British Journal of Education
Vol.10, Issue 10, pp.15-23, 2022
Online ISSN: 2054-636X (Online)
Print ISSN: 2054-6351(Print)

Acquisition of L3 writing proficiency in Hungarian Japanese learners

Wenchao Li
School of International Studies Zijingang Campus,China

Citation: Wenchao Li (2022) Acquisition of L3 writing proficiency in Hungarian Japanese learners, *British Journal of Education*, Vol.10, Issue 10, pp.15-23

ABSTRACT: *The present study investigates writing proficiency in Hungarian Japanese learners to examine how language closeness between the mother tongue and any language acquired later in life might affect the learning quality of the second language. Writing proficiency is measured at two levels: syntactic complexity based on dependency distance, and lexical complexity using moving-average morphological richness (MAMR) and moving-average mean size of paradigm (MAMSP) measures. Study findings suggests that in terms of essay writing (writing style: simple), both MAMR and MAMSP of Japanese written by Hungarians are extremely close to native Japanese writing. Regarding email writing, which requires both honorific and humble forms of language, both lexical and syntactic complexity of Hungarian Japanese learners are characterised by slightly reduced richness than those of native Japanese persons, but remain very close. These outcomes based on a lexical and syntactic examination of Hungarian Japanese learners indicate that the mother tongue (L1) affects the quality of later language acquisitions.*

KEYWORDS: L3 acquisition, writing proficiency, Hungarian, Japanese, syntactic complexity

INTRODUCTION

Hungarian is a member of the Uralic language family. An illustration is provided in (1).

(1)	Egy	lány-t	vár-ok.
	DEM	young lady-ACC	wait I-NOM

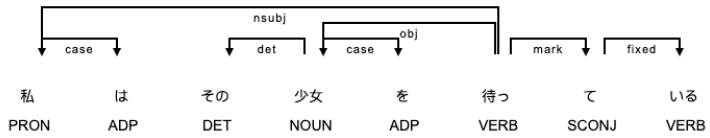
'I am waiting for that young lady.'

Intriguingly, the Uralic family was once grouped with the Altaic language family for its similarities in three respects: (a) phonology: vowel harmony; (b) morphology: agglutinative; the suffix plays the main role in rendering tense, aspect, voice,

derivation, inflection/conjugation; (c) word order: subject-object-verb (S-O-V) and postposition.

These linguistic features are the same as in Japanese, another possible Altaic language member. (1) corresponds to (2) in Japanese.

(2)	私	は	その	少女	を	待っている.
	Watashi	wa	sono	shoojo	o	matteiru.
	I	TOP	DEM	young lady	ACC	wait.PROG.CONCL



Although there is no evidence of the existence of vowel harmony in Modern Japanese, the undeniable similarities in syntax and morphology between the two suggest that the performance of Hungarian learners when acquiring Japanese would be excellent. To examine this hypothesis, lexical and syntactic complexity of 72 narrative essays and emails "asking for a favour" written by Hungarian Japanese learners were studied. The essays were written in plain form. The emails were written in a combination of honorific and humble forms. By examining the morphosyntactic complexity of writings in all forms, we will be able to understand how language closeness between the mother tongue and the later-acquired language may affect the learning quality of the latter.

In this article, Section 2 outlines the Methodology (including the framework, corpora, syntactic parser and mean dependency distance [MDD] calculation), Section 3 is devoted to the Results and Discussion, and Section 4 presents the Conclusion.

DATA AND METHODS

Data

The Hungarian writing data were drawn from the International Cross-Sectional Corpus of Japanese as a second language: 72 compositions written in Japanese, with essays titled "Our Eating Life: Fast Food and Home-Made Food" and emails titled: "asking for a recommendation" were extracted. These were written in the plain, honorific and humble forms. The number of tokens in these compositions totalled 35,028 words. Essays and emails written by native Japanese were also selected as

reference for Hungarian Japanese learners. Details of the study texts are provided in Table 1.

Table 1. The materials

Material	Total words	Title	Writing form
Essays	22932	Our Eating Life	Plain form
Emails	12096	Asking for recommendation Letter	Honorific and humble form

Analysis

The primary aim of the present study was to explore how Hungarian Japanese learners perform in writing tasks. To this end, lexicon richness and syntactic complexity of texts written by Hungarian Japanese learners were employed as metrics. Specifically, MDD was employed for measuring syntactic diversity in the writings. The moving-average morphological richness (MAMR) and moving-average mean size of paradigm (MAMSP) were calculated for measuring lexicon richness of the writings. The MDD, MAMR, MAMSP were computed using self-written computer programme scripts.

Syntactic complexity of writings by Hungarian Japanese learners

The concept, "dependency distance", is part of dependency grammar (Tesnière 1959; Yngve 1960; Hudson 2007; Liu 2009b), and refers to the distance between the governor and the dependent. The governor refers to the core linguistic element in a sentence, such as the verb, or the predicate. The dependent refers to the subject, object, oblique, adverb, post/prepositional phrase, etc. Liu, Hudson and Feng (2009) propose measuring the dependency distance (DD) based on |governor − dependent|. The MDD of the whole sentence would be:

$$\mathrm{MDD} = \frac{1}{n}\sum_{i=1}^{n} |\mathrm{DD}_i|$$

In regard to Example (2), 私はその少女を待っている [Watashi wa sono shoojo o matteiru], there are six types and seven tokens of dependency relationships, e.g., nsubj, case, det, obj, case, mark, fixed. The governor is the verb, *matsu* 'to wait'. The MDD in this case is 1.5.

Table 2. Dependency relation and direction of sentence (2)

Dependency direction	**Dependency relation**
私 ← 待っ	nsubj
は ← 私	case
その ← 少女	det
少女 ← 待っ	obj
を ← 少女	case
待っ ← 待っ	ROOT
て ← 待っ	mark
いる ← て	fixed

Essentially, the Japanese word order seems relatively free. A three-valance benefactive event such as "Mary gives an apple to John" can be encoded in six expressions.

(3) a. Taroo ga Jiroo ni ringo o **ageta**. b. Taroo ga rongo o Jiroo ni **ageta**.
c. Jiroo ni Taroo ga ringo o **ageta**. d. Jiroo ni ringo o Taroo ga **ageta**.
e. Ringo o Taroo ga Jiroo ni **ageta**. f. Ringo o Jiroo ni Taroo ga **ageta**.

That is, provided the verb, '*ageta'*, as the governor, always appears at the end, other elements can move freely, together with the case marker, i.e.,

Possible word orders in Japanese

AがBにCをV；AがCをBにV；BにAがCをV；BにCをAがV；CをAがBにV；

C を B に A が V

In Hungarian, however, the basic word order follows the S-V-O pattern, with negating words being always placed at the beginning. In this study, we aimed to explore how the difference in word orders might affect writing. Dependency direction of the writings were therefore examined.

Lexical complexity of writings by Hungarian Japanese learners

To measure lexical richness of the writings, MAMR (Cech, Kubat 2018) was employed as the metric. It has been suggested by Covington and McFall (2010), Yan and Liu (2021), Li, Liu and Li (2022) that a better average type-token ratio (TTR) can be obtained by using a moving window. The moving window for TTR in terms of word forms can be obtained using the following formula:

$$MATTR\ \ (W)_{word\ \ form} = \frac{\sum_{I=1}^{N-W+1} F_i}{\mathrm{W}\ \ (\mathrm{N}-\mathrm{W}+1)}$$

The moving window for TTR in terms of lemma can be obtained in the same way, i.e.,

$$MATTR\ \ (W)_{lema} = \frac{\sum_{I=1}^{N-W+1} F_i}{\mathrm{W}\ \ (\mathrm{N}-\mathrm{W}+1)}$$

Lexical richness is obtained by $\frac{\sum_{I=1}^{N-W+1} F_i}{\mathrm{W}\ \ (\mathrm{N}-\mathrm{W}+1)} - \frac{\sum_{I=1}^{N-W+1} F_i}{\mathrm{W}\ \ (\mathrm{N}-\mathrm{W}+1)}$

RESULTS AND DISCUSSION

Drawing on the methodology described above, this section provides an assessment of the writing proficiency of essays and emails written by Hungarian Japanese learners. Section 3.1 presents the DD that was calculated using a computer programme. Section 3.2 examines whether writing quality fits a certain distribution model, and if so, the parameters that may suggest a trend in the probability distribution of learning Japanese.

Syntactic complexity of writings by Hungarian Japanese learners

The MDD of essays written by Hungarian Japanese leaners ranges from 1 to 4.1. The distance between most tokens was 2.3 and 2.4 The MDD is 2.71. The MDD of essays written by native Japanese individuals ranges from 0.5 to 4.2. the distance between most tokens was 3.0 and 2.5. The MDD is 2.85, slightly longer than that of essays written by Hungarians. We also studied the dependency distances of 36 emails written

by Hungarian Japanese learners and native Japanese persons. The emails were written in honorific and humble forms. The MDD of emails written by Hungarian Japanese leaners ranges from 1 to 3.4. Most tokens go to the distance of 2.0. The MDD is 2.63, which is shorter than that of the essays. The MDD of emails written by native Japanese persons ranges from 0.5 to 4.4, indicating a larger freedom than emails written by Hungarians. Most tokens go to the distance of 3.0. The MDD is 2.94, which is longer than that of the essays. This is due to the fact that emails asking for favours usually use the honorific and humble forms, whilst narrative essays use the plain form. The foregoing discoveries are summarised in Table 3.

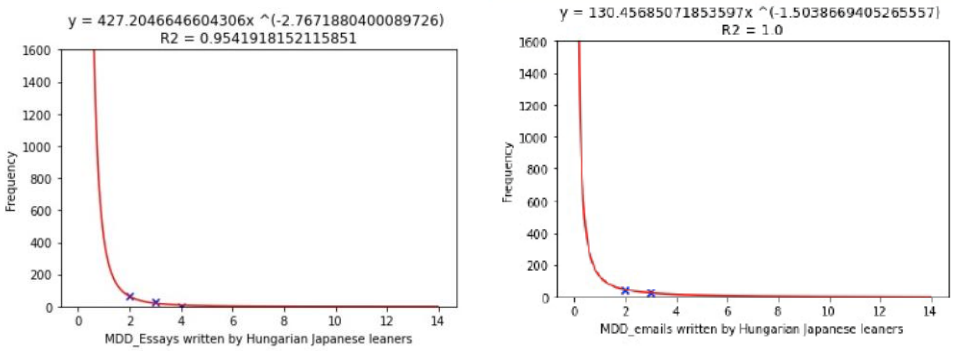
Table 3. Syntactic complexity of writings by Hungarian Japanese learners

	Essays			**Emails**		
	minimal MDD	maximal MDD	the DD that attributes the most tokens	minimal MDD	maximal MDD	the DD that attributes the most tokens
Japanese native	0.5	4.2	2.5	0.5	4.4	3
Hungarian Japanese leaners	1	4.1	2.4	1	3.4	2

The essays and emails written by native Japanese persons closely resemble texts written by native Japanese persons in terms of MDD freedom, maximal MDD and the DD that bears the most tokens. A further study of the regularity of the distributions and frequencies of MDD reveals that MDD-frequency relationships in Hungarian-written data demonstrates a very good fit to the power law function ($y = ax^b$) in both data, with 0.9541 and 1.0 as the determination coefficient, R^2 ($R^2 > 0.90$, very good; $R^2 > 0.80$, good; $R^2 > 0.75$, acceptable; $R^2 < 0.75$, unacceptable). The fitting result of Japanese-written data does not appear to be good.

Table 4. Fitting the power law function to the MDD of Hungarian-written texts

	Essay MDD	***a***	***b***	**R^2**	**Fitting results**
Hungarian	2.71	427.20	-2.76	0.9541	$y = 427x^{-2.76}$
	Email MDD	***a***	***b***	**R^2**	**Fitting results**
Hungarian	2.63	130.45	-1.5	1.0	$y = 130x^{-1.5}$



Lexical complexity of writings by Hungarian Japanese learners

We then estimated lexical complexity, for which MAMR and MAMSP were employed as metrics. The higher the MAMR and MAMSP, the richer the lexicon. A comparison of results between the data by native Japanese people and Hungarians are presented in Table 4.

Table 4. A comparison of writings by Hungarian Japanese learners and native Japanese persons using MAMR and MAMSP as metrics

Hungarian Japanese leaners	**Average MAMR**	**Average MAMSP**	**Native Japanese**	**Average MAMR**	**Average MAMSP**
Essay	0.026	1.053	Essays	0.027	1.054
Emails	0.024	1.050	Emails	0.031	1.068

It seems that in the written essay in plain form, both MAMR and MAMSP of Japanese texts written by Hungarians were extremely close to the writings made by native Japanese persons. With regard to emails "asking for a favour", where both honorific and humble forms were used, Hungarian Japanese learners displayed a close but lower MAMR and MAMSP than native Japanese-written texts. These outcomes lead us to deduce that the mother tongue (L1) does affect later language acquisition at the lexical

and syntactic levels.

CONCLUSION AND FUTURE RESEARCH

This study incorporates a mathematical linguistic approach to examine writing proficiency in Hungarian Japanese learners. Data were extracted from 72 essays and emails written in Japanese by Hungarians, covering the plain, humble and honorific forms of language. Writing proficiency was measured at the lexical and syntactic levels using MAMR and MAMSP for measuring lexical richness, and MDD for measuring syntactic complexity.

Regarding essay writing (writing style: plain form), MAMR and MAMSP of Hungarian-written Japanese closely resemble writings by native Japanese persons. In terms of email writing, involving both honorific and humble forms, both lexical and syntactic complexity of Hungarian Japanese learners suggest slightly less richness than native Japanese persons, but the two were still very close. These outcomes based on our lexical and syntactic examination of essays and emails lead us to conclude that there is a link between L1 mother tongue and the quality of later-acquired language.

A further study of the regularity of the distributions and MDD frequencies revealed that the MDD-frequency relationship in the Hungarian-written data demonstrated a very good fit to the power law function ($y = ax^b$) in narrative essay and "asking for a favour" email writing.

References

Čech, Kubát 2018 — Čech R., Kubát M. Morphological richness of text. *Taming the corpus:*

From inflection and lexis to interpretation. Fidler M., Cvrček V. (eds.). Cham: Springer, 2018, 63–77. DOI: 10.1007/978-3-319-98017-1_4.

Hudson, R. 2007. *Language networks: The new word grammar*. Oxford: Oxford University Press.

Liu, H. 2009. Probability distribution of dependencies based on Chinese dependency treebank. *Journal of Quantitative Linguistics*, 16(3), 256–273

Liu, H., Hudson, R., Feng, Z. 2009. Using a Chinese treebank to measure dependency distance. *Corpus Linguistics and Linguistic Theory* 5 (2), 161–174.

Li W., Liu H., Xiong. 2022. Nihongo ni okeru gojun no jiyuudo to kakuhyooshiki no hoofusa ni kansuru keeryooteki kenkyuu [A Quantitative Study of Word Order Freedom and

Case Marking Richness in Japanese]. Keeryoo kokugo gaku [Journal of Mathematilinguistics] 33 (5), pp. 325-340.

Tesnière, L. 1959. *Éléments de syntaxe structurale*. Paris: Klinksieck.

Yan J., Liu H. Morphology and word order in Slavic languages: Insights from annotated corpora. *Voprosy Jazykoznanija*, 2021, 4: 131–159.

Yngve, Victor H. 1960a. A model and a hypothesis for language structure. *Proceedings of the American Philosophical Society*, 104(5): 444–466.