

英语母语背景汉语学习者书面中介语的词汇发展研究

王文卓

华侨大学华文教育研究院/华文学院，福建 厦门

收稿日期：2025年11月10日；录用日期：2025年12月2日；发布日期：2025年12月15日

摘要

本文基于自建语料库，从计量语言学的角度，以衡量词汇丰富度的三个指标：吉罗指数、密度及单现率，以汉语母语者数据为参照，考察了真实语料中不同水平英语母语汉语学习者在书面语中的词汇使用发展情况。中介语和母语者语料分别源于HSK动态作文语料库及《人民日报》，并使用Python处理语料得到数据后用SPSS进行相关系数检测分析。研究结果表明，随着汉语水平的提高，ECSL学习者书面语中词汇的吉罗指数及单现率也会提高，其输出的词汇密度符合书面语特征，输出词汇的丰富度和复杂度也得到了提升，并逐步向目的语水平靠近。

关键词

汉语中介语，书面语，词汇丰富度，语言水平

A Study on the Vocabulary Development of the Written Interlanguage of Chinese Learners with English Native Language Background

Wenzhuo Wang

Chinese Education Research Institute/Chinese Language and Culture College, Huaqiao University,
Xiamen Fujian

Received: November 10, 2025; accepted: December 2, 2025; published: December 15, 2025

Abstract

Based on a self-built corpus, this study adopts a quantitative linguistic approach to investigate the development of vocabulary usage in the written production of Chinese learners of different proficiency levels with English as their first language. Using three indicators of lexical richness—the Guiraud Index, lexical density, and the hapax legomena rate—the study compares learner data against a reference corpus of native Chinese speakers. The interlanguage data were sourced from the HSK Dynamic Composition Corpus, while the native speaker data came from the People's Daily. Corpus processing was carried out using Python, and correlation analyses were conducted via SPSS. The results show that as Chinese proficiency increases, ECSL learners exhibit higher values in both the Guiraud Index and hapax legomena rate in their writing. In addition, the lexical density of their output aligns with the profile of written discourse, indicating a development toward greater lexical richness and complexity, which progressively approximates target-language norms.

Keywords

Chinese Interlanguage, Written Language, Lexical Richness, Language Proficiency

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来，词汇知识成为第二语言习得研究中的重要内容之一。词汇知识分为词汇数量、质量及接收一产出控制三方面[1]。其中，词汇丰富性是词汇质量研究的重要内容[2]，可以用来测量作者对于词汇运用的广度和深度，是衡量二语写作整体语言水平的重要指标之一[3]-[5]，不仅直接影响学习者的语言表达能力，还在一定程度上反映了学习者的语言水平和语言习得的进程。

通过考察可以发现，前人已利用词语多样性、密度、新颖性、复杂程度和偏误率、词汇异同化等多个指标对汉语二语习得者的词汇深度知识的掌握情况进行了考察[6]-[8]，但同时也存在一些问题，如某些测量指标的计算方法存在弊端(如 TTR)、只是简单的统计词频，但并未检测数据是否有统计学意义上的差异等，难以准确衡量不同水平英语母语汉语学习者在书面语中的词汇使用发展情况。

因此，本研究拟基于自建的语料库，采用定量与定性的研究方法，从吉罗指数、密度及单现率三个维度，对英语母语背景汉语学习者(English native speaking Chinese second language learners，下文均简称ECSL学习者)书面中介语的词汇丰富度进行对比研究，分析不同水平的ECSL学习者在书面语中的词汇使用情况，不仅可以为汉语作为第二语言的教学提供理论支持，还可以为实践教学提供有益的参考，促进二语学习者的语言能力发展。希望通过本研究，可以进一步推动对ECSL学习者词汇发展的深入探讨。本文拟对以下问题进行研究考察：

- (1) ECSL学习者书面中介语词汇的吉罗指数有什么特点？与汉语水平是否相关？说明了什么？
- (2) ECSL学习者书面中介语词汇的密度有什么表现？与汉语水平是否相关？说明了什么？
- (3) ECSL学习者书面中介语词汇的单现率有什么表现？与汉语水平是否相关？说明了什么？

2. 语料与资源

本研究所依托的中介语书面语(W)语料，源自HSK动态作文语料库，选取语料涉及的作文均来自英

语母语背景的汉语学习者。根据选取作文的 HSK 成绩, 将书面语语料分为三个不同的水平等级组, L1 为初级组, L2 为中级组, L3 为高级组, 每个等级组内包括 35 篇作文, 共计 140 篇作文, 涵盖 16988 个词(不含标点), 统称为“二语组”。关于中介语各个语料库的具体词数、作文成绩、以及各二语水平的对应情况, 详见表 1。

Table 1. Basic data of the interlanguage written corpus**表 1.** 中介语书面语语料库基本数据

语料库	等级	词数	作文成绩
中介语书面语语料库(W)	L1	5784	40~60 分
	L2	5679	60~80 分
	L3	5525	80~100 分

为了探究不同水平的汉语学习者与汉语母语者在词汇层面上是否存在差异, 本文选择了与汉语中介语规模相当的汉语母语书面语语料作为参照语料库。本研究从《人民日报》(2010~2017 年)中随机抽取语料, 建立了包含 5288 个词的汉语母语者书面语语料库。《人民日报》作为典型、正式、简洁且规范的书面语体, 其语言风格客观严谨, 具有一定新闻报刊的特色, 能够有效地映射中介语词汇层面的特点, 这些母语者的语料后续将统一称为“母语组”(TL)。

我们将收集到的语料录入计算机, 使用 Python 3.8 中的中文分词组件 Jieba 进行分词、标注词性、频数统计, 得到的结果示例在 excel 表格中呈现如表 2 所示; 使用函数计算 ECSL 学习者不同语言水平下的词汇吉罗指数、密度和单现率。

Table 2. Top 15 high-frequency words, their frequencies and part-of-speech tags in L1 (Based on Jieba)**表 2.** 使用 Jieba 得到的词频、词及词性(以 L1 的前 15 个高频词为例)

频数	词	词性	频数	词	词性
567	的	uj	86	有	v
171	我	r	74	人	n
129	是	v	66	在	p
123	孩子	n	54	他们	r
108	父母	n	53	也	d
95	他	r	53	对	p

在此基础上使用 SPSS 27.0.1 进行相关系数分析, 检验上述关于词汇的三个指标与学习者的汉语水平是否相关。

3. 数据结果与分析

3.1. 吉罗指数

计量语言学中常用型例比(TTR, type-token ratio)来衡量词汇丰富度, 计算的是文本中不同词语(type)在所有词语(token)中所占的比例。TTR 的值越大, 表示文本中所使用的词汇越多样化。但 TTR 的值常受语料库规模大小的影响。为解决这一问题, 本文使用 TTR 的一种变体——“吉罗指数”进行计算, 在

一定程度上减少文本规模对词汇丰富度的影响，其计算公式为：Index of Guiraud = $\frac{\text{Types}}{\sqrt{\text{Tokens}}}$ (Types 是型符，Tokens 是类符)。我们使用 Python 计算二语组和母语组文本中不同的词语记为“类符数”，并把二语组和母语组文本中的所有词语记为“形符数”，计算每一等级的吉罗指数。二语组每一等级和母语组语料中的“类符数”、“形符数”和“吉罗指数”如表 3 所示。

Table 3. Calculation of guiraud's index for ECSL learners and native speaker groups
表 3. 二语组和母语组吉罗指数计算结果

等级	类符数	形符数	吉罗指数
L1	1351	5784	17.76
L2	1591	5679	21.11
L3	1759	5525	23.66
TL	1979	5288	27.21

由表 3 结果可知，母语组的吉罗指数为 27.21，高于二语组三个水平的吉罗指数，说明相对于汉语母语者所输出的书面语内容，ECSL 汉语学习者的书面用语用词的丰富度较低，用词比较集中。

我们使用 SPSS 软件将 ECSL 学习者不同语言等级(L1-L3)和其对应的吉罗指数两个变量进行相关系数检测，计算结果如表 4 所示。

Table 4. Results of the correlation analysis between proficiency levels and guiraud's index for ECSL learners
表 4. ECSL 学习者不同语言等级和吉罗指数的相关性检验结果

变量	等级	吉罗指数
等级	1	0.997*
吉罗指数	—	$p = 0.050$
	0.997*	1
	$p = 0.050$	—

注：**.在 0.05 级别(双尾)上相关性显著。

表 4 中的数据显示 ECSL 学习者的不同水平与吉罗指数指标具有显著相关性($R = 0.997, p = 0.050$)，即随着 ECSL 学习者二语水平的提高，其输出的中介书面语词汇的吉罗指数越来越高，词汇丰富度也越来越高，所使用的单词范围越来越广。

二语学习者的中介语系统并不是恒定不变的，而是在不断发展和变化的。一般来说，具有较高汉语水平的学习者会更善于使用不同的词汇[9] [10]，本研究的数据在一定程度上证明了这一点。随着水平的提高，ECSL 学习者所输出的书面语的吉罗指数呈现靠近目的语的趋势。

3.2. 词汇密度

词汇密度主要指实词在文本中所占比例，即本文中所有实词与总词数之间的比率。实词既有语法意义，又具有完整的词汇意义。因此，如果文本中所使用的实词越多，其语篇信息密度越大，意义范畴越广泛，内容表达越丰富。在汉语中，词类的划分以及实词和虚词的区分仍存在部分争议，本研究拟选取

名词、动词、形容词、副词作为实词[11]，进行词汇密度的计算，计算公式为：词汇密度 = $\frac{\text{实词数}}{\text{总词数}} \times 100\%$

各个等级对应的名词、动词、形容词、副词及实词数、所计算出来的词汇密度计算如表5所示。

Table 5. Content word counts and lexical density of ECSL learners and native speaker groups

表5. 二语组和母语组的实词数及词汇密度

等级	名词	动词	形容词	副词	实词数	总词数	密度
L1	1326	1294	213	510	3343	5784	57.80%
L2	1279	1292	206	505	3282	5679	57.79%
L3	1239	1334	216	449	3238	5525	58.61%
TL	1915	1393	226	240	3774	5288	71.37%

词汇密度的高低可以区分书面语与口语，通常书面语的词汇密度大于口语，书面语篇的词汇密度一般在40%以上，而口语语篇的词汇密度一般在40%以下[12]。由表4的数据可以看出，各个等级的二语学习者的词汇密度均大于40%，符合一般书面语篇的词汇密度，说明ECSL学习者在词汇的使用上具有较强的语体意识。同时，母语者的词汇密度为71.37%，显著高于L1~L3阶段，即相同长度的书面语语料中ECSL学习者所输出的内容实词少于母语者，其包含的信息量较低。我们使用SPSS软件将不同等级和其对应的词汇密度两个变量进行相关系数检测，数据结果显示二者无统计学意义上的相关关系，说明随着等级的提高，学习者实词的使用并无较为显著的发展，且词汇密度不能有效的区分ECSL学习者的不同语言水平，这与吴继峰[8]、李春琳[13]的发现一致。可能是因为选取的语料的其主题词类似，如表6的数据显示，ECSL学习者三个水平组前15位高频词中共享词位较多，如“的”、“是”、“我”、“人”、“父母”等，核心共享实词占实词总类型的54.5%，表明不同水平组在核心实词使用上具有高度同质性，此种主题重叠可能导致三组所输出的书面语内容在信息结构和信息量上类似造成词汇密度差异不显著[13]。同时，这一结果能在一定程度上说明词汇密度可能是反映语体意识而非语言水平的指标。由上述分析可知，二语组与母语组均显著高于书面语词汇密度[12]，表明学习者已形成书面语体意识，实词使用水平也较为固定，这种“中介语固化现象”[14]可能源于英语母语者倾向保留主语、连接词等结构，导致虚词比例难以减少，而母语组词汇密度更高(71.37%)，反映母语者能更高效地传递信息，具有较高的语言整合能力。

为了进一步研究该现象产生的原因及中介书面语中实词的使用情况，本研究分别统计了不同语言水平文本中居前15位的高频(字)词，如表6所示。

Table 6. Top 15 high-frequency words across ECSL learners and native speaker groups

表6. 二语组和母语组中的前15位高频(字)词表

L1			L2			L3			TL		
频序	高频词	占比									
1	的	9.80%	1	的	8.82%	1	的	9.05%	1	的	6.90%
2	我	2.96%	2	我	3.57%	2	是	2.41%	2	和	1.29%
3	是	2.23%	3	是	1.97%	3	我	2.14%	3	是	1.17%
4	孩子	2.13%	4	在	1.94%	4	在	1.30%	4	发展	1.15%

续表

5	父母	1.87%	5	人	1.13%	5	对	1.12%	5	在	1.12%
6	他	1.64%	6	父母	1.11%	6	也	1.05%	6	中国	1.08%
7	有	1.49%	7	了	1.09%	7	了	0.98%	7	了	0.96%
8	人	1.28%	8	他们	1.00%	8	父母	0.90%	8	经济	0.89%
9	在	1.14%	9	有	0.95%	9	人	0.89%	9	要	0.62%
10	他们	0.93%	10	不	0.93%	10	孩子	0.89%	10	碳	0.53%
11	也	0.92%	11	都	0.93%	11	和	0.76%	11	国家	0.51%
12	对	0.92%	12	他	0.85%	12	都	0.74%	12	低	0.51%
13	了	0.92%	13	也	0.74%	13	有	0.69%	13	我们	0.47%
14	都	0.85%	14	孩子	0.70%	14	一个	0.67%	14	为	0.47%
15	会	0.83%	15	就	0.70%	15	自己	0.62%	15	与	0.45%

就居前 15 位的高频词而言, 二语组中 L1 实词(是、孩子、父母、有、人、他们、都、会)的总占比为 11.6%, L2 中实词(是、人、父母、他们、有、不、都、孩子、就)的总占比为 9.44%, L3 中实词(是、也、父母、人、孩子、都、有)的总占比为 7.57%, 即随着 ECSL 学习者语言水平的提高, 其前 15 位的高频词占比逐渐降低。

我们随机选取二语组和母语组中的语料对此现象进行说明:

二语组:

L1: 小孩子是最受外部影响的人。人长得越大从别人受的影响越小。长大了人们的性格就基本上固定下来了, 但是小的时候我们的性格还在形成, 而那时候当然我们是跟父母接触的最多。

L2: 我开始学习汉语的时候, 真的不知道以后在我的生活中怎么用汉语, 可我已经在中国的时间不算少, 我现在知道我学汉语的目的是什么。

L3: 本来, 要学习一种外语, 对每一个初学者来说, 都有一定程度上的障碍和苦恼, 但是只要能把握学习的方法, 在学习中肯求一些趣味, 那么, 不但可以提高学习的效率和水平, 还可以为自己增添一些情趣吧!

母语组:

TL: 科技创新正在成为推动经济高质量发展的核心动力。人工智能与大数据技术的深度融合不仅提升了产业效率, 也为社会治理提供了智能化解决方案。这种技术驱动的转型需要平衡发展与伦理的关系。

Ellis [15] 提出了以频率为中心的二语习得理论, 即学习者习得接收到某种高频的语言特征信息, 相应地就会高频地输出相关语言项目, 频率是语言习得的关键[16]。ECSL 学习者在语言习得的早期阶段, 在书面语中更常见并倾向于使用一些基础的实词(如 L1 中的“孩子”、“人”等; L2 中的“我”、“汉语”等), 这些词汇通常是日常交流生活中必需的基础词汇, 出现的频率较高, 其词义含有更多的信息性, 可以加深学习者对书面语信息的理解。到了高水平写作阶段, ECSL 学习者可以逐渐从集中使用典型的实词转向使用难度较高的实词或虚词(如 L3 中的“障碍”、“苦恼”、“趣味”、“效率”等)。而母语组的书面语内容词语丰富度较高, 如 TL 中例句中实词(科技、创新、高质量、发展、核心、动力、人工智能、大数据)的总占比较低, 且由于新闻体裁的信息性、客观性和规范性, 出现了难度较高且带有强烈话题色彩的词语(如“深度融合”、“产业效率”、“社会治理”、“智慧化解决方案”等), 其不易在文本中重复, 因此其前 15 位的高频实词占比较低, 且主题词较多。

其次，汉语的主题词可以将其语义范围与后续句子共享[17]，如 TL 例句所示，整句话以“科技创新”为话头展开，突显性较强，后续小句与其共享话头，形成了一个完整且连贯的语言结构。而中介语中的流水句较多，语篇语言连贯性较弱，如 L1 例句所示，三个短句中话头分别为“小孩子”、“人”、“人们的性格”，且第三个短句分句中的话头也由“人们的性格”、“我们的性格”变为了“我们”，以此表现出其相同话头的持续性较弱，且如果在连续小句中频繁重复话头或创建新的话头过程，不仅会容易加重读者的认知负担，还会降低语篇的连贯性，带来语篇内容的冗余、单调，不符合语言的经济原则[18]，随着 ECSL 学习者水平的增高，此现象得到了一定缓解，如 L2、L3 例句所示。因此，随着汉语水平的提高，中介书面语中出现的话头可以在一定程度上延续，增加了书面语的可读性和连贯性，相同实词在书面语中的使用频率也随之降低，语言的经济性得以体现，逐渐向目的语水平靠近。

第三，英语中多出现从属结构，其书面语句繁长且复杂[19][20]，而汉语句子较短[21]，分句或流水句较多，汉语母语者倾向于逐层叙述思维的整个过程[20]，因此，ECSL 学习者容易受到英语句子结构特征的影响，在汉语写作时仍倾向使用较长的从句及长句[22]，ECSL 学习者话题链的平均长度与话题链的平均分句长度在整个发展过程中均低于汉语母语者水平，且其发展过程中存在暂时性的停滞[23]，因此，我们可以猜测，受英语母语影响，ECSL 学习者可能会出现词汇石化的现象[14]，其在实词话题链产出数量上和长度上均不及汉语母语者，使用频率也高于母语者水平。

3.3. 单现率

单现词是文本中只出现一次的词语，文本的词汇丰富度与单现词呈正比，即单现词越多，其文本的词汇丰富度越高。单现率即单现词所占比率，计算公式为 $\text{单现率} = \frac{\text{词频为1的词数}}{\text{总词数}} \times 100\%$ ，各个语言水平的单现词总数及计算得出的单现率如表 7 所示。

Table 7. Hapax legomena and their percentage in ECSL Learners and native speaker groups
表 7. 二语组和母语组的单现词及单现率

等级	单现词	总词数	单现率
L1	739	5784	12.78%
L2	934	5679	16.45%
L3	1107	5525	20.04%
TL	1249	5288	23.62%

由表 7 结果可知，母语组的单现率为 23.62%，高于二语组三个水平的单现率，说明相对于汉语母语者所输出的书面语内容，ECSL 汉语学习者的书面用语用词的丰富度较低，用词也比较集中。

我们使用 SPSS 软件将 ECSL 学习者不同语言等级和其对应的词汇单现率两个变量进行相关系数检测，计算结果如表 8 所示。

Table 8. Results of the correlation analysis for proficiency levels and hapax legomena rate in ECSL learners
表 8. ECSL 学习者不同语言等级和词汇单现率的相关性检验结果

变量	等级	单现率
等级	1 —	1.000** $p = 0.004$

续表

单现率	1.000** <i>p</i> = 0.004	1 —
-----	-----------------------------	--------

注：**在0.01级别(双尾)上相关性显著。

表8中的数据显示ECSL学习者的不同水平与单现率指标具有显著相关性($R = 1.000, p < 0.010$)，即随着ECSL学习者二语水平的提高，其输出的中介书面语的词汇单现率越来越高，词汇丰富度也越来越高。在良好的学习氛围中，ECSL学习者运用语言知识的能力会越来越强，二语水平与学习时间大多为正向相关趋势[15]。二语水平的提升对于词汇的习得明显具有积极的影响。与前文计算“吉罗指数”所得到的结论一致，随着学生汉语水平的提高，其写作产出的词汇多样度和词汇复杂度都显著上升。

4. 结语

本研究基于自建语料库，从词语的吉罗指数、密度及单现率出发，考察了真实语料中不同水平英语母语汉语学习者在书面语中的词汇丰富度的使用发展情况。研究结果表明，二语水平的提升是循序渐进的，充满了动态性与复杂性，随着汉语水平的提高，ECSL学习者书面语输出词汇的丰富度和复杂度得到了提升，并逐步向目的语水平靠近，词汇的吉罗指数及单现率提高，且词汇密度符合书面语特征，但其词汇密度差异并不显著。

本研究也存在一些不足之处，如ECSL学习者书面语写作中的词汇丰富性可能会受话题熟悉度、写作策略等因素的影响，导致某些衡量指标难以得到更加清楚的体现；其次，本研究主要是比较不同学习者水平的作文文本，缺乏纵向观察；除本研究提到的三个指标，还有其他的指标可以进行计算。希望后续的研究可以对ECSL学习者的词汇习得达到更有效、全面的考察。

参考文献

- [1] Henriksen, B. (1999) Three Dimensions of Vocabulary Development. *Studies in Second Language Acquisition*, **21**, 303-317. <https://doi.org/10.1017/s0272263199002089>
- [2] Nation, P. and Webb, S. (2011) Researching and Analyzing Vocabulary. Heinle Centre Learning.
- [3] Laufer, B. and Nation, P. (1995) Vocabulary Size and Use: Lexical Richness in L2 Written Production. *Applied Linguistics*, **16**, 307-322. <https://doi.org/10.1093/applin/16.3.307>
- [4] 万丽芳. 中国英语专业大学生二语写作中的词汇丰富性研究[J]. 外语界, 2010(1): 40-46.
- [5] 朱慧敏, 王俊菊. 英语写作的词汇丰富性发展特征——一项基于自建语料库的纵贯研究[J]. 外语界, 2013(6): 77-86.
- [6] 黄立, 钱旭菁. 第二语言汉语学习者的生成性词汇知识考察——基于看图作文的定量研究[J]. 汉语学习, 2003(1): 56-61.
- [7] 曹贤文, 邓素娟. 汉语母语和二语书面表现的对比分析——以小学高年级中国学生和大学高年级越南学生的同题汉语作文为例[J]. 华文教学与研究, 2012(2): 39-46.
- [8] 吴继峰. 英语母语者汉语写作中的词汇丰富性发展研究[J]. 世界汉语教学, 2016(1): 129-142.
- [9] Ortega, L. (2003) Syntactic Complexity Measures and Their Relationship to L2 Proficiency: A Research Synthesis of College-Level L2 Writing. *Applied Linguistics*, **24**, 492-518. <https://doi.org/10.1093/applin/24.4.492>
- [10] Pallotti, G. (2014) A Simple View of Linguistic Complexity. *Second Language Research*, **31**, 117-134. <https://doi.org/10.1177/0267658314536435>
- [11] 胡显耀. 基于语料库的汉语翻译小说词语特征研究[J]. 外语教学与研究, 2007(3): 214-220, 241.
- [12] Ure, J. (1971) Lexical Density and Register Differentiation. In: Perren, G.E. and Trim, J.L.M., Eds., *Applications of Linguistics: Selected Papers of the Second International Congress of Applied Linguistics*, Cambridge University Press, 443-452.

-
- [13] 李春琳. 汉语二语学习者产出型词汇水平和写作质量相关关系分析[J]. 华文教学与研究, 2017(3): 54-61.
 - [14] Selinker, L. (1972) Interlanguage. *IRAL—International Review of Applied Linguistics in Language Teaching*, **10**, 209-231. <https://doi.org/10.1515/iral.1972.10.1-4.209>
 - [15] Ellis, N.C. (2002) Frequency Effects in Language Processing. *Studies in Second Language Acquisition*, **24**, 143-188. <https://doi.org/10.1017/s0272263102002024>
 - [16] 文秋芳. 频率作用与二语习得——《第二语言习得研究》2002年6月特刊评述[J]. 外语教学与研究, 2003(2): 151-154.
 - [17] 曹逢甫. 主题在汉语中的功能研究——迈向语段分析的第一步[M]. 北京: 语文出版社, 1995.
 - [18] 朱钰麒, 熊文新. 基于小句复合体理论的汉语中介语语篇连贯性探讨[J]. 语言教学与研究, 2023(1): 1-11.
 - [19] Crews, F.B. (1974) *The Random House Handbook*. Random House.
 - [20] 连淑能. 英汉对比研究(增订本) [M]. 北京: 高等教育出版社, 2010.
 - [21] 王寅. 英汉语言宏观结构区别特征[J]. 外国语(上海外国语学院学报), 1990(6): 38-42, 26.
 - [22] 吴继峰. 英语母语者汉语书面语句法复杂性研究[J]. 语言教学与研究, 2016(4): 27-35.
 - [23] Hao, Y., Wang, X. and Yu, Q. (2021) Typological Characteristics of Interlanguage: Across Native Language Types and L2 Proficiency Levels. *Lingua*, **257**, Article ID: 103085. <https://doi.org/10.1016/j.lingua.2021.103085>