合理合法性的各项标准的基础上,提出针对有声思维法口译研究的“量尺”,并以此来考察和分析当今主要有声思维法口译研究成果实验细节,指出其不足及需改善之处。

### 3.1 确立“量尺”

有声思维法符合实证研究的自然主义维度,旨在通过口头报告的方式收集有声数据、形成主题范畴,通过观察和分析寻找事物发展的模式或规律,属于定性研究范畴。上世纪八十年代,随着社会科学领域里自然主义研究方法的广泛使用,其本身的合理合法性引起争议。Lincoln等人提出的自然主义研究的衡量标准影响尤其巨大,为多个学科广泛采纳。他们认为,一项可信的研究应当具有“真实性、适用性、一致性及中立性”四项标准,前两者属于内在及外在的效度,后两者涉及研究的信度和客观性(Lincoln & Guba 1985:290-293)。

Li(2004)运用上述衡量标准,考察了有声思维笔译研究的信度与效度,并提出了相关的细化评价指标。受其启发,本文针对口译研究的特点,借鉴以上具

体标准并作了修改,提出了用于考察有声思维口译研究文献信度和效度的“量尺”——自然主义研究衡量标准(见表 1):

表 1. 自然主义研究衡量标准<sup>1</sup>

项目	衡量范畴(4项)	衡量标准(4项)	具体评价指标(10项)
1	内在效度 (internal validity)	真实值;可信度 (truth value; credibility)	A. 受试选择 B. 语境选择 C. 选材标准 D. 长期接触 E. 三方验证
2	外在效度 (external validity)	适用性;可转换性 (applicability; transferability)	F. 受试确认 G. 同行征询 H. 详细描述
3	信度 (reliability)	一致性;可靠性 (consistency; dependability)	I. 逐步复制 J. 编码互信 E. 三方验证 F. 受试确认 H. 详细描述
4	客观性 (objectivity)	中立性;可证实性 (neutrality; confirmability)	E. 三方验证 F. 受试确认 H. 详细描述

表 1 有四大衡量范畴,每个范畴有一项衡量标准,这些标准下又有 10 项具体的评价指标,用于从设计细节的角度考察某项实证研究是否达到衡量标准的要求。10 项具体的评价指标的定义如下:

- A. 受试选择:说明选择受试的原因、目的及标准,并交代其个人背景等相关信息;
- B. 语境选择:说明选择语境的原因,以及研究设计是否能够尽量贴近自然真实场景(包括事先对受试进行有声思维法的介绍及培训);
- C. 选材标准:所用语料的选择标准及目的,包括对材料字数、内容、难度等的细节评估;
- D. 长期接触:为了避免研究者在场的负面影响,是否有与受试长期接触,以消除陌生感;
- E. 三方验证:研究中是否对已有数据进行三方验证,如采用辅助的研究方法(如问卷调查、后续个人采访、焦点小组访谈、文本分析、话语分析等);
- F. 受试确认:是否曾与受试确认研究收集到有声数据或其他资料的真实性;

<sup>1</sup> 此表根据 Lincoln & Guba(1985)以及 Li (2004)提出的自然主义研究衡量标准及评价指标修改而得。

- G. 同行征询:是否就研究项目相关问题与业界同行进行过专业征询;
- H. 详细描述:在文献中详细描述从框架设计、实施过程、数据分析到导出结论的整个研究过程;
- I. 逐步复制:设置对照组的受试,重复各项研究步骤,对结果进行相互对比及讨论;
- J. 编码互信:避免研究者预设立场因素干扰,将原始数据的编码工作交由第三方处理。

确立“量尺”标准之后,我们从当前具有代表性的国际翻译学术期刊(包括 *Meta*, *Babel*, *Target*, *Interpreting*, *Across Languages and Cultures*, *The Translator*, *The Interpreters' Newsletter*, *Translator's Journal*)及 Benjamins Translation Library, AICC Bibliography, IRN Bulletin 等数据库资源,筛选出 9 篇采用有声思维法作为主要研究方法且影响力较大的口译研究文献(包括 Bartłomiejczyk 2006; Gumul 2006; Ivanova 2000; Kohn 1996; Mead 2002; Monacelli 2009; Ng & Obana 1991; Pio 2003; Vik-Tuovinen 2002),以表 1 中的 10 项具体评价指标和 4 项衡量标准为量尺,考察这些研究信度和效度。

### 3.2 发现与讨论

从 9 篇文章看来,绝大多数研究者(8 篇)都对其研究程序及步骤进行了详细描述,且非常注重“受试选择”、“语境选择”和“逐步复制”环节。有 4 名研究者详细提及了实验选择口译材料的根据和标准,但有超过一半的研究者并未提及或未考虑口译材料难易度对数据收集和研究结果的影响。这 9 项研究都未提及是否曾经与受试沟通确认所收集数据的准确性,或有无向同行就研究问题进行征询,或请第三方人员进行数据编码处理。仅有一篇文章使用了“问卷调查”的方法进行三方验证,一名研究者谈到自己在约一年的时间内与受试多次接触,并完成实验。分析内容详见图 1:

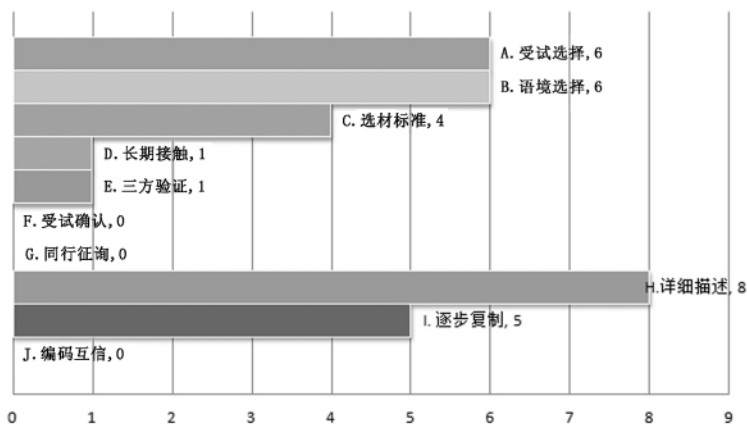


图 1. 10 项具体评价指标在 9 篇论文中的体现

这 9 项研究几乎都谈及了受试(研究对象)的一般信息,但有些研究者并没有详细介绍受试的背景,重要信息有如职业年限、工作或学习经历、语言组合等常常略去。Monacelli (2009)对受试的介绍比较全面。她在其研究中清晰列出受试的编号、具体教育背景及年限(例如,“3 年制口译方向硕士学位”)、从业年限(例如,“自 1990 年起”)及其职业经历(例如,“意大利议会专职译员”、“自由职业者”)等(同上:34)。除此之外,研究者甚少谈及受试的选择标准及目的,而受试的总数量有时候亦低于 10 人以下。例如,Ng & Obana(1991)在其实验中仅采用了 7 名学生(2 名本科三年级、2 名口译一年级、2 名口译 2 年级以及 1 名刚毕业的兼职口译员)。极少研究者提到过是否就受试口译能力展开过前测,以确认同组受试水平相当。这些问题都可能在一定程度上对研究的可信度产生影响。

在语境选择方面,大多数研究者都尽量在实验中还原真实口译场景(如采用真实会议的视频、音频,提前为译员提供背景资料等),以让受试在接受实验时处于自然的情境之下。例如,Mead(2002)在文中谈到,研究者曾考虑使用志愿听众以更接近口译现场,但由于录音地点的限制和观众水平的参差,最终没有使用志愿观众。然而,有声思维法口译研究的场景毕竟与普通口译任务不同。口译研究中常常使用的(即时性)反省法中,译员在口译过程及有声思维过程中的自发性远不如笔译研究中(同时性)内省法。因此,除了还原真实口译情境之外,研究者也有必要对参与研究的受试进行一定的有声思维法培训,帮助他们更加适应在该模式下工作,降低陌生感。

考察还发现,这些研究文献在表 1 的 4 项衡量标准方面仍有待提高(见图 2)。在内在效度的真实值层面,5 项评价指标均在相关研究中有所体现。其中,有 6 篇论文提到或描述了对于语境及受试的选择;4 篇在文中解释了实验所用口译语料的选材标准;分别各有 1 篇谈到与受试长期接触(如 Mead 2002)以及采用三方验证。在反映信度及客观性的一致性中立性层面,均有过半的指标能够在 9 篇论文中找到对应。然而,体现外在效度的适用性层面,9 篇文章仅符合 1 项指标。究其原因,一方面,这是由于口译活动本身的复杂性决定的。口译任务本身可能涉及多个领域(如政治、经济、教育等)、牵涉多种角色(如译员、听众、发言人等),其思维过程又因口译形式的不同而有所区别(如同声传译及交替传译)。因此,一般情况下,这些研究结果较难适用于其他语境或其他受试。另一方面,这种现象也反映了前人的研究设计中欠缺与受试确认有声数据的环节,并常常忽略了与业界同行展开专业征询。

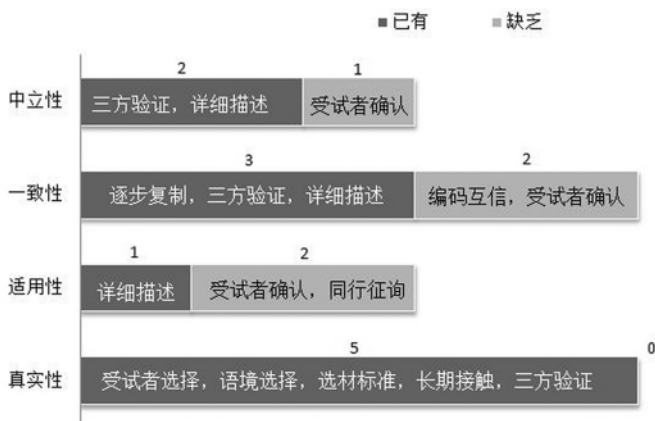


图 2. 4 项衡量标准在 9 篇论文中的体现

#### 4. 有声思维法口译研究应考虑的其他因素

基于表 1 提出的衡量标准考察前人有声思维口译研究实验的信度和效度, 我们认为, 有声思维在口译研究中的应用除要遵循一般的研究程序设计步骤外, 研究者还应考虑以下几个问题:

##### 1) 研究目的

早期的有声思维法口译研究主要以方法论探索为目的, 常常较为空泛。过于笼统的研究目的可能会导致受试的口头报告漫无边际, 造成数据收集的失焦。相比之下, 近期相关文献的研究目的则有明晰化的趋势, 这无疑令研究更有的放矢。因此, 研究者应从更微观的视角来细化研究目的, 寻找可测量的数据单位, 从而让受试在更明确的口头报告方向指引下进行有启发性的有声思维活动。

##### 2) 研究对象

受试的选择对实证研究结论的可靠性起着至关重要的作用。在有声思维法口译研究中, 有必要使用前测、资历认定等手段确认受试的口译能力, 以保证相关数据的可比性, 同时亦应阐明受试的遴选标准及其依据。这一点是在目前的有声思维口译研究中较为缺乏的。在这点上, Gile *et al.* (2001:9-10) 提出了选择受试的两点基本原则: 1) 根据所有可用的信息判断受试是否在所研究的领域具有代表性; 2) 对所得数据结果谨慎推断, 阐明存在的局限。

有声思维法要求受试用口头报告的方法描述、解释或评论自己的思维过程。对绝大多数受试而言, 这并非是表达思想的一种自然的方式。因此, 研究者应提前给予受试适当的有声思维训练和练习时间, 使其适应实验的程序。由于受试有声数据的个性化特点, 研究者在列出受试的个人信息时, 除了详细介绍受试的



语言组合、教育背景及职业经验之外,有时也应根据具体研究项目考虑译者个人性格、抗压能力、抗干扰能力等因素对研究结果的影响。同时,在条件允许的情况下,应尽量增加受试的数量。在一般社会科学研究中,具有统计学意义的受试人数应为 30 人或以上。然而,由于口译活动的特殊性,研究者往往受制于从业人员稀少、地域分散、时间不定以及保密性等原因,难以召集到 30 人以上的译者。探索性的案例研究中,有时仅有一个对象,但却能够引发重要的科学发现。因此,受试人数较少并不会直接影响结论的可信度,口译研究者应尽力保证研究的科学严谨,避免草率下结论(同上:7-12)。

### 3) 口译材料

口译材料虽然是开展口译实证研究的核心载体之一,其实际难易程度的确定与测量常常缺乏客观性。在许多的实验设计和质量评估中,口译研究者或者凭借个人经验决定材料的选择,或者征询专家小组意见定夺材料的难易度与適切性。事实上,即便是专家小组,也因个人对文本难易度的关注点和评价标准不同,难以达成一致(Liu & Chiu 2009)。除了字面意群的难度之外,口译材料的口语性也是重要因素。比如,一篇写起来行云流水的美文,读起来却可能让屈膝咬牙,令译者听得云里雾里,更不知如何翻译是好。简言之,口译材料难易度缺乏一致性可能会影响研究者对译者口译表现的评价,进而干扰到研究结论的可信度。确立客观的口译材料难易度标准,也是确保有声思维口译研究具有可重复性的保证。

### 4) 有声数据的收集与处理

受试进行有声思维时,应确保有声数据得以清晰记录,并尽量避免去干扰或引导。如果此时有任何问题亟待澄清,研究者可以先行记录下来,在后续访谈或问卷中进行解答或验证。数据转写完成后,研究者应与受试确认有声数据转写稿的准确性,以减少转写者在原始数据认知理解上的误差,确保撰写内容的客观性及完整性。在整理有声数据的过程中,研究者除了关注与研究问题直接相关的信息外,也不应忽略有声数据与其他数据(问卷结果、访谈内容、话语分析等)的关联性,寻找研究对象各因素之间的内在关系。

## 5. 结语

本文从文献梳理及研究选样考察中发现,不少有声思维口译研究在信度、效度及整个设计环节都存在着不足,制约了该方法在口译研究中的进一步发展。针对此方面缺陷,本文提出了一系列的原则与措施,以确保有声思维口译研究在方法论上能有充分的理论基础和足够的解释力。

## 参考文献

- Bartłomiejczyk, M. 2006. Strategies of simultaneous interpreting and directionality [J]. *Interpreting* 8: 149-174.
- Bernardini, S. 2001. Think-aloud protocols in translation research achievements, limits, future prospects [J]. *Target* 13: 241-263.
- Choi, J. 2006. Metacognitive evaluation method in consecutive interpretation for novice learners [J]. *Meta* 51: 273-283.
- Dancette, J. 1994. Comprehension in the translation process: An analysis of think-aloud protocols [A]. In C. Dollerup & A. Lindegaard (eds.). *Teaching Translation and Interpreting 2: Insights, Aims, Visions* [C]. Amsterdam: John Benjamins. 113-120.
- Ericsson, K. & H. Simon. 1980. Verbal reports as data [J]. *Psychological Review* 87: 215-379.
- Førch, C. & G. Kasper. 1983. Plans and strategies in foreign language communication [A]. In C. Førch & G. Kasper (eds.). *Strategies in Interlanguage Communication* [C]. London: Longman. 20-60.
- Fraser, J. 1996. Mapping the process of translation [J]. *Meta* 41: 84-96.
- Gile, D. 1992. Basic theoretical components for interpreter and translator training [A]. In C. Dollerup & A. Lindegaard (eds.). *Teaching Translation and Interpreting: Training, Talent and Experience* [C]. Amsterdam: John Benjamins. 185-194.
- Gile, D. 1997. Methodology [A]. In Y. Gambier *et al.* (eds.). *Conference Interpreting: Current Trends in Research* [C]. Amsterdam: John Benjamins. 118-199.
- Gile, D. *et al.* 2001. *Getting Started in Interpreting Research* [M]. Amsterdam: John Benjamins.
- Gumul, E. 2006. Explication in simultaneous interpreting: A strategy or a by-product of language mediation? [J]. *Across Languages and Cultures* 7: 171-190.
- Ivanova, A. 2000. The use of retrospection in research on simultaneous interpreting [A]. In S. Tirkkonen-Condit & R. Jääskeläinen (eds.). *Tapping and Mapping the Processes of Translation and Interpreting* [C]. Amsterdam: John Benjamins. 27-52.
- Kalina, S. 1998. *Strategische Prozesse Beim Dolmetschen. Theoretische Grundlagen, Empirische Fallstudien, Didaktische Konsequenzen* [M]. Tübingen: Gunter Narr.
- Kohn, K. & S. Kalina. 1996. The strategic dimension of interpreting [J]. *Translator's Journal* 41: 11-138.
- Li, D. 2004. Trustworthiness of think-aloud protocols in the study of translation processes [J]. *International Journal of Applied Linguistics* 14: 302-313.
- Li, D. 2011. Think-aloud teaching in translation class: Implications from taps translation

- research [J]. *Perspectives: Studies in Translatology* 19: 109-122.
- Lincoln, Y. & E. Guba. 1985. *Naturalistic Inquiry* [M]. Newbury Park: Sage.
- Liu, M. & Y. Chiu. 2009. Assessing source material difficulty for consecutive interpreting [J]. *Interpreting* 11: 244-266.
- Lörscher, W. 1991. *Translation Performance, Translation Process, and Translation Strategies: A Psycholinguistic Investigation* [M]. Tübingen: Gunter Narr.
- Mead, P. 2002. Exploring hesitation in consecutive interpreting [A]. In G. Garzone & M. Viezzi (eds.), *Interpreting in the 21st Century* [C]. Amsterdam: John Benjamins. 73-83.
- Monacelli, C. 2009. *Self-preservation in Simultaneous Interpreting* [M]. Amsterdam: John Benjamins.
- Ng, B. & Y. Obana. 1991. The use of introspection in the study of problems relating to interpretation from Japanese to English [J]. *Meta* 36: 367-381.
- Pio, S. 2003. The relation between ST delivery rate and quality in simultaneous interpretation [J]. *The Interpreters' Newsletter* 12: 69-100.
- Shlesinger, M. 2000. Interpreting as a cognitive process: How can we know what really happens? [A]. In T. Sonja & R. Jääskeläinen (eds.), *Tapping and Mapping the Processes of Translation and Interpreting: Outlooks on Empirical Research* [C]. Amsterdam: John Benjamins. 3-15.
- Toury, G. 1991. Experimentation in translation studies: Achievements, prospects, and some pitfalls [A]. In S. Tirkkonen-Condit (ed.), *Empirical Research in Translation and Intercultural Studies: Selected Papers of the Transif Seminar, Savonlinna 1988* [C]. Tübingen: Gunter Narr. 45-66.
- Vik-Tuovinen, G. 2002. Retrospection as a method of studying the process of simultaneous interpreting [A]. In G. Garzone & M. Viezzi (eds.), *Interpreting in the 21st Century* [C]. Amsterdam: John Benjamins. 63-71.
- Wakabayashi, J. 2003. Think-alouds as a pedagogical tool [A]. In B. Baer & G. Koby (eds.), *Beyond the Ivory Tower* [C]. Amsterdam: John Benjamins. 61-83.

收稿日期: 2010—05—26; 修改稿, 2011—10—02; 本刊修订, 2011—10—10

通讯地址: 香港九龙红磡 香港理工大学中文及双语学系(李)

510420 广东省广州市 广东外语外贸大学翻译学研究中心(王)

to investigate the impact of language on proficient Chinese-English bilinguals' scene perception. Results showed that (1) properties of language and phrase patterns affected the bilinguals' scene perception; (2) second language proficiency also affected the subjects' scene perception; (3) the effect of language was more direct than cultural influence.

**A study on the variations of the plural marking in Chinese EFL student speech** (p. 864)

*HU Min* (School of Foreign Languages, East China Normal University, Shanghai 200062, China)

Based on data from Chinese EFL-student speech, this article explores variations in plural marking and possible factors contributing to the variations. Ninety-one college students at three proficiency levels participated in the study, which yielded the following findings. First, variants of plural marking were not limited to marked plural nouns and unmarked plural nouns, but also included bare nouns and marked singular nouns. Second, there were systematic variations, of different degrees, owing to phonological, semantic, syntactic, and lexical factors. Semantics was found to produce the most important effect on plural marking while syntax had the weakest correlation with plural marking. Third, factors that conditioned plural marking did not show great differences between speech and writing. Fourth, results provided partial support for the priming hypothesis instead of efficiency arguments and pointed to a close relationship between redundant plural information and plural marking, a relationship which reflects a developmental process of the acquisition of plurals.

**An empirical study of the involvement of working memory components in Chinese-English bilinguals' writing output** (p. 877)

*HUO Jing & WANG Qi* (College of Foreign Languages and Literature, Northwest Normal University, Lanzhou 730070, China)

The study is aimed at testing Kellogg's model (1996) of the involvement of working memory components in writing output processes and exploring the features of phonological loop and visual-spatial sketchpad in Chinese-English bilinguals' writing output processes. A dual-task experiment was conducted, which consisted of examining the reaction times (RTs) of working memory components and the performance of written output. The research findings are: 1) Phonological loop and visual working memory are involved in the "transform" and "plan" processes respectively; 2) The involvement of visual working memory is significant in Chinese writing output; 3) The effect of language switch due to differences between the two languages in which the writing and rhyming tasks were performed is found both in Chinese and English writing, but the effect is significant only in English writing.

**A study of the suppression of story character goals during foreign language narrative comprehension** (p. 887)

*MA Zheng & WU Shiyu* (School of Foreign Languages, Shanghai Jiao Tong University, Shanghai 200240, China)

Based on Gernsbacher's Structure Building Framework, this study explored foreign language (FL) learners' narrative comprehension processes. By using the E-prime software package as a tool to present experimental materials, we conducted two experiments, one being the within-subjects design, and the other the between-subjects design. The results show that 1) the FL learners use the mechanism of suppression to dampen the interference of irrelevant information in order to keep track of the character goals and understand the "causal path" in narratives; 2) their L2 proficiency influences the efficiency of suppression; 3) their L2 proficiency co-effects with suppression, resulting in the "threshold effect" in their suppression of character goals.

**Think-aloud protocols and interpreting research: A critical review** (p. 900)

*LI Dechao* (Dept. of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China)

*WANG Weiwei* (Center for Translation Studies, Guangdong University of Foreign Studies, Guangzhou 510420, China)

The paper reviews the major studies of the past two decades that have applied think-aloud protocols (TAPs) to the study of interpreting. It summarizes TAPs development in the process, strategy and quality of interpreting. Drawing on the safeguards for the trustworthiness of naturalistic inquiries proposed by Lincoln & Guba (1985), the paper specifies a set of criteria which are

then used to evaluate the research designs of nine interpreting studies that adopted TAPs as their research method. It is shown that most of the surveyed studies did not meet these criteria. The paper concludes by giving suggestions on how to ensure the reliability and validity of future TAPs interpreting research.

**Corpus-based translation studies: Issues and progress** (p. 911)

*HUANG Libo* (School of English Studies, Xi'an International Studies University, Xi'an 710128, China)

*WANG Kefei* (National Research Centre for Foreign Language Education, Beijing Foreign Studies University, Beijing 100089, China)

Corpus-based Translation Studies (CTS) has gained marked achievements in recent years, which is manifested in two aspects; one is a deepened understanding of such old topics as translation universals, translator's style, etc.; the other is the development of new topics, such as corpus-based explorations of language changes, construction of the multimodal corpus for interpreting studies, etc. The future development of CTS presents some new tendencies including 1) a shift of focus from translation proper to the external contexts of translation, that is, from translated texts to factors that constrain translation text production and to the changes brought about to the source or target languages; 2) a shift from description to explanation with an increase in empiricism and multi-disciplinarity; 3) the simple mono-lingual comparable, interlingual parallel research model or an integrated one of both has been replaced by the multiple-complex models in relation to research needs with a tendency toward process and causal model.

**Foreign language use in Chinese metropolises: An analysis of evidence from a national survey with special references to Beijing, Shanghai, Tianjin, Guangzhou, Shenzhen, Chongqing and Dalian** (p. 924)

*WEI Rining* (Dept. of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China)

*SU Jinzhi* (Institute of Applied Linguistics, Ministry of Education, Beijing 100010, China)

This paper analyzes the raw self-reported data from the national "Survey of Language Situation in China" concerning residents' foreign language reading proficiency, spoken proficiency and frequency of language use, with special references to seven metropolises. In the three surveyed areas, the national averages were rather low, while the averages of most of the seven cities were statistically significantly higher than the national averages, probably because respondents with relatively high education qualifications were over-represented in the city samples; the effect sizes for the differences between the cities' averages and their national counterparts differed amongst those cities, which might be partially attributed to the factor of education qualification.

**Advances in conceptual metaphor studies in neuroscience: A focus on the issue of psychological reality** (p. 934)

*JIANG Guijing & LI Heng* (College of Foreign Languages and Cultures, Xiamen University, Xiamen 361005, China)

Conceptual metaphor theory put forward by Lakoff via phenomenology has been challenged ever since its advent. One challenge concerns the existence/non-existence of its psychological reality. In earlier times, owing to the limitations in methodology and research perspectives, the arguments on both sides were confined to ontological verifications and hence failed to reach any final conclusion beyond the level of logical reasoning. Findings from recent studies in neuroscience show that conceptual metaphor is based on mental simulation function of the neurons and that psychological reality does hold in behavior, neuropsychology and individual development. In an epistemic stance, the paper gives an overview of the latest explorations into the psychological reality of conceptual metaphor in neuroscience, exposes the existing problems in the previous experimental studies and provides a prospective outlook for further study. Language is embodied in the mind and materialized in the brain. Studying metaphor via neuroscience helps consolidate and further develop the theory.