

基于语料库的偏误分析中的模糊性小考¹

金兑垠 朴智英 徐美灵

韩国梨花女子大学 中文系

tekimi@ewha.ac.kr

lychee@ewha.ac.kr

xiaoling518@nate.com

摘要: 本报告就偏误分析(EA)中出现的模糊性问题进行了探讨。我们通常把偏误分析分为五个阶段,即收集学习者的语言资料、鉴别偏误、描写偏误、解释偏误和评价偏误等。其中的解释偏误阶段可以说是第二语言习得中最重要的一个阶段,这是因为这一阶段是找出学习者为什么会产生这些偏误的过程。既这是与第二语言习得的最终目标关系最为紧密的阶段。但研究者在揭示学习者的偏误原因和偏误类型的过程中经常发生一些因研究者的主观判断而产生的模糊性偏误用例。这很有可能导致偏误分析本身就是偏误的问题。因此本报告将提出偏误分析中产生的这些模糊性偏误问题,进行初步的分析,并寻求有效的解决方案。

关键词: 第二语言习得;偏误分析;学习者;错误;模糊性;失误;母语;迁移

A discussion on the ambiguity of Error Analysis based on the corpus

Tae Eun Kim, Jee Young Park, Meiling Xu

Ewha Womans University, Chinese language and literature

Abstract: This study aims on the problem of the errors of EA(Error Analysis). The steps of EA can be classified into 5 different stages: collection of a sample learner language, identification of errors, description of errors, explanation of errors, and evaluation of errors(Ellis 1994:48). Among them, the step of explanation of errors is the most important in SLA, since the answer for the question why learners make the errors are investigated in the step, and that is the most closely related to the ultimate goals of SLA. However, many ambiguous situations occur when researchers find the reasons for the errors of learners, and that makes researchers judge the types of errors very subjectively. As a result, errors of EA can be made as well. This paper discusses about the ambiguous situations in EA with examples of learners' errors. Also, it suggests some remedies to solve the problems.

Key words: SLA, EA, learner, error, ambiguous, mistake, mother tongue, transfer

1 引言

第二语言习得(SLA, Second Language Acquisition), 即真正的外语习得的研究首先是从偏误分析(Error Analysis, 以下简称为EA)开始的。²20世纪50年代是对比分析(Contrastive Analysis, 以下简称为CA)的兴盛时期, 对比分析主要是受到行为主义心理学和结构主义语言学的影响, 但CA强调的是对学习者的母语和目的语的语言体系进行一系列的对比研究。CA认为母语的干扰是学习者出现偏误的唯一理由,因此把偏误看成必须纠正的对象。与此相反,EA认为学习者的偏误原因不仅仅来自于母语的干扰, 从而开始关注学习者自身的语言体系。即开始对学习者的语言体系中既不

¹ This work was supported by the National Research Foundation of Korea Grant Funded by the Korean Government(NRF-2014S1A2A2057389)

² 本文认为对于学习者来说, 如果不是学习母语, 那么在学习其他语言时, 外语(foreign language)和第二语言(second language)这两个概念相差不大, 因此本文没有区分这两个概念。

是来自于他们的母语也不是来自于目的语的偏误进行系统误差(systematic error)分析,来揭示学习者的语言习得过程和习得机制。³

Corder(1971)首次提出了学习者偏误这一问题,还提出了偏误分析的五个阶段,即收集资料,鉴别偏误,描写偏误,解释偏误和评价偏误等。我们认为研究过EA的学者都已认识到这几个阶段的难度相当的大。在这五个阶段中SLA研究的核心是解释偏误,即回答‘为什么会产生这样的偏误’。本文的研究重点就在于揭示发生这种解释偏误的原因。从目前的研究文献来看,与EA相关的诸多研究在解释偏误时许多研究者事先制定出一个基准,然后据此揭示其来源。Richard(1971)提出了干扰误差(interference error),语内错误(intralingual error),发展错误(Developmental error)等三个原因,⁴Richard(1974)又把语内错误和发展错误合二为一,解决了语内错误和发展错误区分不够明确的问题。Dulay and Burt(1974)则分为发展错误,干扰误差(interference error)和独特误差(unique error)。Brown(1980)则把偏误分为干扰传递(interference transfer),语内传递(intralingual transfer),学习语境(context of learning)和交际策略(communication strategies)等。⁵之后就有许多学者在这些先行研究的基础上进行相关研究或在此基础上提出新的分类法。但问题是这种分类法中存在不可能解决和不易避免的模糊性,我们也无法排除研究者的主观能动性的介入。有时出现的偏误不只是一个,而是两三个共同干扰的结果,研究者们为了分类,有可能选择性地分析其中最主要的原因,而排除另一个原因。⁶有时一个误差是制造另一个误差的来源⁷,但实际上EA的分析中并没有考虑到这种偏误发生的过程。这种分析上的模糊性或不彻底性让我们无法保证EA结果的客观性,进而让人很难把握EA的最终目标,即学习者的习得过程和机制。此外还有一个主要的局限性就是如果学习者因不熟悉这些词项而故意使用回避策略的话,我们就很难把握学习者的语言体系。因为EA只能对学习者的部分加以分析,但是克服EA的这种局限性并不是一件容易的事情,因此许多论文明知有这种局限性,还不得不继续进行研究,并对其结果进行分析和讨论。

虽然EA自身存在很多难以克服的局限性,但仍有很多学者以不同语言背景的学习者为对象进行了研究。⁸就如Corder(1967)所说,学习者的偏误对于教授者来说是了解学习者习得的阶段,对于研究者来说是了解学习者的语言习得过程和学习战略的阶段,对于学习者本身来说是了解自己的外语习得程度⁹的最直接的方法。近年来EA研究处理偏误的方法也有了新的进展。即一直以来CA研究者都把偏误看成是必须纠正的对象,如果说初期的EA是通过偏误分析来揭示学习者的语言体系,并进行语言学方面的理论性研究,那么现在的EA研究是利用大量的学习者语料库,来揭示外语学习者习得过程中的难点,最后再应用到教学大纲,教学法以及教材的编写上。总而言之,这些EA在语料库语言学的新的方法论的指导下得到了迅猛的发展。¹⁰

看来今后在SLA领域EA会持续发展下去。本文针对以往西方语言学界指出的EA的基准和

³ Selinker(1972)主张把学习者的语言体系看成一种独立的结构体系,即看成中介语(interlanguage),而不是一个孤立的偏误的集合体。之后许多研究者接受了Selinker的这种观点,把学习者的偏误看成是中介语的某一个阶段,偏误不是一个单纯的纠正对象,而是学习者在外语习得过程中出现的一种语言现象。

⁴ 这一部分的内容参考了Heydari and Bagheri(2012:1583-1585)。

⁵ Brown(1980)又把交际策略分成回避(avoidance),预制模式(prefabricated pattern),认知和人格风格(cognitive and personality style),诉诸权威(appeal to authority),语言转换(language switch)等。

⁶ 参考Long and Sato(1984:257)。

⁷ 参考Nzama(2010:43)。

⁸ 在韩国学界经常可以看到以中国学习者和韩国学习者为对象的EA研究,此外还有López(2009)针对西班牙学习者、Sawalmeh(2013)针对沙特阿拉伯学习者、Jabeen(2015)针对巴基斯坦和伊朗学习者等各种母语背景的学习者为对象的EA研究。

⁹ 参考Corder(1967:161)。

¹⁰ 韩国的中国学界也紧跟这一趋势,进行了很多利用语料库语言学的EA研究,取得了一定成效。如전기정(2012), 박용진(2015), 박지영 외(2016)等。

分类上的局限性,从汉语作为外语教学的角度来进行了详细的讨论。虽然国内外的中国语言学界已对偏误进行了很多具体,有深度的讨论,但还没有学者从偏误分析的局限性出发,并以此为切入点,探讨相应的改善方案。因此我们将对目前的EA相关文献进行回顾与总结,为了今后更好地开展EA研究,深入挖掘解决问题的新思路,好方法打下坚实的基础。

2 对外汉语教学中的偏误分析¹¹

中国的对外汉语教学相较于西方国家起步时间较晚,以鲁健骥(1984)的研究为起点,直到20世纪80年代才基本成型。¹²80年代的EA研究大体上跟CA相同,进行的是母语和目的语之间的初步分析,而中国的偏误分析的成果应该是到了90年代才正式形成。自从鲁健骥(1984)提及中介语的概念和起源以来,很多学者陆续发表了EA论文,相继出现了鲁健骥(1987)、孙德坤(1993)、鲁健骥(1994)、王建勤(1994)、李大忠(1999)等的研究,到了2000年以前已经打下了偏误分析和中介语的理论基础。EA研究涉及的范围不仅仅是语音,词汇,语法等领域,还延伸到语用领域,这一时期的汉语习得研究发展为具有独具特色的领域。近几十年,对外汉语教学迎来了新的鼎盛时期¹³,不仅有了质的发展,还有了量的发展,EA研究也取得了一些可喜的成果。¹⁴

考察2000年以后中国的偏误分析论文可知研究的主题和对象呈细化,多样化面貌,数量上也有了令人瞩目的飞速成长。在中国文学以及汉语教育方面,最权威的中国期刊《语言教学与研究》《世界汉语教学》《汉语学习》《语言文字应用》中2000年以后发表的EA论文数量竟达232篇,这是20世纪80年代到现在为止发表的EA论文总数的86.9%。¹⁵论文数量快速增长的最重要的原因在于国内外的汉语教育需求在持续增长,因此EA的研究也随之不断增长。可以说2000年以后是中国EA研究真正的发展时期。但是就如前所述EA研究须分收集材料,鉴别偏误,描写偏误,解释偏误,评价偏误等阶段进行,中国EA研究的大多数为解释偏误,既比起阐明偏误来源,更侧重于解释。解释偏误时大都没有客观性的判断基准或统计,只凭研究者的语感以及主观上的判断,也没有明确注明偏误句的出处,分析上也有很多有待补充的部分。¹⁶

肖奚强(2000)考察了韩国汉语学习者的偏误问题,但没有把偏误问题进行类型化分类,也没有分析偏误来源,只凭自身的教学经验,罗列了韩国学习者的典型的偏误句,并提出几项针对韩国学习者的教学法,这一研究只能属于比较初级的EA研究。肖奚强(2001)为了深入分析EA,采

¹¹ “对外汉语教学”这一概念是中国学界使用的一个术语,指的是汉语作为一门外语的教学和学习。韩国也沿用了这一术语,因此本文也采用了这一概念。

¹² 鲁健骥发表了一系列论文,如《中介语理论与外国人学习汉语的语音偏误分析》(1984),《外国人学汉语的语用失误》(1993),《外国人学习汉语的词汇偏误分析》(1994)等,这些论文可说是中国EA的先行研究。

¹³ 参考김혜령(2015:32-38)。

¹⁴ 考察《语言教学与研究》《世界汉语教学》《汉语学习》《语言文字应用》等与汉语及教育相关的代表性期刊结果表明,1980年至2016年发表的偏误分析论文数量达267篇。以下《世界汉语教学》是从1987年开始统计的数据,《语言文字应用》是1992年开始统计的数据。

期刊	《语言教学与研究》	《世界汉语教学》	《汉语学习》	《语言文字应用》
论文数量	59篇	60篇	87篇	61篇

¹⁵ 2000年以后发表的EA论文篇数如下,跟1980-1999年的论文篇数相比,增加了六倍。

期刊	《语言教学与研究》	《世界汉语教学》	《汉语学习》	《语言文字应用》
论文数量	47篇	56篇	75篇	54篇

¹⁶ 在中国EA领域中引用指数高的研究者,如肖奚强(2000)、赵葵欣(2000)、赵金铭(2002)、袁毓林(2002)、周文华(2009)、谢福(2010)等也没有具体探讨偏误来源,大都注重偏误解释。

用了崭新的 EA 研究方法,主张区分偏误和失误,根据偏误分析的目的进行类型化分类之后,再对偏误用例进行分析。肖奚强(2005)在学习者语料库中抽取‘除了’句式,并对此中介语进行了详细的研究,不仅描写了外国学习者使用‘除了’句式的情况,还按语言水平等级统计了使用分布情况。这一系列的研究虽然在试图采用多种方法来研究 EA 这一点上是值得肯定的,但仍然处于片面的解释偏误的阶段。即还是没有到通过分析学习者的偏误来源来解释第二语言习得过程的阶段。

方绪军(2001)倒是提到了产生偏误的诱因。这篇论文研究了中介语中带有偏误的动词句,即动词和补语的偏误问题,指出动词句产生配价偏误的诱因是学生母语的干扰、教学引导的不足、不适当的类化、学生个人因素等。但他在解释产生偏误的诱因时只是举出简单的偏误用例,因此实际上他的这些研究很难说是解释偏误。

2000 年以后的中国 EA 研究在偏误分析方法上仍然没有摆脱鲁健骥(1994)提出的遗漏、误加、误代、错序等四个模式。赵金铭(2002)根据学习者的偏误与母语话者可能出现的句子相似度来解释偏误。假如错句和正确句最为相似则说有“最小差异”,并根据可接受性分了 5 个等级。¹⁷他还提出发生这些错句的原因很可能是目的语学习的不足和母语干扰,只可惜没有进一步对其进行更为详细的分析,而且偏误用例大都来自转引材料。崔希亮(2005)则研究了欧美学生汉语介词习得的偏误。他说从偏误类型上看共有九种情况:介词冗余,框式介词缺少呼应词语,介词结构出现的位置不当,结构错位,结构不完整,体貌标记错误,词语搭配问题,语义模糊,介词混用等。实际上这些偏误都是从添加,遗漏,错序,混用等几项大类里细分出来的,但是在尝试提出具体的类型并进行细致的分析这一点上是值得肯定的。虽然他的研究比起之前的研究在解释偏误上有了一定的进步,即对各种偏误进行了更为详细的说明,但令人遗憾的是没有对发生偏误的原因进行深入考察,也没有对母语和目的语进行比较分析,更没有提出适用于各个学习者的教学方案。

袁毓林(2005)则用最近在 EA 中使用最多的语料库语言学来进行研究。这个方法是在语料库中抽取大量的偏误用例进行统计,从而确保研究的客观性,从这一点上看,他的研究有一定的意义。他把学习者在使用“不”时出现的偏误分为句法和语义两个类别进行议论,但这一研究同样也没有详细分析偏误原因,只注重于解释偏误类型。谢福(2010)在对较大规模的学习者语料进行统计的基础上,区分了“是……的”句的正确句和偏误句,还研究了初、中、高三等级的留学生“是……的”句的习得顺序。他将“是……的”句分为八个句式,进而对学习者的使用频度加以比较。这一研究的意义在于可让我们了解学习者和母语者的“是……的”句习得情况,但他没有尝试解释为什么会出现这种状况,出现偏误的主要原因是什么。周文华(2014)尝试使用类型学的分析方法,在大量的语料库里抽取偏误用例,考察了 SVO 型和 SOV 型母语者的介词及介词句的语序上的差异。他得出的结论是母语的语序带给汉语语序的影响跟学习者的水平有关。即随着学习者水平的提高,母语的干扰将逐步减少。比起以往的研究只把偏误的原因归为母语的干扰有了一些进展,但是最终得出的结论仍然是不必对母语迁移过度泛化,没有提及母语以外的其他原因引起的学习者的偏误。

如上所述,中国的 EA 研究到现在为止,在量化和质化上都有了飞速的增长,研究范围或对象也变得更为集中细致,在使用学习者语料库时采用的数学统计等研究方法也比之前客观化。但是研究的焦点仍然停留在具体地描述偏误的阶段,还没有到揭示偏误原因改善并运用到教学上的阶段。初步揭示偏误原因的几篇论文大多数仍为母语和目的语的差异对比。

3. 通过偏误用例来看偏误分析的模糊性

这一部分将以韩国大学生学习者的作文语料库为基本材料,讨论在 EA 中解释偏误时出现的模

¹⁷ 引用了袁博平(2001)在《汉语中的两种不及物动词与汉语第二语言习得》(演讲稿)一文中提到的“等级”概念。

糊性。这并不是假设性的推论,也不是回顾教学中碰到的偏误,而是讲学习者在作文课中出现的偏误,因此这是个有用的信息。¹⁸而且因为这是以韩国学习者的作文为对象的,还可为国内学界进行的偏误分析研究提供可靠的(authentic)资料。

如前所述,EA中说的解释偏误就是揭示偏误原因,因此可根据出现的偏误原因进行分类。据被许多EA研究者奉为指针的Ellis(1994)的说法,偏误分析是SLA中揭示其本质的阶段,是EA中最最重要的一个阶段。¹⁹因此解释偏误时,研究者们应该非常慎重,必须确定偏误的类型。²⁰西方的许多EA研究者大都把学习者的偏误放在语际错误(interlingual error)和语内错误(intralingual error)这两个大的范畴里进行分析。后来随着EA研究的发展这两个类型分得更为细致。周小兵、朱其智、邓小宁(2007)是最具代表性的中国EA研究者之一,他们把偏误发生的原因分成母语的负迁移、目的语规则泛化、教育误导、交际策略运用、偏误的认知解释等五个类型。²¹最近朴智英等(2016)对韩国汉语学习者的动宾结合关系进行了分析,他们根据偏误原因分为自然句、简单偏误句、复合偏误句、无法推测的偏误句等四个类型。²²

尽管研究者们为了合理地解释EA而做出了诸多的努力,但是解释偏误时最先出现的问题仍然是偏误句里不是只有一个偏误,而是会出现两三个复合型的偏误,因此我们面临的问题是如何处理这些偏误用例。或许研究者们只能选择最具代表性的偏误。那么我们又面临谁来定这个所谓的代表性偏误呢,又有谁来保证这个判断结果具有客观性呢等等问题。或者也有人可能为了消除这样的问题把所有的复合偏误放到各个偏误类型里。不过即使使用这种权宜之计处理复合性偏误,还是有一个更大的问题,那就是怎么解释这个偏误,即在判断偏误原因时,会有研究者的主观性干扰。这就是本文将要重点讨论的话题,即解释EA时其本身带有的偏误和模糊性等诸多问题,本文旨在把这几个问题进行分类并加以分析。

3.1 分析材料的来源

本论文分析的学习者作文为ColloConc²³资料。ColloConc是可以输入韩国汉语学习者的作文,构建学习者语料库的一个系统,还可以输入教材上的材料,因此可以有效地分析学习者使用搭配的情况与教材的相关性,是一种concordance。²⁴输入时期为2015年9月到2016年6月,学习者

¹⁸ 有关偏误分析的研究中有不少偏误的出处是不够明确的。这些偏误用例有可能是研究者根据自己的记忆而写的,也有不少可能是研究者为了便于研究而进行处理的(manipulate)偏误句。这样的偏误用例也不是没有研究的必要。在理想的分析中设定可预料的偏误并对此进行分析也有一定的意义。但是用学习者语料库来研究过EA的研究者都会觉得,事实上很难找到这种理想的偏误句。

¹⁹ 参考Ellis(1994:57)。

²⁰ 最近的韩国汉语EA研究大都处于鉴别偏误、描述偏误阶段。(진기정(2012)、윤유정(2013)、손경애(2014)、박지영 외(2016))这些研究在揭示学习者偏误形式这一方面具有一定的价值,但是没能揭示EA的核心问题,即解释偏误原因,第二语言习得过程本身,从而最终为学习者服务,因此从这一角度看还是有所不足。

²¹ 母语的负迁移指的是学习者把母语的语法规则直接沿用到目的语中,从而引发的偏误,这就是我们通常所说的语际错误。目的语规则泛化指的是没有完全理解目的语的特定语法现象和相关规则,并错误地以此类推,把范围扩大到不适合的范畴。教育误导指的是教学误导引发的偏误。交际策略运用指的是学习者在目的语知识不足的情况下使用交际策略时出现的偏误。偏误的认知解释指的是语言的普遍性、标记性、自然度等。유재원 외(2014:84-93)再引用。

²² 自然句指的是正确的句子,简单偏误句指的是有错别字的句子。复合偏误句可分为遗漏、添加、混用、语序、其他等五项。遗漏指的是词语的一部分或全部被省略或脱落的偏误。添加指的是插入不必要的语素或词语的偏误。混用指的是对词语的用法和词汇搭配关系理解不够而出现的便误。语序指的是在句子成分的排列上出现的偏误。此外不属于以上四种情况的偏误处理为其他偏误。最后无法推测的偏误句指的是研究者无法从中推测并分析出规律的偏误用例。박지영 외(2016:126)再引用。

²³ <http://colloconc.ewha.ac.kr/>

²⁴ 有关ColloConc的详细资料与功能请参考朴智英外(2015)。

作文中出现的词汇类符(type)的总数为 3,168 个, 形符(token)总数为 72,262 个。²⁵

收集到的学习者作文句子总数为 11,616 个。有句号、问号、叹号等文章符号则看成一个句子, 不包括有逗号的句子。学生的作文按学习者的等级进行分类。等级分为初中级(low intermediate)阶段²⁶及中级(intermediate)阶段, 收集初中级阶段的语料时再按学生使用的汉语教材细分, 这是为了便于观察教材的内容和学习者作文的相关性。初中级和中级的作文主题都定为是‘我的现在和未来’, 比较宽泛, 为了收集更多的文章, 让负责该课程的老师提示学习者即使跟主题的相关性不大也可尽量多写。之所以选择这种开放的作文形式, 是为了让学习者自由选择词汇, 尽量避免因教授者的课程设计而引发的偏误可能性。

本文中涉及的内容为偏误分析时呈现复杂性面貌的句子以及带有模糊性的句子, 因此本文的分析对象排除了使用比较简单的句子的初中级阶段的作文, 只选取了中级阶段的作文。本文搜集的作文对象是 2016 年第一学期(即 3 月)在梨花女子大学上《汉语阅读》和《汉语口语I》(二年级专业课, 原语讲座)的 50 名学生。²⁷我们在这些学生作文中选取的用例是汉语学习期间为一年以上的学习者的作文。我们认为在EA的解释偏误阶段, 一般情况下学习者的水平越高, 出现的问题越复杂, 因此我们没有选择这一阶段的学习者。我们认为学一年以上的汉语, 才可定为中级水平, 而处于这一阶段的学习者中介语较有规律性, 其原因也多种多样。

3.2 失误和偏误

Corder早就指出应区分偏误和失误(mistake), EA研究须排除失误用例。这里所说的学习者的偏误指的是由于学习者对目的语的知识不足而反复出现, 学习者又没有意识到那是偏误, 也不能

²⁵ 类符指的是使用的个别词汇的数量, 形符指的是使用某一个别词汇的总数。例如, 学习者作文中出现的动词“去”的类符虽为一个, 而收集的作文中出现“去”的形符为 2,517 个, 即学习者作文中使用“去”的频率。



图 1 “去”的使用频度 (类符频率)



²⁶ 收集的初中级学习者的作文是 2015 年 9 月到 12 月之间, 为期一学期。参与作文的学习者为位于首尔市内的 4 所大学听<중국어 2>的学生, 大部分是学习<중국어 1>时才接触汉语的学生, 学习期间为 6 个月到 8 个月。开学后严格限定了听课学习者的汉语学习经验, 排除已有丰富学习经验的学生, 因此可以说学习者之间的汉语水平偏差不是太大。初中级作文的收集一共为 3 次, 每次平均收集了 375 名学习者的作文。中级学习者的作文是在 2016 年 3 月到 6 月在梨花女子大学收集的。中级汉语一个学期开设的课程数较少, 听课的学生也相对地少一些, 因此比起初级汉语学习者的作文要少很多, 有一定的局限性。

²⁷ 虽然我们可根据课程的性质来推断选择这门课的学习者大概是具有中级汉语水平的学生, 但笔者还是进行问卷调查了解了学习者的个别学习期间。其结果发现六个月至一年的学习者六名, 因此我们在研究中排除了他们的作文, 在 44 名学生的作文中抽取了偏误句。

自觉地改正过来的情况。而失误指的是偶然出现,学习者可以自觉地找出错误,并立即改正的情况。偏误和失误的定义非常明确,人们对此也没有什么异议。实际上在EA中识别出这两种情况并不是件容易的事。²⁸以下是学习者语料库中出现的偏误句中写错字的例子,研究者可以非常容易且明确地指出此类句子是错句。²⁹

(1)* 我们发现其中 300 个人,只 11 个个吃这个药。(我们发现 300 人中,只有 11 个人吃这个药。)

看例(1)“11 个个”中的第二个“个”字,我们可以断定这是把“人”字错写成“个”字的情况,为什么呢?因为前一个分句中“300 个人”中的“人”字并没有写错。³⁰因此此类句子我们不能称其为偏误,原则上不应该包括在EA范围里。但是下面的例(2)就不能等同于例(1),一概而论。

(2)*我只在学校上○(写错的“课”字),但我不太喜欢我的专业,所以我的成绩也不好。

(我只在学校上课,但我不太喜欢我的专业,所以成绩也不好。)

(3)*我最近的生活过得很忘。(我最近的生活过得很忙。)

看例(2),如果我们根据这个学生写的其他句子的水平来判断,不像是因为学习者不会写“上课”而写错的。即这好像是错字,是一种失误。但是光靠分析学习者的作文,我们没有方法证明这是失误。因为这个学生只写了一次“上课”,而这仅有一次的“课”是写错的。因此我们不能明确地说是刚开始学时没有学会写“课”字,还是在写作的过程中出现了这种失误。如果是前者的话属于偏误,如果是后者的话属于失误,但我们没有明确指名其类型的根据。下面例(3)也属于这种情况,就如例(2)一样,不知道是把“忘”字错写成“忙”字,还是因失误写成“忘”字,我们不能光凭学生作文的分析来妄下定论。跟例(2)不同的是,我们不知道是把“忘”字写错,还是把另一种意思的字写在了“忙”的位置上,也有可能是因“忘”和“忙”的汉字构成要素(亡+心,亻+亡)相似而产生的,因此根本无法判断是偏误还是失误。下面是好像事先学过汉字的学生的作文中常常出现的简体和繁体混用的句子。

(4)*现在我是大學三年级的学生。(现在我是大三的学生。)

在这个句子中,学生把“大学”的“学”写成了繁体字,“学生”的“学”写成了简体字。即在一个句子中同时混用“學(学)”的繁体和简体字。那我们现在来设想一下所有可能发生的情况:第一,这个学生是因为事先学过汉字,因此受其影响把“大学”的“学”写成了繁体字。第二,这个学生有可能认为“大学”的“学”是繁体字,“学生”的“学”是简体字。第三,这个学生混淆繁体的“學”和简体的“学”,并不知哪一个是正确的。如果是第一种情况应该看成是失误,如果是第二和第三种情况应该看成是偏误。因此可以说研究者在解释偏误时,不仅很难鉴别失误和偏误,而且也很难解释偏误原因。

在EA中鉴别偏误和失误是识别偏误的最为基础的阶段。但是至今为止的研究中连这种最为基础的阶段也很难下结论。如果所有的鉴别和分析需要靠研究者的主观性的话,那门我们只能把它看成是为了分析而做的分析,而不能看成是真正意义上的SLA研究。

3.3 语际错误和语内错误

现在开始讨论一下在鉴别偏误的过程中不属于失误,而归为偏误,描述时很难找出其原因的情况。如前所述, Ellis(1994)把偏误分为语际错误和语内错误两类,前者是因母语的干扰导致的错误,后者是学习目的语时遇到的内部干扰而发生的错误。如果把这两个概念放到韩国学习者的情况来看,语际错误是因韩国语的干扰而发生的错误,语内错误是因汉语本身具有的难点而产生的

²⁸ 例如 박지영 外(2016:120)没有区分偏误和失误。因为研究对象为大规模语料库,使用汉语搭配分析系统(ColloConc)抽取句子,因此即使是失误也没有确认的方法。

²⁹ 例文使用简体字。因为学习者的偏误中常常出现简体和繁体混用的情况,因此我们需要对此加以区分。而且纠正偏误句的时候,如果是正确句,就放在括号里。

³⁰ 实际上这个学生的作文中除了本文中提示的句子以外,还出现了很多次写对的“人”字。

错误。这两个类型的偏误，因发生偏误原因各有不同，应该明确加以区分，实际上经常因此出现解释偏误的模糊性。下面举个例子详细地说明一下。

(5) *我不用心什么成绩。(我不用在成绩上费什么心。)

一看例(5)，我们就可知道这是因词汇的混用而发生的偏误。在汉语语法项目中离合词是韩国汉语学习者的难点之一。造例(5)这个句子的学生就是因为不知道“费心”是动宾离合词，因此在“费心”后面使用了“什么成绩”做宾语，造了这样一个错句。如果把例(5)的错句放到韩国语里也非常不自然，因此也不会怀疑是母语的迁移。下面例(6)和例(7)的例子跟上面的例子又有所不同。

(6) *我家人也同感我的意见。(我家人也同意我的看法。)

看例(6)，学生造的这个句子韩国语的意思大概是“우리 가족도 나의 의견에 동감한다”。看来是联想到自己母语中与这个词语相对应的词，即韩国语里的“동감하다”一词，进而把二者完全等同起来造成的错句。那么这种错误可看成是母语迁移，既究其原因应该是语言之间的干扰引起的偏误。反之，如果学习者是因不知道汉语的“同感”不能用作动词，只能用作名词而发生的偏误，那么对此进行的原因分析也应有所不同。因为汉语里有很多兼词，学习者很容易把“同感”看成是可带宾语的动词。在没有对学习者的跟踪采访，问为什么会把“同感”用作动词的情况下，我们不能百分之百地肯定这究竟是因何原因而发生的偏误。例(7)在解释偏误时同样存在模糊性。

(7) *还有我对化妆品和衣服有关心。(还有我很关心化妆品和衣服。)

例(7)是韩国学习者的作文中常常出现的偏误类型，究其原因可知是误用“关心”而发生的偏误。首先把偏误的原因看成是母语迁移，看成韩国语的“게다가 나는 화장품과 옷에 관심이 있다”，因而产生了像例(7)这样的偏误句。那么这就属于语际错误。反之，如果是学习者没有正确掌握“关心”这一词语的词类，把“关心”看成是名词，而不是看成动词，从而造了这样一个错句，那么这就属于语内错误。例(8)的情况更为复杂。

(8) *我常常生活得很忙，做着各种各样的活动。(我过得很忙，参加各种各样的活动。)

看例(8)，学习者想说的意思大概是他过得很忙，参加了很多活动。如果把这一句子逐一翻译并造句，“생활하다”是“生活”，“활동을 하고 있다”是“做着~活动”，可看成是因母语迁移而发生的错误。但是如果再仔细分析的话，可看出学习者显然不知道“지내다”用汉语说应该是动词“过”，“生活”后写了用作程度补语时使用的助词“得”，看来是不知道“生活”是名词。而且显然也不知道汉语的“活动”是名词，跟动词“参加”搭配。很明显这个学生还没有正确掌握，表示汉语的动作的进行或持续的时候应该使用助词“着”。就是说例(8)是涉及汉语的好几个语法项目的复合型错误。因此如果从这一角度分析，这个句子的错误应该属于语内错误。像这样只要翻译成韩国语时是通顺的句子，就无条件地把学习者的错误原因看成是母语的迁移，那么有可能解释偏误本身就是一种错误。

3.4 不自然的表达

本小节要讨论的是怎样处理句子本身在语法上没有问题，因此可以说是正确句，但是从母语话者的角度来看又有些不自然的句子。如果这是在教室里，学习者问老师自己造的句子是否正确时，教授者有可能说类似于“中国人不这么说”之类的模棱两可的话。因为这一问题不是语法上的问题，而是社会语言学中所说的可接受性(acceptability)和适当性(appropriateness)问题。EA 研究者也会碰到这种因可接受性和适当性问题而发生的显得不够自然的句子，母语话者都觉得某些地方不够自然的句子。首先我们会面临是否该把这个句子看成是错句的问题，如果看成是错句，那么哪个部分是错误的，应该如何描述，究其原因是什么，应如何解释等等问题接踵而来。我们先来看一看识别偏误阶段的一个用例。

(9) *我的最好朋友是我的姐姐。(我最好的朋友就是我的姐姐。)

我们可以肯定例(8)是个错句。因为助词“的”的位置弄错了。即使改成“我最好的朋友”，

如果没有“就”字,学生造的句子还是不够通顺。虽然不能说这是语法上的问题,但只要是使用汉语的母语话者肯定会说该使用“就”字。那么这个句子应该处理成一个错误还是两个错误呢?

(10)?虽然我不知道,也不确定未来的我会做什么。(虽然我不知道,也不确定我未来会做什么。)

实际上例(10)是正确句,因此我们不能把它放到错句里来分析。但是母语话者认为像例(10)这样的句子不够自然,应该说成括号里的句子。那么我们应该把这个句子看成错句呢,还是应该看成正确句呢?如果看成是前者,韩国汉语学习者有可能把韩语的“미래의 내가 무엇을 할 것인가”翻译成“未来的我会做什么”,这就成了因母语迁移而发生的偏误。如果是后者研究者仍然无法排除句子不够通顺的感觉。

(11)?一半是中国人,一半是韩国人。(所以我觉得自己一半是中国人,一半是韩国人。)

例(11)里想表达的意思是从小住在中国,因此觉得自己一半是中国人,另一半是韩国人。这个学生造的句子光看两个分句意思很明确,但例(11)没能准确地表达学生想说的真正的意思。事实上例(11)从语法的角度看是正确句,如果考虑上下文应该改成括号里的句子,那么我们会遇到这个句子是正确句还是错句的问题。如果是错句,那么其原因是什么?应该归为哪一个类型的偏误呢?研究者该如何处理这一类型的偏误呢?例(12)跟例(11)一样,有相同的问题。学生造的句子从语法角度看,没有任何问题,但是母语话者会觉得括号里的句子更为自然。

(12)?我在考虑去公司工作,还是考研。(我在考虑是去公司上班,还是去考研。)

如何把自然度反映到偏误分析里,是留给EA研究者的一个难题。请看下面的句子,模糊性更强。

(13)?我现在比较忙,期末考试快要到了。(期末考试快要到了,我现在比较忙。)

例(13)的句子翻译成韩语就是“나 지금 좀 바빠. 곧 기말고사야.”,从逻辑角度上看似乎没有什么问题是在母语话者看来,这个语序是相当别扭的。因为从中国人的逻辑角度上看应该先说“期末考试快要到了”,然后再说“我现在比较忙”,才会觉得自然。那么这个用例也要看成是因母语迁移而发生的句子排序上的错误吗?最近有关EA的研究大部分是分析大规模的学习者语料库,把偏误进行分类,再各自进行统计,那么我们能否把这种句子排序上的问题放到一系列的类型中呢,这也是我们会遇到的问题。而且在此之前这类句子本身就没有问题,能否成为EA的对象也是一个需要考虑的问题。反之,如果看成是正确句,那么只能无视交流式教学法中追求的符合上下文的话语,只是考虑单个句子的EA研究。

3.5 学习者个性差异上的问题

为了正确的EA分析,大部分的研究者们一般搜集同一水平学习者的作文或语音材料。尽管如此事实上我们很难考虑到成人学习者的先行学习水平和学习者本身的思考方式等。因此在这种情况下搜集的学习者语料库中不可避免地出现学习者个性上的差异。根据最近的EA研究经验来看,韩国汉语学习者中最突出的现象就是先行学习的英语对汉语习得的影响。³¹

(14)*所以就打消了当一名翻译的念头。(所以就打消了当翻译的念头。)

如例(14)所示,学习者经常在一般名词前面加“一+量词”。其实这个句子不需要说明是一个人,因此加数词“一”就显得很不自然。韩语也有这样的情况,因此我们不能把这种偏误的原因看成母语的迁移。故意加不必要的数词“一+量词”,这种偏误可能不是由母语引起的,我们只能从另外一个角度找出其原因。如果学习者为大学生或成人,那么我们推测这种滥用“一”的现象可能是受先行学习的英语的影响。即学习英语时学到在一般名词前加不定冠词“a”这一语法的学生很可能在学习另一门外语——汉语时把其规则过度泛化。但是我们也无法单凭搜集到的作文就轻率地做出受英语影响的结论。写出这种句子的学习者是否都学过英语,即使学过英语,我们也无法确

³¹박지영외(2016:130)中提出介词句出现在句尾的情况好像是受先行学习的英语的影响,而在分析时却把这一现象归类于复合型偏误中语序偏误。

认是否是受了英语不定冠词“a”的影响。

(15) *而且我打算努力打工为的是挣很多钱。(而且为了挣更多的钱,我打算努力打工。)

例(15)的偏误原因也很难分析。学习者大概是想说“돈을 많이 벌기 위해서 열심히 아르바이트를 할 계획”之类的话,这一点并不难猜测。但是为什么把“为的是挣很多钱”这一形式放在句尾?其偏误的原因到底是什么?我们很难下结论。这跟韩国语的语序也完全不一样,因此我们也不能把这个句子看成是母语迁移,虽然可以简单地说可能是没有正确掌握“为了~”句式的用法,但这么说又无法说服自己。因此在这里我们可以多加一个理由,那就是英语的影响。即需要解释原因的语句可能是受和to(to earn lots of money)或for(for lots of money)一起处于句尾的英语句型的影响。此外,虽然可能性比较小,但也有可能是因为学习者的逻辑思考本身较为特殊。因为有如此多的情况,因此我们不能在此断言其中的哪一个才是正确的原因。

(16) *毕业以后的时候我有想学什么是别的我的专业。(快要毕业的时候,我又想学一个别的专业。)

例(16)很难看出其意图和句型。学习者原来的想法很可能是笔者放在括号里的句子,韩国语的意思最有可能是“졸업한 후에 나는 또 다른 전공을 하나 더 공부하고 싶다”。但即使想表达的意思是这样的,但是写出“我有想学什么是……”这样一个句型非常独特。这时如果我们想一想英语的“what I want to study more is...”这个句型可能有所启发。可见我们不能完全排除英语的影响。或许这种假设本身就是研究者出于主观性的过度分析,倒不如简单地说这也是学习者个人逻辑上的问题。

以上第三部分探讨的问题都是在研究EA时可能因存在好几种复合原因而产生的偏误情况。如果研究者任意选择某一个最具可能性的一个原因,那么其结论可能是客观性的偏误,也有可能是研究者主观性的偏误。而且这对统计数字的可信度也有直接的影响。因此直到最近,尽管有很多有价值的EA研究,但我们也有必要重新考虑一下是否应对偏误本身多加重视,并进行分析与讨论。

4. 结语

EA的最大目的是集中研究学习者的语言体系,理解第二语言的习得过程,为学习者提供帮助。除了本文第三章提到的情况以外,如果学习者因为是难点,而有意回避特定表达方式或语法项目的话,研究者则无法对此类词汇和语法进行偏误分析。事实上最近大部分EA研究者在利用大规模的语料库语言学上的研究方法来进行研究,而这种大规模的学习者语料库的统计数字确实提高了客观性和可信度,却忽视了在搜集过程中如何找到真正能揭示其原因的具体线索。因此我们也无法识别在语法上没有错误,但在上下文衔接上不通顺或不自然的错句。这样一来我们有可能造出跟学习者的意图完全不一样的句子,因为只要语法上没有问题,那我们就无法分析出其真正的偏误原因。例如学习者想表达的是“张三把李四推倒了。”,但是说成“李四把张三推倒了。”,在不能确认上下文的情况下很难识别偏误。EA的先行者Corder或Selinker提出的系统的偏误³²指的并不是因人而异的个别偏误,而是共同出现的偏误,但这不是表面上看起来相似的情况。本文旨在揭示正确的偏误原因,有规律性的偏误。因此本文认为在揭示正确的偏误原因的过程中有可能发现相当多的模糊性偏误,而且已确定为有规律可循的偏误当中也有可能是被忽视或分类上有错的偏误。在利用大规模语料库资料的研究中,如果出现这种问题,有可能出现大量的偏误。

因此本论文提议以学习者语料库为基础的EA研究必须搜集母语句子,以此作为深入分析学习者意图的依据。不管是作文还是口语语料库,我们都需要搜集学习者的目的语和母语两个句子。

³² Corder 或 Selinker 都主张不应该研究个别的偏误,而是通过学习者偏误中出现的系统误差(systematic error)来分析中介语。

这种方法可能利用在大规模语料库时相当繁琐,也会有诸多困难,尽管如此,为了量化的可信度,更为了质化的正确分析,我们必须要考虑的这一点。

除了搜集母语,如果可以最好采访学习者,这样更有利于分析。因为我们可以通过采访解决研究期间无法找出确切答案的学习者个人差异等问题。如果采访所有的学习者在时间上或空间上受到一定的限制,研究者可先进行初步研究,对需要重新确认的资料进行小规模追踪调查。

还有一种妨碍正确的 EA 研究的要素,那就是想在常见偏误句型或语法项目中找出偏误时,学习者有意使用回避策略,那么我们就无法找出偏误。如果想从这个学习者身上找出偏误,我们只能对这个学习者进行多次材料搜集。学习者的输出(output)也有可能跟所给的话题或语法项目无关,跟研究者的意图无关。因此我们可以对各种话题进行多次提示,从而引出学习者能够表达的母语。与此同时我们可通过 EA 研究考察学习者真正的第二语言学习的过程,比起横向研究,纵向研究更为有用。对同一学习者对象的纵向研究中,我们可以了解学习者的阶段性习得水准,具体剖析发展过程。

1960年代以西方语言学为中心发展起来的EA研究针对的是多种语言,近来揭示出很多有价值的结果。本文的探讨不是否定这些已有的研究成果,而是为了在已有研究成果的基础上,进行更为准确、更为深入的研究。本文的意义在于虽然是初步的考察,但是指出了身为EA研究者都认同的分析时出现的模糊性部分,并把这一问题拿到桌面上,让大家对此都有所了解,并希望以此为契机,为更好地研究EA共同思考,尽快找到有效的解决方法。

参考文献

- [1] 김해령 역(王建勤 著).제 2 언어 습득 및 중국어 학습. 한국문화사, 2015
- [2] 박지영 · 김태은.한국인 중국어 학습자의 동목결합관계 양상 고찰 - 초중급 (low-intermediate) 학습자의작문 분석을 중심으로. 중국어문학논집, 2016(98): 119-156
- [3] 손정애. 현대중국어 어기조사 “了”의 오류 분석-“첨가와 누락”의 오류를 중심으로. 중국언어연구. 2014(51): 231-259
- [4] 유재원 · 최재영 · 정연실 · 노지영 · 백지훈 역(周小兵 · 朱其智 · 邓小宁 著 (2007)). 중국어학습자의 문법오류연구. 한국문화사, 2014
- [5] 윤유정. 한국학생 중국어 주·빈어 어순 오류 분석. 중국어문학논집. 2013(80): 219-249.
- [6] 전기정. 학습자 말뭉치를 통한 한국인 학습자의 오류 분석-이중목적어 관련 구문을 중심으로. 중국어문논총, 2012(53): 151-172
- [7] 박용진. (한국인 중국어 학습자를 위한) 중국어 교육 이론과 중간언어 말뭉치 분석, 역락, 2015
- [8] 崔希亮. 欧美学生汉语介词习得的特点及偏误分析. 世界汉语教学. 2005(3): 83-96
- [9] 方绪军. 中介语中动词句的配价偏误分析. 语言教学与研究. 2001(4): 39-47
- [10] 鲁健骥. 中介语理论与外国人学习汉语的语音偏误分析. 语言教学与研究, 1984(3): 44-
- [11] 鲁健骥. 中介语研究中的几个问题. 语言文字应用, 1993(1): 21-25
- [12] 鲁健骥. 外国人学习汉语的语法偏误分析. 语言教学与研究, 1994(1): 49-64
- [13] 李大忠. 偏误成因的思维心里分析. 语言教学与研究, 1999(2): 110-119
- [14] 吕文华、鲁健骥. 外国人学汉语的语用失误. 汉语学习, 1993(1): 41-44
- [15] 朴智英、金兑垠. 关于构筑 ColloConc 的报告——面向韩国学生的汉语搭配偏误分析. 中文教学现代化学报(4), 1-10
- [16] 孙德坤. 中介语理论与汉语习得研究. 语言文字应用, 1993(4): 82-92
- [17] 王建勤. 中介语产生的诸因素及相互关系. 语言教学与研究, 1994(4): 105-120

- [18] 肖奚强.韩国学生汉语语法偏误分析.世界汉语教学,2000(2): 95-99
- [19] 肖奚强.略论偏误分析的基本原则.语言文字应用,2001(1): 46-52
- [20] 肖奚强.外国学生“除了”句式使用情况的考察.语言教学与研究,2005(2): 34-40
- [21] 谢 福.基于语料库的留学生“是……的”句习得研究.语言教学与研究,2010(2): 17-24.
- [22] 袁毓林.试析中介语中跟“不”相关的偏误.语言教学与研究,2005(6): 39-47.
- [23] 赵金铭.外国人语法偏误句子的等级序列.语言教学与研究,2002(2): 1-9.
- [24] 赵葵欣.留学生学习和使用汉语介词的调查.世界汉语教学,2000(2): 100-106.
- [25] 郑艳群.中介语中程度副词的使用情况分析.汉语学习,2006(6): 66-72.
- [26] 周文华.基于语料库的外国学生兼习得研究.语言教学与研究,2009(3): 40-47.
- [27] 周文华.母语语序类型对目的语习得研究.语言教学与研究,2014(5): 10-17.
- [28] 周小兵、朱其智、邓小宁,外国人学汉语语法偏误研究.北京语言大学出版社,2007.
- [29] Corder, P.(1967), The significance of learners' errors. *IRAL*, 5, 161-170.
- [30] Darus, S. and Subramanian, K.(2009), Error analysis of the written English essay of secondary school students in Malaysia. *European Journal of Social Science*, 8:3, 483-495.
- [31] Dulay, C. and Burt, K.(1974), Error and strategies in child second language acquisition. *TESOL Quarterly*, 8, 129-138.
- [32] Ellis, R.(1994), The study of second language acquisition. London: Oxford University Press.
- [33] Heydari, P. and Bgheri, M.(2012), Error Analysis: Source of L2 Learners' Errors. *Theory and Practice in Language Studies*, 2:8, 1583-1589.
- [34] Jabeen, A.(2015), The Role of Error Analysis in Teaching and Learning of Second and Foreign Language. *Education and Linguistics Research*, 1:2, 52-61.
- [35] James, C.(2001), Errors in Language Learning and Use: Exploring error analysis. Beijing: Foreign Language Teaching and Research Press.
- [36] Khansir, A.(2012), Error Analysis and Second Language Acquisition. *Theory and Practice in Language Studies*, 2:5, 1027-1032.
- [37] Lennon, P.(1991), Error: some problems of definition and identification. *Applied Linguistics*, 12:2, 180-195.
- [38] López, W.(2009), Error Analysis in a learner corpus. What are the learners' strategies?, A survey of corpus based research, 675-690.
- [39] Nzama, V.(2010), Error Analysis: A study of Errors Committed by Isizulu Speaking Learners of English in Selected Schools. Unpublished M.A. Dissertation, University of Zululand.
- [40] Richards, C.(1971), A non-contrastive approach to Error Analysis. *English Language Teaching Journal*, 25, 204-219.
- [41] Richards, C.(1974), Error Analysis: Perspectives on second language acquisition. London: Longman.
- [42] Sawalmeh(2013), Error Analysis of Written English Essays: The case of Students of the Preparatory Year Program in Saudi Arabia. *English for Specific Purpose World*, 40:14, 1-17.