

LANGUAGE DISTANCE AND L3 JAPANESE ACQUISITION IN MORPHOSYNTACTIC MODULE

Wenchao Li

Department of Japanese Studies, Zhejiang University, Hangzhou City, China

ABSTRACT

This study applies mathematical linguistics to explore how language distance plays an essential role in third language acquisition in terms of a morphosyntactic module. Data were drawn from 3410 essays written in Japanese by low, middle and high levels of learners from 12 first-language (L1) backgrounds who acquire English as a second language (L2)-interlanguage and Japanese as a third language (L3). The findings indicate that (a) mean dependency distance is an efficient indicator for syntactic complexity of writing proficiency. In both elementary and intermediate groups, learners of highly agglutinative languages are likely to show higher dependency distance than learners from isolated-language and fusion-language backgrounds. (b) The frequency and dependency distance are distributed in Power Law Function. Fitting Right truncated Good to the dependency distances indicates that the values of the parameter p ascend as the degree of agglutination of learners' mother tongue increases. (c) The syntactic complexity in multi-background Japanese learners' essays highlights that no matter how diverse the learners' native and target languages are, the syntax is always constrained by universal law, namely, minimising dependency distance. This is in accordance with existing findings in second language acquisition of inflectional languages.

KEYWORDS

language distance, writing proficiency, mathematical linguistics, L3 acquisition, mean dependency distance

1. INTRODUCTION

Typological distance and learning order have been deemed the most essential factors in language acquisition. Many scholars from different camps have attempted to determine which factor is more important and to what extent transfer from previously acquired languages happens (wholesale or piecemeal). Existing assumptions fall into three groups: (a) the second language (L2) has a significant influence at the early stage regarding phonological and syntactic acquisition (Hammarberg and Hammarberg 1993; Marx 2002; Tremblay 2007; Williams and Hammarberg 1998; Onishi 2016; Bardel and Falk 2007; Falk and Bardel 2011; Forsyth 2014; Archibald 2019; Dziubalska-Kolaczyk and Wrembel 2017; Onishi 2016; Wrembel 2015), (b) typological distance plays an essential role in lexis learning (Cenoz 2001; Rossi et al. 2006; Rothman 2010; Singleton 1987), and (c) the Cumulative-Enhancement Model (Flynn et al. 2004), Scalpel Model (Slabakova 2016) and Linguistic Proximity Model (Westergaard et al. 2017) show that language transfer is selective.

Owing to the development of natural language processing, another line of notable research contributes a great deal to the field: the quantitative approach (Hunt 1965; Lu 2011; Jiang et al. 2019). Quyang and Jiang (2018) examined the syntactic complexity of Chinese L2 English writing and arrived at regularity: with increased learners' grades, writing quality improves. Jiang et al. (2019) explored writing development across beginner and intermediate L2 English learners,

finding that the mean T-unit length, mean sentence length, and dependent clauses per clause are effective measures of writing quality. In line with the research on the quantitative approach, however, there remains room for further investigation. First, the target language of previous work seems to be limited to English, a (relatively) morphologically isolated language for which writing proficiency appears to be better captured at the syntactic level (e.g., relative, coordinate, and subordinate clauses). Furthermore, existing studies tend to focus on second language acquisition (SLA) more than third language acquisition (TLA). To our knowledge, the only study to explore third-language (L3) Japanese writing quality is that of Komori et al. (2019), which deemed the mean hierarchical distance an efficient index of Japanese learners' writing. Based on Quyang and Jiang's (2018) insights, Li and Yan (2021) examined three levels of Japanese English learners and demonstrated that interlanguage follows certain linguistic laws regardless of the learner's native language. This inspires us to consider the role a learner's first language plays in learning Japanese as a third language. Moreover, thus far, the metrics in measuring learning proficiency include the mean sentence length, mean clause length (Hunt 1965), T-unit length (Hunt 1970), noun phrases (Biber et al. 2011), number of coordinate/subordinate clauses (Bulte and Housen 2012), degree of sophistication (Ai and Lu 2013), mean dependency distance (MDD) (e.g., Liu 2008, 2017; Quyang and Jiang 2018), and mean hierarchical distance (Jing and Liu 2015; Komori et al. 2019).

Given that Japanese is morphologically agglutinative, one or more suffixes (causative, negation, voice, tense, or honorification) are added to a verb/adjective stem to create complex predicates. The dependency direction at the lexical level is thus head-initial, contrary to syntactic structure (head-final). Incorporating this, it would be necessary to consider lexical complexity when indexing the writing proficiency of agglutinative languages (e.g., Altaic, Korean, Indonesian, Hungarian). For instance, the verb 返す [*kaesu*.plain form], its honorific form *okaeshininaru* [*kaesu*.honorific form], and its humble form *okaeshisuru* [*kaesu*.humble form] have the same semantic meaning "to return sth" but different pragmatic functions. Word type-token ratio has a great deal to do with the degree of honorification and, thus, should be considered a crucial candidate for indexing spoken and written quality.

The present study investigates 360 essays written by L3 Japanese learners from 12 L1 language backgrounds (Vietnamese, Thai, Chinese, English, Russian, Indonesian, French, Turkish, German, Spanish, Hungarian, and Korean). By calculating the differences in lexical and syntactic complexity between writings at different levels and learners' mother tongues, this study aims to understand the association between L1-L3 language distance and L3 acquisition proficiency. The following questions are addressed:

Question 1: Can mean dynamic mean dependency distance, type-token ratio, and the distribution of long dependency relationship types indicate L3 writing proficiency?

Question 2: While Japanese is agglutinative and has an SOV word order, the research target in this study (i.e., Japanese learners from 12 L1 backgrounds) bears different morphosyntactic features (see Table 1). Is there an indicator in the probability distribution of L3 proficiency? If so, is it the *morphological distance* (distance between mother tongue and L3 language: agglutination/fusion/isolation) that suggests a trend in learning level?

Question 3: It has been explicated that the power law function may reflect the probability distribution of the dependency distance of second language learners' language proficiency (e.g., Quyang et al. 2018). Does L3 Japanese acquisition across diverse L1 backgrounds follow a similar regularity? Are there parameters suggesting a trend in learning quality?

Table 1. Japanese learners from multiple backgrounds

L1	Family	Languages	Morphology	Word order	Adjective Noun	Adposition
HHG	Uralic	Hungarian	Highly agglutinative	SVO	Adj-Noun	Postposition
TTR	Altaic	Turkish	Highly agglutinative	SOV	Adj-Noun	Postposition
RRS	Slavic	Russian	Inflectional	SVO	Adj-Noun	Preposition
IID	Austronesian	Indonesian	Agglutinative	SVO	Noun-Adj	Preposition
TTH	Kra-Dai language	Thai	Isolating	SVO	Noun-Adj	Preposition
VVN	Austroasiatic	Vietnamese	Isolating	SVO	Noun-Adj	Preposition
CCH	Sino-Tibetan	Chinese	Isolating	SVO	Adj-Noun	Preposition
EUS	Indo-European (Germanic)	English	Relatively isolating	SVO	Adj-Noun	Preposition
GAT	Indo-European (Germanic)	German	Inflectional	SVO	Adj-Noun	Preposition
FFR	Indo-European (Italic)	French	Inflectional	SVO	Noun-Adj	Preposition
SES	Indo-European (Italic)	Spanish	Inflectional	SVO	Noun-Adj	Preposition
KKR	unknown	Korean	Agglutinative	SOV	Adj-Noun	Postposition
JJJ	unknown	Japanese	Agglutinative	SOV	Adj-Noun	Postposition

2. DATA AND METHODS

2.1. Data

Data were drawn from the International Cross-Sectional Corpus of Japanese as a Second Language¹. Following the scores of Japanese Computerized Adaptive Test², this study extracts 1705 low, middle and high learners from Hungary, Russia, Indonesia, Thai, Vietnam, Chinese, England, Germany, France, Spain Korea and Turk. A totally of 3410 compositions written in Japanese titled “Our Eating Life: Fast Food and Home-Made Food” was extracted (1705 learners × 2 tasks). The tokens of these essays amount to 633,000 words. 100 essays by native Japanese were collected for reference.

2.2. Procedures

The present study is designed to examine third language acquisition proficiency with a focus on the morphosyntactic module. To this end, writing proficiency is measured via syntactic complexity. Essays written by Japanese learners across different L1 backgrounds are classified into two levels: primary and intermediate. The classification is made by the “Learner Text Evaluation System”. By looking into the writing proficiency of different levels, we may understand whether language distance plays a role in learning quality. Moreover, considering conjunctions might be missing in compositions, leading to an undetected dependency relationship (e.g., the run-on sentence “ringo ga suki, hoshii” [apple-NOM-like, want]; I like apples, so I want [apples])—sentence length alone may not be sufficient to represent syntactic complexities. This

¹ Sakoda (2020:10) points out that SLA refers to learning a language other than the first language.

² https://j-cat.jalesa.org/?page_id=168

study thus delves into the distribution of MDD for determining the level of writing proficiency. The following procedures were carried out:

Step 1: Draw raw data from the corpora

Step 2: Classify the writings into two levels: elementary and intermediate via the Learner Text Evaluator

Step 3: Parse each sentence via the GiNZA v4 Parser

Step 4: Produce a computer program to calculate the dynamic MDD from the parsed outputs

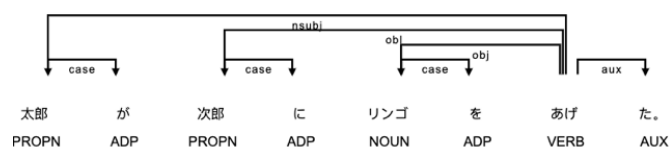
Step 5: Explore the probability distribution and the parameters that index writing proficiency

Step 6: Produce a computer program to verify the Spearman correlation and Euclidean distance clustering between the essays written by Japanese natives and learners from different L1 backgrounds

2.3. Data Analysis

As previously noted, this study employs dependency distance to indicate the syntactic complexity of essays written in Japanese. The framework is Dependency Grammar (Tesnière 1959; Hudson 2007; Liu 2009b), which contains three concepts: dependency relationship, dependency direction, and dependency distance. A dependency relationship is characterized by being binary and asymmetrical. Among syntactic structures, the verb is the GOVERNOR, and all elements are connected via a ‘governor-dependent’ relationship. For instance, in *Taroo ga Jiroo ni ringo o ageta* [Taroo-NOM-Jiroo-DAT-apple-ACC-give.PAST], the verb *ageta* ‘gave’ behaves as the governor on which the subject *Taroo*, the direction object *ringo* ‘apple,’ and the indirect object *Jiroo* depend. The straight arrow is directed from the governor to the dependent, as shown in (1).

(1) Dependency structure of ‘*Taroo ga Jiroo ni ringo o ageta.*’



Dependency direction is a means of word-order typology (Liu 2010). Regarding the research targets of the present study, as Liu (2010) suggests, Chinese, English, Hungarian, Japanese, and Turkish tend to favor a head-final structure, while German and Spanish generally prefer a head-initial structure. Dependency distance refers to the linear distance between governor and dependent. This concept was initially proposed by Yngve (1960) and developed by Hudson (1995). Liu, Hudson, and Feng (2009) proposed measuring the mean dependency distance for a sentence in terms of dependency direction. The calculation procedure is as follows: subtract the position numbers of the governor and the dependent, and assume words in a sentence are assigned in a string, that is, $W_1 \dots W_i \dots W_n$. In any dependency relationship between words W_a and W_b , W_a is a governor, and W_b is its dependent. When W_a and W_b are adjacent, the distance between them is $|1|$. To put it another way, the dependency distance (DD) between the two words is $|\text{governor} - \text{dependent}|$ (the absolute value), and the MDD of the whole sentence would be

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

In this formula, n is the number of words in the sentence, and DD_i is the dependency distance of the i th dependency relationship of a sentence. Building on this, the annotation of sentence (1) was:

Sentence number	Dependent			Governor			Dependency type
	Order number	Word	POS	Order number	Word	POS	
S1	1	Taroo	NOUN	7	age	VERB	nsubj
S1	2	ga	CASE	1	Taroo	NOUN	case
S1	3	Jiroo	NOUN	7	age	VERB	obl
S1	4	ni	CASE	3	Jiroo	NOUN	case
S1	5	ringo	NOUN	7	age	VERB	obj
S1	6	o	CASE	5	ringo	NOUN	case
S1	7	ageta	VERB	0	/	/	root
S1	8	ta	TENSE	7	age	VERB	aux

Incorporating this, the MDD of sentence (1) would be 2. Theoretical and empirical studies from different camps have confirmed that dependency distance is a metric of language production and comprehension (Gibson 1998, 2000; Temperley 2007; Liu 2008; Gibson et al. 2013; Scontras et al. 2015; Rispen and De Amesti 2017; Fang and Liu 2018; Wang and Liu 2019): the longer a dependency distance is, the harder this dependency is to access (input and output). Inspired by these contributions, a general law of human language is proposed: minimising dependency distance (Liu 2008; Temperley 2007; Futrell et al. 2015). Against this background, this study wishes to determine whether there is a trend for Japanese learners with a subject-object-verb (SOV) language background to create a longer dependency distance while learners from a subject-verb-object (SVO) language background are likely to produce a shorter dependency distance. There are about 14 levels of dependency relationships in Japanese (Table 2). The following long-distance relationships will be targeted: ccomp (object subclause), advcl (adverbial clause), conj (conjunction), and acl (adjectival clause).

Table 2. Selected dependency relationships in Japanese

	Dependency relations	Grammatical structures
Clause level	nsubj	Subject of noun phrase
	ccomp	Complementizer of object clause
	obl	Oblique element
	advcl	Adverbial clause
	advmod	Adverbial modifier
	cop	Copular
	conj	Conjunction
Noun phrase level	nmod	Noun modifier
Others	acl	Adjectival clause

3. RESULTS AND DISCUSSION

3.1. The Reliability of MDD

This section examines whether MDD could index the syntactic complexity of writing quality. A Spearman rank correlation test between MDD and score of J-CAT score was conducted. A scatterplot with a regression line was plotted in Figure 1.

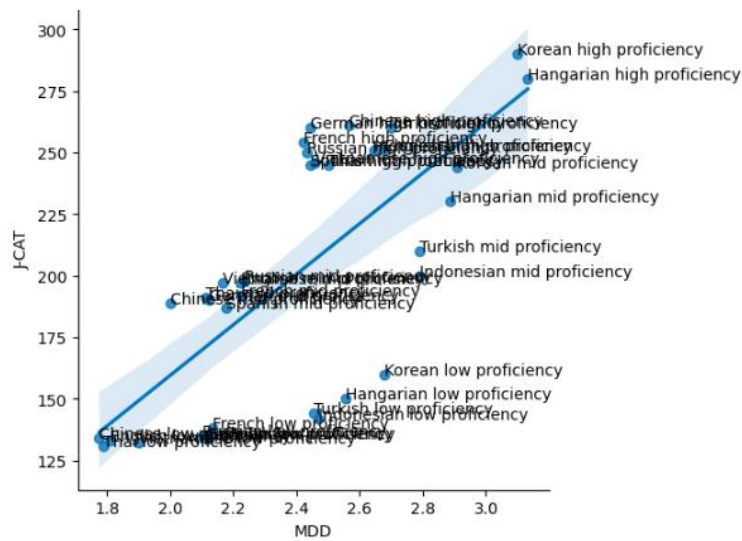


Figure 1. Spearman rank correlation coefficient between J-CAT scores and proficiency indicator MDD

As indicated, MDD and learners' acquisition proficiency well fitted the regression line. The low, middle and high proficiency groups show extremely harmonic correlations, with $\rho = 0.9992$, $p = 0.024$. Given this, we contend that MDD is efficient for indexing Japanese writing proficiency. Research question 1 was, thus, answered.

3.2. Associations of Language Distance and Acquisition Proficiency

With the metrics MDD in place, now we are in the position of exploring how L1's morphological features might affect Japanese acquisition proficiency. Figures 2 present the boxplots of MDDs of proficiency low, middle and high.

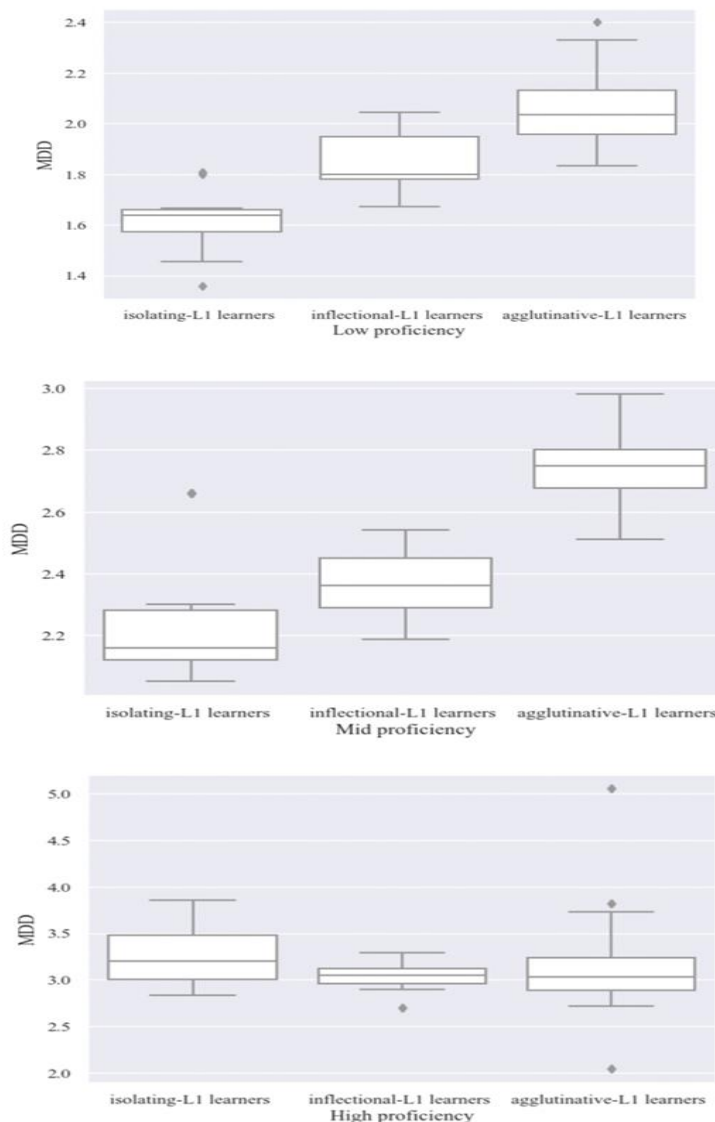


Figure 2. Boxplots of MDDs in different learning levels

Figure 2 indicates that (a). in early acquisition stage, agglutinative-L1 learners' essay (Turkish, Indonesian, Korean, Hungarian) presents a higher MDD value than inflectional-and isolating-L1 learners' essay (Chinese, Vietnamese, Thai, French, German, Russian, Spanish, English). (b). this advantage retains in middle learning level. By high learning level, learners' essay achieved similar MDD. Since MDD is an index of discriminating writing proficiency, we deduce that non-native Japanese learning quality at the early learning stage was influenced by the mother tongues. Euclidean-distance clustering based on MDD of two levels (elementary and intermediate) was carried out. The results further suggested variations within inflectional languages. The Slavic-and Romance-L1 learners produce a higher MDD than Germanic-L1 learners; Germanic-L1 learners are more close to insulating-L1 learners, as seen in low learning level, the essays written by French, Spanish, Turkish and Russian learners are clustered together; the essays written by Indonesian and Hungarian are clustered together; the essays written by Vietnamese, Thai, English, Chinese, German learners are clustered together.

(2) Euclidean-distance clustering

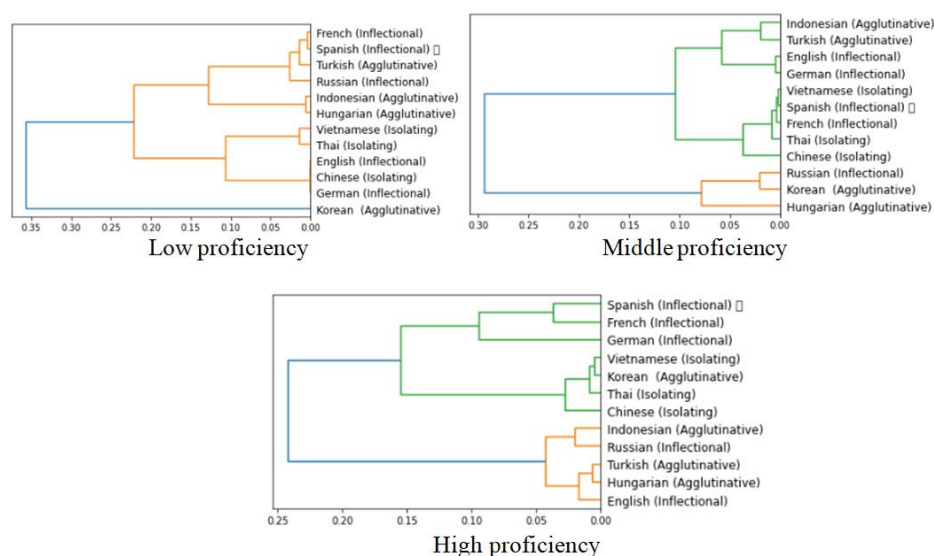


Figure 3. A cluster tree of essays written by Japanese learners from 12 L1 backgrounds based on MDD

Altmann-Fitter is further employed to fit the distribution of MDD, and it was found that the data of the isolating, agglutinative, and inflectional groups were fitted to *Right truncated Good*, as shown in Table 3.

Table 3. Fitting the *Right truncated Good* to the dependency distances of Japanese learners from different L1 backgrounds

L1	a	p	X^2	$P(X^2)$	DF	C	R^2
Isolating group	-0.6229	0.8404	0.1380	0.9869	3	0.0097	0.8462
Inflectional group	0.0013	1.0237	0.0735	1.0000	24	0.0012	0.9584
Agglutinative group	-0.0279	1.0346	0.0834	1.0000	19	0.0015	0.9778

The parameters in *Right truncated Good* are a and p . The values of the parameter p increase as the degree of agglutination of learners' mother tongue increases, that is, from isolating to inflectional to agglutinative. The MDD results further clarify that the minimizing dependency distance shown in the Japanese writings (Liu 2008; Temperley 2007; Futrell et al. 2015) reflects that of the learner's mother tongue. Research question 3 is, thus, answered.

3.3. The Probability Distribution of L3 Learning Proficiency

Previous sections have shown that MDD is a good indicator of writing proficiency. A further step forward in the relationship between MDD and frequency is attempted. A Python program is produced to fit the MDD-frequency by learners across different L1 backgrounds to the Power Law Function ($y = ae^{-bx}$). The finding suggests that apart from essays written by Hungarian Japanese learners, the MDD-frequency of essays written by the rest 11 L1-language learners fit appear to fit the power law function, with 0.7992 as the lowest value of the determination coefficient R^2 and 0.9621 as the highest ($R^2 > 0.90$, very good; $R^2 > 0.80$, good; $R^2 > 0.75$, acceptable; $R^2 < 0.75$, unacceptable). Figure 3 demonstrates the fitting outcomes of MDD-frequency.

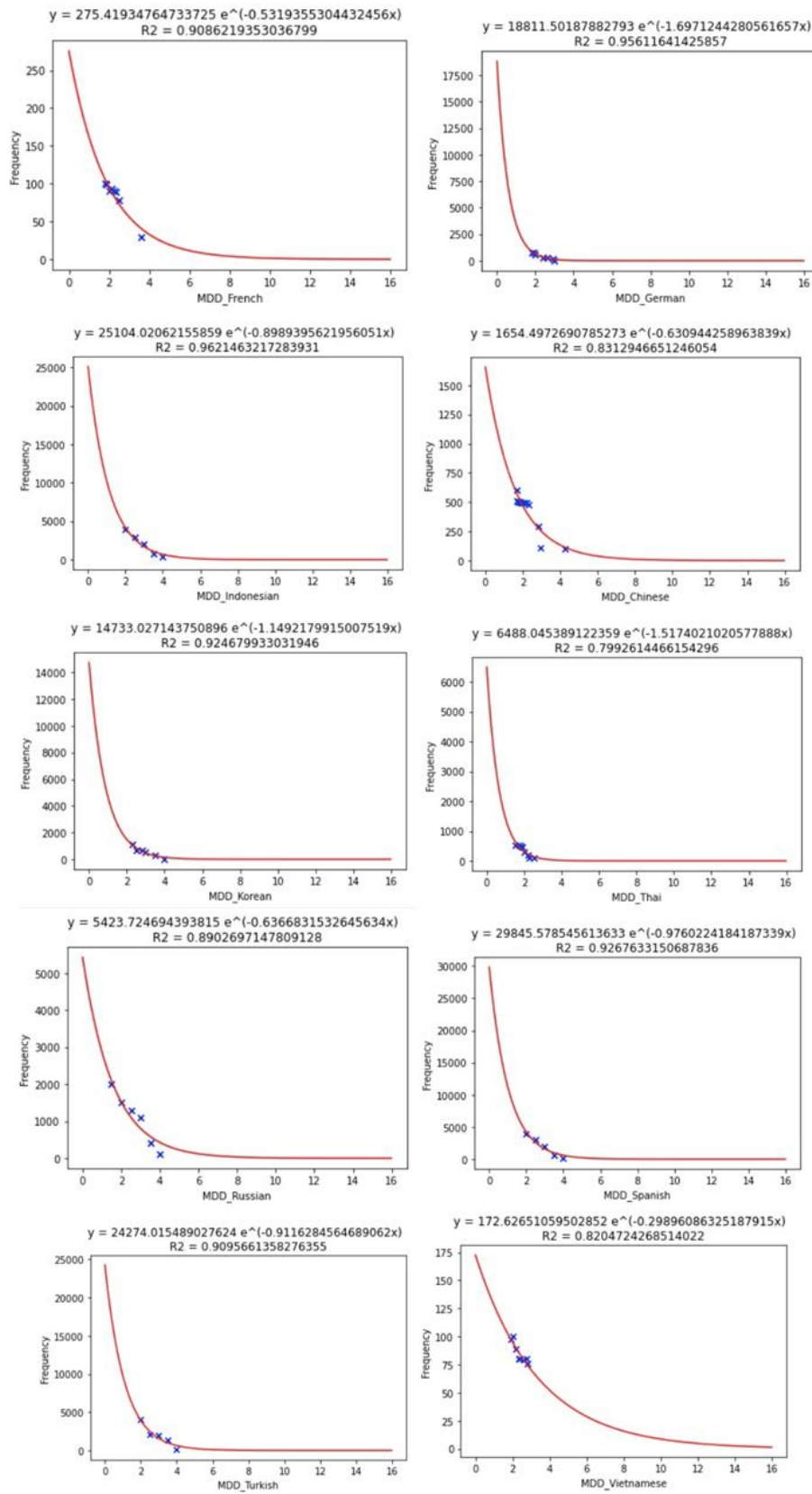


Figure 3. Fitting Power Law Function to MDD-frequency relationship of essays

Table 4 presents the variation of the parameters that contribute to the power law function. As suggested by the fitting results, the learning proficiency is predictable via ($y = ae^{-bx}$).

Table 4. Parameters that contribute to the power law function

L1	<i>a</i>	<i>b</i>	<i>R</i> ²	Fitting results
French	275.41	-0.53	0.9086	$y = 275.41e^{-0.53x}$
Vietnamese	172.62	-0.29	0.8204	$y = 172.62e^{-0.29x}$
Indonesian	251044.02	-0.89	0.9621	$y = 25104.02e^{-0.89x}$
Chinese	1654.49	-0.63	0.8312	$y = 1654.49e^{-0.63x}$
Korean	14733.02	-1.14	0.9246	$y = 14733.02e^{-1.14x}$
Thai	6488.04	-1.51	0.7992	$y = 6488.04e^{-1.51x}$
Russian	5423.72	-0.63	0.8902	$y = 5423.72e^{-0.63x}$
Spanish	29845.57	-0.97	0.9267	$y = 29845.57e^{-0.97x}$
Turkish	24274.01	-0.91	0.9095	$y = 24274.01e^{-0.91x}$
German	18811.50	-1.69	0.9561	$y = 18811.50e^{-1.69x}$

4. DISCUSSION AND CONCLUSION

This study explores how the language distance between L1 and L3 influences L3's writing proficiency. The data comprised 3410 essays produced by learners from 12 mother tongues. Mathematical linguistic approach is incorporated. Statistical analysis revealed that mean dependency distance is an efficient indicator for syntactic complexity. The finding also revealed that learners of highly agglutinative languages are likely to show higher MDD values than learners from isolated-language and fusion-language backgrounds at initial learning stages. Syntactic distance, that is, word order, also facilitates a distinction in MDD: the learners from SOV language backgrounds show longer MDD and more dependency relationship types than the learners from SVO language backgrounds. A Python program was used for fitting the probability distribution of writing proficiency in Japanese learners from different L1 backgrounds. The results indicated that the frequency and dependency distance is distributed in the Power Law function. Building on this, we employed the Altmann-Fitter software to verify the indicator in the probability distribution of L3 proficiency. Fitting *Right truncated Good* to the dependency distances of Japanese learners from different L1 backgrounds shows that the values of the parameter *p* increase as the degree of agglutination of learners' mother tongue increases. That is, L3 Japanese writing proficiency is influenced by language distance. Specifically, the typologically closer the learner's previously known languages are to the third language, the better the learner can perform in L3 acquisition. This is in accordance with the Zipf Principle of Least Effort (Zipf, 1949), which suggests that any human action aims to lighten the processing load. In this study, with the degree of agglutination of native language being higher, accessibility to the learner's knowledge of their previously known language is quicker, leading to a strong transfer from native language to subsequent language. The syntactic complexity in multi-background Japanese learners' essays highlights that no matter how diverse the learners' native and target languages are, the syntax is always constrained by universal law, minimizing dependency distance. This is in accordance with existing findings (i.e., SLA in English (Quyang et al. 2018; Lu and Liu 2016)).

While our findings could provide deeper insight into Japanese education, much remains to be explored in the future. First, the present study only examined narrative essays. To draw a more unified picture of how language distance may influence learning proficiency, other writing genres, such as emails, should be examined in more detail. Second, a further look into the role of interlanguage English in L3 Japanese acquisition is necessary (positive/negative transfer). Finally,

the speaking proficiency of learners across diverse L1 backgrounds could shed more light on the link between L1–L3 distance and acquiring proficiency.

ACKNOWLEDGEMENTS

This paper is based on work that was supported by the National Foundation of Social Sciences of China (22BYY186).

REFERENCES

- [1] Ai, H., & Lu, X. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. In A. Di Iaz-Negrillo, N. Ballier, & P. Thompson (Eds.), *Automatic treatment and analysis of learner corpus data* (pp. 249–264). Amsterdam: John Benjamins.
- [2] Archibald J. (2019). Assessing linguistic I-proximity in L3 phonology. *Workshop on Third Language Acquisition*. June. Konstanz.
- [3] Baixeries J, Elvevåg B, Ferrer-i-Cancho R. (2013). The Evolution of the exponent of Zipf's law in language ontogeny. *PLOS ONE* 8 (3).
- [4] Bardel, C., & Falk, Y. (2007). The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research*, 23(4), 459-484. L1 and L2.
- [5] Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45(1), 5-35.
- [6] Bulte, B., & Housen, A. (2012). Defining and operationalizing L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity Accuracy and Fluency in SLA* (pp. 21–46). Amsterdam: John Benjamins.
- [7] Cenoz, J. (2001). The effect of linguistic distance, L2 status and age on cross-linguistic influence in third language acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp. 8- 19). Clevedon, UK: Multilingual Matter.
- [8] Chen, H., Liang, J., Liu, H. (2015). How Does Word Length Evolve in Written Chinese? *PLOS ONE* 10 (9).
- [9] Dewaele, J. (1998). Lexical inventions: French interlanguage as L2 versus L3. *Applied Linguistics*, 19, 471-490.
- [10] Dziubalska-Kolaczyk K and Wrembel M. (2017). Natural growth model: Explaining third language phonological acquisition. Unpublished paper presented at *Societas Linguistica Europaea (SLE)*, 2017, Zürich.
- [11] Falk, Y., & Bardel, C. (2011). Object pronouns in German L3 syntax: Evidence for the L2 status factor. *Second Language Research*, 27, 59–82.
- [12] Flynn S, Foley C, and Vinnitskaya I. (2004). The cumulative-enhancement model for language acquisition: Comparing adults' and children's patterns of development in first, second and third language acquisition of relative clauses. *The International Journal of Multilingualism* 1: 3–16.
- [13] Forsyth, H. (2014). The Influence of L2 Transfer on L3 English Written Production in a Bilingual German/Italian Population: A Study of Syntactic Errors. *Open Journal of Modern Linguistics*, 4, 429-456.
- [14] Futrell R, Mahowald K, Gibson E. (2015). Large-scale evidence of dependency length minimization in 37 languages. *PNAS* 112(33): 10336-10341.
- [15] Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition*, 68, 1-76.
- [16] Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain: Papers from the first mind articulation project symposium*: 94–126. The MIT Press.
- [17] Gibson, E., Tily, H., Fedorenko, E. (2013). *The Processing Complexity of English Relative Clauses Language Down the Garden Path: The Cognitive and Biological Basis for Linguistic Structure*. Oxford University Press.
- [18] Grzybek, P. (2013). Homogeneity and heterogeneity within language(s) and text(s): Theory and practice of word length modeling. *Issues in Quantitative Linguistics Vol 3*. eds Reinhard K, Altmann G. (RAM-Verlag, Lüdenscheid). pp 66–99.

- [19] Hammarberg, B. and Hammarberg, B. (1993). Articulatory re-setting in the acquisition of new languages. Reports from the Department of Phonetics, PHONLIM 2, 61-67.
- [20] Hudson, R. (2007). Language networks: The new word grammar. Oxford: Oxford University Press.
- [21] Hunt, K. (1965). Grammatical structures written at three grade levels. NCTE Research Report No. 3. Champaign, IL, USA: NCTE.
- [22] Hunt, K. (1970). Recent measures in syntactic development. In M. Lester (Ed.), Reading in applied transformation grammar. New York.
- [23] Jiang, J., & Liu, H. (2015). The effects of sentence length on dependency distance, dependency direction and the implications-based on a parallel English–Chinese dependency treebank. *Language Sciences*, 50, 93–104.
- [24] Jiang, J., Bi, P., & Liu, H. (2019). Syntactic complexity development in the writings of EFL learners: Insights from a dependency syntactically-annotated corpus. *Journal of Second Language Writing*, 46.
- [25] Komori, S., Sugiura, M., Li, W. (2019). Examining MDD and MHD as syntactic complexity measures with intermediate Japanese learner corpus data. *Proceedings of the Fifth International Conference on Dependency Linguistics (Depling, Syntax Fest):130–135.*
- [26] Li, W. & Yan, J. (2021). Probability Distribution of Dependency Distance Based on a Treebank of Japanese EFL Learners' Interlanguage, *Journal of Quantitative Linguistics*, 28:2, 172-186.
- [27] Liu, H. (2008). Dependency distance as a metric of language comprehension difficulty. *Journal of Cognitive Science*, 9 (2), 159–191.
- [28] Liu, H. (2009). Probability distribution of dependencies based on Chinese dependency treebank. *Journal of Quantitative Linguistics*, 16(3), 256–273
- [29] Liu, H. (2010). Dependency direction as a means of word-order typology: A method based on dependency treebanks. *Lingua*, 120 (6):1567–1578.
- [30] Liu, H. 2017. An introduction to quantitative linguistics. The Commercial Press.
- [31] Liu, H., Hudson, R., Feng, Z. (2009). Using a Chinese treebank to measure dependency distance. *Corpus Linguistics and Linguistic Theory* 5 (2), 161–174.
- [32] Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, 45(1), 36–62.
- [33] Marx, N. (2002). Never quite a 'native speaker': Accent and identity in the L2 – and the L1. *Canadian Modern Language Review* 59, 264–281.
- [34] Williams, S., & Hammarberg, B. (1998). Language switches in L3 production: Implications for a polyglot speaking model. *Applied Linguistics*, 19, 295-333.
- [35] Onishi, H. (2016). The effects of L2 experience on L3 perception, *International Journal of Multilingualism*, 13:4, 459-475.
- [36] Ouyang, J., & Jiang, J. (2018). Can the probability distribution of dependency distance measure language proficiency of second language learners? *Journal of Quantitative Linguistics*, 25(4), 295–313.
- [37] Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, 18, 2030–2048.
- [38] Rothman J. (2010). On the typological economy of syntactic transfer: Word order and relative clause high/low attachment preference in L3 Brazilian Portuguese. *IRAL: International Review of Applied Linguistics in Language Teaching* 48: 245–73.
- [39] Sakoda, K. (2020). *Kaiteeban Nihongo Kyooiku ni Ikasu Dainigengoshuutokukenkyuu (Second Language Acquisition: an implication to Japanese education, revised edition)*. ALC PRESS INC.
- [40] Sikogukira, M. (1993). Influence of languages other than the L1 on a foreign language: A case of transfer from L2 to L3. *Edinburgh Working Papers in Applied Linguistics*, 4, 110-132.
- [41] Slabakova, R. (2016). *Second Language Acquisition (Oxford Core Linguistics)*. Oxford University Press.
- [42] Temperley, D. (2007). Minimisation of dependency length in written English. *Cognition* 105, 300-333.
- [43] Tesnière, L. (1959). *Éléments de syntaxe structurale*. Paris: Klincksieck.
- [44] Tremblay, M.C. (2006). Cross-linguistic influence in third language acquisition: The role of L2 proficiency and L2 exposure. *Cahiers Linguistiques d, Ottawa*, 34, 109-119.

- [45] Westergaard M, Mitrofanova N, Mykhaylyk R, and Rodina Y. (2017). Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism* 21: 666–82.
- [46] Yngve, Victor H. (1960a). A model and a hypothesis for language structure. *Proceedings of the American Philosophical Society*, 104(5): 444–466.

AUTHOR

Wenchao Li is Associate professor, Deputy director of Department of Japanese Studies, University of Zhejiang, and an executive director of Japanese Education Research Association (China). She received her M.Phil. and Ph.D. in Linguistics at the University of Tohoku (Japan). She also holds a M.St. in Japanese Studies (linguistic track) from the University of Oxford (College: Hertford). Her main research interests are mathematic linguistics. She has published 40 papers in English in international linguistic journals, and has served as a member of editorial board of *The European Journal of Theoretical and Applied Sciences*; reviewer for *Journal of International Linguistics Research* and *International Journal of English Linguistics*. Her three books were published in English by LINCOM Press (Germany).