

# 基于语料库的 同声传译停顿现象成因分析

刘 媛

重庆邮电大学，重庆

**摘 要** | 本研究的语料来源为2018年“第四次工业革命中塑造中国创新社会”国际会议，根据会议现场实录，分析多名同声传译员输出的译文质量，推测了各译员不同的停顿原因。根据本项研究的结果，由于原文速度过快、信息密度大或者是发言者逻辑不清等原因，大多数口译员在听辨过程中认知负荷过重，导致口译员的注意力分配不均，造成某些语句翻译质量较差。

**关键词** | 语料库；同声传译；停顿；认知负荷模型

Copyright © 2023 by author (s) and SciScan Publishing Limited

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/). <https://creativecommons.org/licenses/by-nc/4.0/>



## 1 引言

同声传译是一种复杂的语言活动，译员对源语言的理解和输出几乎是同时发生的，多个待处理任务在同时争夺译员有限的注意力资源（Gile，2009）。语言产生的过程也就是说话人思想的形成、组织和表达，同时说话人自己的言语监测机制也在不断工作，克服或减少语言错误的频率，确保能够输出准确、高效的言语。判断流利与否的主要标志包括非流利特征，如停顿、重复和更正

(Rennert, 2010)。其中停顿是言语交际过程中不可避免的语言现象,因为它象征着许多情况,输出句中的短暂停顿,有时也是断句叙事的一种方法。在同声传译的动态交流过程中,译员必须具备多任务处理的能力。事实上,同声传译人员在口译过程中经常会因为各种原因而出现停顿和输出不流畅的情况,这往往会导致听者无法准确理解输出的译文。停顿是口译输出过程中流畅性的重要信号,也是判断口译质量的重要标准。

本研究的语料来自2018年“第四次工业革命中塑造中国创新社会”国际会议,该会议所有发言的英汉互译由两名经验丰富的专业翻译人员交替进行同声传译工作。本研究以吉尔(Gile)的认知负荷模型为基础,采用定性与定量相结合的方法,对本次会议口译人员出现不流畅停顿的原因进行了简要分析,旨在找出主要影响因素,为解决口译工作中出现的停顿现象提供一些解决思路。

## 2 理论框架

### 2.1 非流利因素分类

蒂西(Tissi, 2000)根据卡尔多涅托、佐迪和科拉(Caldognetto, Zordi, and Corrà, 1982)对语言不流利特征的分类,提出了口译环境中非流利因素的分类,包括沉默停顿和中断。而沉默停顿中又包括初始停顿、衔接停顿和子句内部停顿,而中断(重复、纠正和错误的开始)包括填充停顿、插入性评论和语篇中断。

为了更加突出停顿在沟通中发挥的作用,切科特(Cecot, 2001)对蒂西提出的分类进行了修改,他将沉默停顿进一步细分为了交际性停顿和非交际性停顿。切科特(Cecot)的分类可能更有效地涵盖了在早期研究中出现过的大多数不流利情况。

之后,在高瑟(Gósy, 2007)对语言非流利因素的分类中,指出非流利因素还包括口误、语法错误、单词激活错误和词序问题等口头错误,不论这些错误是否得到补救,他都将这些错误视为非流利现象的特征。

### 2.2 非流利因素成研究

戴朝辉(2011)基于PACCEL-S语料库,研究了顺序口译模式下,性别对

大学生口译非流利性特征的影响。他的研究中发现,在顺序口译中,女学生会比男学生更多地使用自我纠正方法。

林、吕和梁(Lin, Lv, and Liang, 2018)发现,无论口译的工作方向是A语言译入B语言还是B语言译入A语言,大脑的工作记忆能力和语言能力的区别,使得学生同声传译中出现中断和犹豫的频率也出现明显的区别。

普勒夫茨和德弗兰克(Plevoets and Defrancq, 2016)使用欧洲议会共译语料库研究了信息负荷和理解不流利之间的联系。研究中选择了四个指标——语速、词汇密度、数字内容和平均句子长度——分别表征同声传译中理解和表达这两个智力任务的信息负荷,并用语音停顿表征翻译的不流畅性。源语的语速和译文的词汇密度是造成同声传译非流畅性的两个不同因素。

根据现有文献分析,影响翻译语音流畅性的变量可能来自译员、源文本或口译方式(例如,工作方向)。本研究从信息负荷的角度集中分析了说话人语速变化与非流利性特征出现次数之间的关系,也就是主要分析源文本信息密集程度与译员翻译质量之间的关联。

### 3 研究数据和研究方法

#### 3.1 语料处理

本研究中的语料库数据来自2018年“第四次工业革命中塑造中国创新社会”国际会议,会议中包含中国和其他国家几位行业代表对一些前沿科学技术的看法,以及对未来发展的相关政策的解读。本研究把会议视频转录成中英文书面文本,然后创建英汉口译语料库。在语料库中,源语和译文的标点符号(逗号和句点)是基于语义和语音停顿来分组。双向语料库共有23,565个单词,根据语言类别分为中文源语语料库(5,479个单词)、中文译文语料库(9,415个单词)、英语源语语料库(6,118个单词)和英语译文语料库(2,553个单词)。

#### 3.2 停顿定义

停顿通常分为恰当停顿和不恰当停顿,本研究主要针对翻译表达阶段的不恰当停顿进行研究。不恰当的停顿会导致语言输出过程中的不流畅性,因此也

被称为“不流畅停顿”。不流畅的停顿又进一步划分为有声停顿和无声停顿。在本研究中的有声停顿指的是“呃”“嗯”“这个”等在译文中无实际语义的字词，或者是对词语的重复和更正（重新措辞）；而本研究中的无声停顿指句子或字词之间的间隔时间超过 2 秒，出现明显语义中断，即视为无声停顿。

### 3.3 研究方法

吉尔建立了同声传译的认知负荷模型： $SI=L+M+P+C$ ，其中 SI 代表同声传译，L 代表倾听与分析，M 代表短时记忆努力，P 代表言语生产，C 代表协调。同时，在认知资源的功能方面，吉尔认为成功的口译必须满足两个关键要求。首先，译员的 TA（可用处理能力总量）必须大于口译操作过程中所需的 TR（总需求），即  $TA>TR$ 。其次，口译每一步所使用的认知资源总量（LA）也必须大于口译中特定加工步骤所需的加工资源（LR），也就是  $LA>LR$ ；短时记忆能力（MA）必须大于特定口译文本对短时记忆的要求（MR），即  $MA>MR$ ；语言产出能力（PA）也必须能够覆盖口译任务对语言产出的要求（PR），即  $PA>PR$ ；有限精力的协调能力（CA）要大于口译任务中具体对协调能力的要求（CR）， $CA>CR$ 。

吉尔认为认知负荷模型可以很好地解释口译过程中发生失误的主要原因。本研究基于吉尔的认知模型，探讨了同声传译中译员停顿的发生是否与说话人的语速变化密切相关。

## 4 案例分析

### 4.1 数据分析

本研究中作为参考的语速如表 1 所示。

表 1 参考语速

Table 1 Reference speed

英文	英文	中文
VOA 美国之音新闻广播 2.167 ~ 2.667 词 / 秒	英国 BBC 新闻广播 2.5 ~ 3 词 / 秒	CCTV 新闻广播 3 ~ 5 字 / 秒

在本研究的会议语料库中，第 1、2、3 位演讲者的母语是英语，在会议中用英语发言，第 4 号和 5 号发言者是中国人，用中文发言。因此，在本研究中，对发言速度和口译不流畅停顿的横向比较必须限制于同一种语言内，研究的主要结果如表 2 所示。

表 2 研究语料语速

Table 2 Speed of the corpus studied

发言者	1 (英文)	2 (英文)	3 (英文)	4 (中文)	5 (中文)
语速 (字、词 / 秒)	2.580	2.594	3.207	3.972	5.888
非流利停顿 (次)	65	76	51	58	56

由于本次会议有两名译员交替工作，考虑到两名译员在性别、翻译能力、工作经验等方面的差异，横向对比的数据比较局限于研究同一名译员在面对不同语速时，不流畅停顿现象的差异。数据收集的结果如表 3 所示。

表 3 译员停顿频率

Table 3 Frequency of interpreter pauses

发言者	1 (英文)	2 (英文)	3 (英文)	4 (中文)	5 (中文)
语速 (字、词 / 秒)	2.580	2.594	3.207	3.972	5.888
口译员 1 (平均每百词停顿次数)	4.9	5.54	4.56	3.13	2.76
口译员 2 (平均每百词停顿次数)	1.61	1.55	2.34	1.9	1.09

本研究发现语速与非流利停顿的出现频率并不是绝对的正相关关系。研究中发现，相较于参考语速标准来说，所有的英文发言人的语速都是比较快的，其中发言人 3 的平均语速是最快的，但译员 1 的在口译时产生的非流利停顿的频率却相对较低。研究也发现两位译员在同传汉译英时，发言人语速与译文非流利停顿次数呈明显的反比趋势。这些结果明显与普勒夫茨和德弗兰克 (Plevoets and Defrancq, 2016) 的研究成果不一致。也就是说在翻译过程中，说话人的语速与词密度的确会对译文质量产生一定的影响，但并不是决定性因素。笔者针

对上述两种特别现象，对造成非流利停顿的因素做了一个简单的分析。

## 4.2 非流利停顿原因分析

### 4.2.1 等待策略

汉英语序的重大差异增加了汉英同传的难度。英语句子很长时，句子的核心成分不一定会在句首出现。同传时，如果一直等待核心词出现才开始翻译，就容易错过前面的补充信息。所以这时候要采用灵活的处理方法，需要译员在未听到英语中心词时就开口翻译。

如表 4 所示，这一句中并没有难词和专业术语，但是句子很长，还含有一些从句结构。译员 1 在理解句子的同时用了等待策略，等待句子核心意义的出现，导致刚开始出现了较长时间的有声停顿。同时很容易发现，译员的等待策略也导致译文中漏译了很多信息。这也可以说明由于发言人语速过快，词密度大，导致译员用漏译代替了非流利停顿，从而出现了高语速发言中译文停顿频率反而低的情况。

表 4 案例 1

Table 4 Example 1

发言人 3	译员 1	参考译文
Government has a very important role to play here as well, because, retraining is something that they can do right, not by themselves but in partnership with the private sector, right?	我觉得，呃……当然需要政府和私营部门共同来合作。	政府在这里也有一个非常重要的作用，因为培训是政府要做的事情，不是政府自己做，而是与私营部门合作，对吗？

如表 5 所示，译员 2 在翻译时出现了三处短暂的有声停顿，因为原句很长，发言人 5 在说话时有明显的思考和犹豫，导致语句不连续，语义碎片化呈现。所以传译员的精力分配需要兼顾理解原文、整合信息和短期记忆三方面。对于一些冗余信息，译员采取了等待策略，等核心词“the owner of this data”出现之后，才开始组织句式。在所有导致停顿类型中，出现频数最高的就是等待策略这一类。

译员尽管理解了原文，但由于汉语表达能力有限，很难在瞬间产出恰当的表达方式。这意味着译员迅速提取汉语表达的能力还有待提高。

表 5 案例 2  
Table 5 Example 2

发言人 5	译员 2	参考译文
我觉得实际上呃我非常同意刚才的这个观点，就是数据的拥有权，包括我们怎么样，尤其是从公众的角度的话，怎么样更好地去控制数据，然后呢同时又利用到数据对我们带来的巨大的科技优势。	I quite agree with the point. The on the <u>uh</u> the owner of this data need to be <u>uh</u> better defined, and we need to make better use of the <u>uh</u> data.	I think I actually agree with this point, especially from the public's point of view, about data ownership, how we can better control data, and then at the same time take advantage of the great technological advantages that data can bring to us.

4.2.2 语义理解困难

听辨理解困难导致的口译停顿仅次于表达措辞导致的口译停顿。

通过译文可以看出译员 1 在同传时并没有完全理解原文意义，只能选择译出字面意义。在句首出现了短暂的有声停顿，反映出译员根据听到的信息在整理措辞。就在停顿的同时，译员忽略了后面的关键信息，导致理解有误。也就是译员的注意力分配不均，把更多的精力分配给了 M（短时记忆），导致 L（听辨能力）的精力分配不足以支撑译员充分理解原文。听力理解能力是口译成败的一个关键因素，也是一名译员的综合语言 and 知识水平的反映，如表 6 所示。

表 6 案例 3  
Table 6 Example 3

发言人 3	译员 1	参考译文
And the government obviously would foot part of the bill right for some of these expenses in a new type of job with a new type of job description.	然后呢我们也需要去对新的就业岗位进行一个很好的描述。	同时政府显然也需要创造更多的新兴工作和新兴岗位，来解决制造行业人员面临的工作机会丢失的问题。



译员 2 在同传中出现了多次明显的有声停顿, 观察原文可以发现, 原文这一句的逻辑比较混乱, 经过修改后为: “那在宏观层面的话呢, 我现在感到非常乐观的一点是, 我们进入到被大量的低成本的成本的传感器所围绕的时代。我们在之前可能没有这样的一个时代, 我们人产生这么多生理数据的同时, 这些数据都可以被传感器捕捉到”。正因为发言人的语言组织比较混乱, 译员在 L (听辨) 环节消耗了大量的精力, 一方面需要理解发言人的说话内容, 另一方面要整合听到的信息形成短时记忆。所以在 P (产出译文) 环节的精力就不足以保证产出译文的流畅性, 如表 7 所示。

表 7 案例 4

Table 7 Example 4

发言人 4	译员 2	参考译文
那在宏观层面的话呢, 我觉得现在其实我非常乐观的一点是, 我们进入到对大量的低成本的成本的传感器所围绕, 我们在之前可能没有这样的一个时代, 我们人在产生这么多生理数据的同时, 这些数据可以被捕捉到。	And at the macro level, I'm very optimistic. We have entered into an era <u>uh</u> where we have a lot of low cost <u>uh</u> . And we can capture <u>this data</u> this biometric data.	At the macro level, I am very optimistic that we are entering an era where we are surrounded by a large number of low-cost sensors. We may not have had an era before where we had so much physiological data that could be captured by sensors.

通过以上案例我们也可以发现, 除了语速的快慢会影响译文质量, 发言人的说话逻辑是否清晰也直接关系到译员对内容的理解程度, 从而影响译文质量。

在以上的案例中, 不论是语速、词密度、语句长度还是发言人的语言逻辑性都与原文的语言质量 (口音、语速、语误和逻辑性等) 紧密相关。但作为译员没办法改变原文的质量, 所以当原文质量偏低时, 考验的就是译员的听辨能力和语言组织水平。也就是说, 译员需要在短时间内分配好注意力模型, 当原文质量较低时, 需要分配更多精力进行听辨, 相应地减少输出模块的精力投入, 借助笔记等工具保证翻译质量。

5 结论

口译产出过程中的停顿现象与译员的即时认知加工和策略使用高度关联。



译员工作记忆能量的大小决定了思想组织的顺畅与不顺畅,从而导致了产出的流利性和不流利性。译员的流利程度反映了译员处理信息的效率,即解码—编码—再编码—再解码的速度。流利度是同传评价标准的一个重要方面,也决定了社会对同传译员工作质量的好坏评价。因此,同传中的口译停顿的研究格外具有价值。同传训练中听辨理解和口语表达的训练是核心内容。目前的口译(尤其是同传)培训中,大多强调英语的口语水平,而没有对译员的汉语口语水平提出严格的要求。这导致了許多译员母语口头表达能力低下,甚至出现许多低级的用词、语法错误,这应该引起我们高度重视。

在本研究中,分析了这场国际会议的口译译员在双语同传过程中表现出的停顿现象和停顿背后的不同原因。研究表明,译员多数情况下是因为原文的语速过快、信息密度过大、逻辑不严密等,给听辨环节造成了过量的负荷,导致译员的精力分配不均而产出质量不佳的译文。也就是译员的 TA(可用处理能力总量) $<$ TR(总需求),无法解读出原文的完整信息。在未来的口译译员培养方向上,应当更加重视译员的英文的听辨能力培养和中文的逻辑组织能力的培养,任何能力的缺失都会导致最终译文出现问题,努力追求 L(听辨能力),M(短时记忆),P(产出译文)三个方面能力的均衡发展。

## 参考文献

- [1] Cecot M. Pauses in simultaneous interpretation: A con-trastive analysis of professional interpreters' perform-ances [J]. The Interpreters Newsletter, 2001 (11): 63-85.
- [2] Chang C. Directionality in Chinese/English Simultane-ous Interpreting: Impact on Performance and Strategy Use [D]. Austin: The University of Texas at Austin, 2005.
- [3] Gumul E. Explicitation in simultaneous interpreting: A strategy or a by-product of language mediation? [J]. Across Languages and Cultures, 2006 (2): 171-190.
- [4] Hsieh J. Comparing Cognitive Process during Pauses be-tween Experienced

- Interpreters and Novice Interpreters: Eye Movements in Sight Translation [D]. Taipei: Taiwan Normal University, 2014.
- [5] Mead P. Control of pauses by trainee interpreters in their A and B languages [J]. The Interpreters' Newsletter, 2000 (10): 89-102.
- [6] Tissi B. Silent pauses and disfluencies in simultaneous interpretation: A descriptive analysis [J]. The Interpreters' Newsletter, 2000 (10): 103-127.
- [7] 戴朝晖. 中国大学生汉英口译非流利现象研究 [J]. 上海翻译, 2011 (1): 38-43.
- [8] 符荣波. 英汉双向交替传译中译语停顿的对比研究 [J]. 外语教学与研究, 2012 (3): 437-447.
- [9] 谭艳珍, 朱叶秋, 严艳. 汉英交替传译中停顿位置的对比研究 [J]. 外国语文, 2013 (3): 122-127.
- [10] 杨承淑. 口译的信息处理过程研究 [M]. 天津: 南开大学出版社, 2010.
- [11] 杨军. 口语非流利产出研究述评 [J]. 外语教学与研究, 2004 (4): 278-284.

## The Causes of Pause Phenomenon in Simultaneous Interpreting: Corpus-based

Liu Yuan

*Chongqing University of Posts and Telecommunications, Chongqing*

**Abstract:** The World Economic Forum's 2018 Shaping China's Innovative Society in the Fourth Industrial Revolution conference serves as the study's corpus, and the causes for the pauses are examined using the performance of multilingual simultaneous interpreters during the conference. According to the study's findings, the majority of interpreters were unable to interact with one another. The study's findings revealed that most interpreters were overburdened with the listening and discriminating process due to the source text's excessive speed, information density, and logical confusion, resulting in an unequal distribution of interpreters' effort and poor translation quality.

**Key words:** Corpus-based; Pause; Simultaneous interpreting; Effort models