Differences of Mean Dependency Distances of English Essays Written by Learners of Different Proficiency Levels

Masanori Oya[1*]

[1] School of Global Japanese Studies, Meiji University
[*] Corresponding author's email: masanori_oya2019@meiji.ac.jp

DOI: https://doi.org/10.53482/2022_53_400

ABSTRACT

This study investigates the differences in the mean dependency distances (MDDs) of the English essays in a learner corpus, focusing on the different proficiency levels of learners, and the different dependency types. This study is based on the following three assumptions. Firstly, the MDDs of learners' production increase as proficiency levels increase. Secondly, there is an upper limit over which MDDs do not exceed, as predicted by the Dependency Distance Minimization principle. Finally, different types of dependencies show different tendencies across learners of different proficiency levels. This study attempts to verify these assumptions with substantial learner corpus data, categorized into subcorpora according to learner proficiency. Corpus analyses yield results that support these assumptions. These results are expected to constitute a prerequisite for employing the MDD of an individual learner's production to evaluate his or her proficiency level.

Keywords: Dependency grammar, dependency distance, learner corpus, dependency distance minimization.

1 Introduction

Dependency distance is the linear distance between two words (or linguistic units) within a sentence structure, known as the dependency head and the dependent (Hudson, 1995; Liu, 2007, 2008; Tesnière, 1959). For example, in the sentence *Sarah read three articles yesterday*, the noun *Sarah* depends on the verb *read* as the subject. The noun *Sarah* is the dependent and *read* is dependency head, and the dependency distance between them is 1. The noun *articles* depends on *read* as the object, and the dependency distance between them is 2 (the direction of dependency is ignored here, and dependency distances are given as absolute values).

Dependency distance has attracted the attention of researchers in the field of quantitative linguistics, because it is assumed to function as an important measure of memory burden and syntactic complexity (Fang and Liu 2018; Futrell et al. 2015; Gibson, 1998, 2000; Gildea and Temperley, 2010; Grodner and Gibson, 2005; Li and Yang 2021; Liu 2007, 2008; Liu et al., 2017; Ouyang and Jiang, 2018; Ouyang, Jiang and Liu 2022; Oya, 2013; Yang and Li 2019; Wang and Liu 2017, inter alia).

It was discovered that dependency distance shows a certain aspect of universal properties of human languages, in that, natural languages show a certain preference for shorter dependency distances, regardless of the difference between them; this preference in natural languages is called Dependency Distance Minimization (DDM) (Ferrer-i-Cancho, 2004; Gildea and Temperley 2007; Liu 2008) For example, the Dependency Locality Theory (DLT) proposed by Gibson (2000) assumes that natural languages prefer shorter dependency distances because there is a certain limit of short-term memory in the process of integrating words into the structure of the whole sentence. Liu (2008) investigated the corpus of 20 languages and found that there was a threshold of dependency distances, which is 4. Liu (2008) compared the mean dependency distance (MDD) of natural languages and that of an artificial language in which the dependency relationships are randomized, and found that the MDD of the natural languages was shorter than that of the artificial language. In the same vein, Futrell et al. (2015) investigated multilingual corpus data with 37 natural languages and reached a similar conclusion as Liu (2008).

Previous studies have claimed, supported by various evidence, that longer dependency distances mean heavier memory load for comprehension. In the same vein, it is natural to suppose that sentences with longer dependency distances would require a heavier memory load for production. Then, learners of English with higher proficiency might produce sentences with longer dependency distances, because they can employ their knowledge about the target language with less pressure on working memory than those with lower proficiency. If this is the case, then the essays produced by learners with higher proficiency must have a longer MDD than those produced by learners with lower proficiency. In order to support this claim, we need to calculate the MDD of the production (spoken or written) data of learners with different proficiencies and compare them to test whether they are statistically different from each other.

In addition to this trend of higher MDD at higher learner proficiency, it may be possible to find a certain MDD which can be mainly found in a group of high proficiency learners and rarely in low proficiency learners. Such a figure can serve as a criterion to distinguish groups of learners in terms of their proficiency levels; if one subcorpus contains a large proportion of individual essays whose MDDs are above the criterion thus found (say, 3.5), then it is highly likely that the individual learners belonging to that subcorpus can be considered to have a high proficiency. To support this claim, we investigate whether such criteria can actually be found in learner corpora across different proficiency levels by calculating and comparing the MDDs of individual essays from each subcorpus.

As already pointed out above, Liu (2008) has found that the threshold on dependency distances, beyond which the frequency of dependency distances lowers dramatically, is 4. Though the existence of such thresholds may contradict the claim in the previous paragraph, it must be interesting to investigate whether or not that threshold is applicable to learner corpora, regardless of the difference in their proficiency levels. Such a threshold across learner data, if it exists, constitutes further evidence for the claim that the threshold on MDD might reflect a part of the universal characteristics of natural language.

The issue of the MDDs in learner data has been investigated in terms of their frequency and its fitting to a certain mathematically-defined distribution (Jiang and Liu 2015; Li and Yan 2021; Ouyang and Jiang 2018). Ouyang and Jiang (2018), for example, found that the probability distribution of dependency distance of leaners' production data shows a good fit to the right truncated modified Zipf-Alekseev distribution (Popescu et al. 2014), and the parameters of the distribution show different changes along with the proficiency levels of learners. Based on these findings, they argue that the parameters of the right truncated modified Zipf-Alekseev distribution can function as indicators of the language proficiency of learners. In a replication study (Oya 2022), a similar result was obtained from a dataset different from Ouyang and Jiang (2018).

From the viewpoint of language education, it may be valuable to know the distribution of MDDs, while the notion of parameters fluctuating with proficiency levels seems too abstract for practical use. For example, if learners of lower proficiency are found to use many short-distance dependencies, it indicates that they can only handle less complex syntactic structures; this finding can be applied to language teaching by teaching them about long-distance dependencies. In this context, the MDDs and their differences across different proficiency levels seem to be more transparent than the parameters of the right truncated modified Zipf-Alekseev distribution, and comparing and contrasting the MDDs themselves will be presented as something more valuable for practical reasons.

Along with these aspects of the MDDs of learner data, I propose a more fine-grained investigation of dependency distance in terms of different dependency types. Dependencies can be classified according to the part of speech of their head words or dependent words, or according to their functions in the sentence in which they are used. For example, in the example sentence, *Sarah has written three articles this year*, the dependency between the noun *Sarah* and the verb *written* is categorized as nominal subject, and the dependency between the verb *written* and the noun *articles* is categorized as the direct object. To the best of my knowledge, previous studies have not accounted for the difference in dependency distances in these different dependency types. In this context, it may be valuable to focus on the difference in dependency types and investigate whether MDDs of different dependency types show different increases, decreases, or no changes across proficiency levels of learners. For example, it is expected that learners of different proficiencies may show different dependency distances of different dependency types, such as those for clause embeddedness (e.g., relative clauses or that-clauses) because these dependency types are concerned with the syntactic complexities of their productions. In this study, all the dependencies are typed based on Universal Dependencies (UD) (de Marneffe et al. 2021; Zeman 2015; Zeman et al. 2020, inter alia).

Based on the background described above, this study raises and attempts to answer the following research questions:

R.Q. 1: Do the different MDDs of learner essays reflect different proficiency levels?

R.Q. 2: Is there an MDD criterion that may distinguish one group of learners from another (e.g., B2 and others, non-native learners and natives)?

R.Q. 3: Is there any threshold, or an upper limit, on the dependency distances of individual learners' essays?

R.Q. 4: Are MDDs of different dependency types different across different learner proficiency levels?

2 Material

The learner corpus used in this study is the International Corpus Network of Asian Learners of English (ICNALE) (Ishikawa 2013), an English corpus of production data (spoken monologue, spoken dialogue, written essays, and edited essays) by college-level and graduate-school-level learners from different countries and regions in Asia (China, Hong Kong, Indonesia, Japan, Korea, Pakistan, the Philippines, Singapore/ Malaysia, Taiwan, and Thailand), along with those by native speakers of English. The ICNALE contains more than 15,000 speeches and essays. Their topics are controlled. This study has used written essays from the corpus.

The production data in the ICNALE are controlled in terms of the proficiency levels of the learners; the data are categorized, according to their CEFR proficiency levels, into A2, B1_1 (B1 low), B1_2 (B1 high), and B2+. The categorization is based on the learners' scores on various English proficiency tests. Along with these learner data, the ICNALE also contains essays written by native speakers of English for comparative studies between non-native and native production data. They are categorized into ENS1 (students), ENS2 (teachers), and ENS3 (others). The detail of the ICNALE, including the relation between the CEFR levels and the test scores, is available on their web page (http://language.sakura.ne.jp/icnale/).

3 Method

3.1 Analysis 1: MDDs across different CEFR levels

All the essays in a CEFR level in the ICNALE are parsed by the Stanford Dependency Parser (de Marneffe, MacCartney, and Manning 2006). The output contains the dependency type of the words in each dependency relationship, along with the information on the order of the words in a sentence. The format of the output of the Stanford Parser for an input sentence *Sarah has written three articles this year* is shown below:

nsubj(read-3, Sarah-1)
aux(read-3, has-2)
root(root-0, written-3)
nummod(articles-5, three-4)

dobj(read-3, articles-5)
det(year-7, this-6)
advmod(read-3, year-7)

The line "nsubj(read-3, Sarah-1)" can be read as "nsubj is the dependency type of the dependency between *written*, which is the 3rd word of this sentence, and *Sarah*, which is the 1st word of this sentence." For all the other dependency types, refer to the references specified on the web page of the UD (https://universaldependencies.org/). The dependency distance between the dependency head and its dependent can be calculated from the output; for example, the dependency distance between *Sarah* and *written* above can be calculated as 2, the absolute value of extracting 3 (on "written-3" at the first line) from 1 (on "Sarah-1" at the same line). All the dependency distances can be calculated through this process. In UD, the main verb of a sentence is dependent on the *root*, which is an imaginary word representing the root of the dependency tree. In the example above, the main verb *written* depends on the root, and the dependency distance between them is 3. This study does not consider the direction of dependencies; all the dependency distances are above zero.

Based on the parsed output of the ICNALE, their MDDs are calculated by dividing the sum of all the dependency distances in the essays in a group by the number of the dependency distances in the same group. This is repeated for all the essays at all levels in the ICNALE.

This study includes the dependency distance between the main verb and the root when calculating the MDD, based on the assumption that the distance between the main verb and the root of a sentence represents its structural characteristics, such as preverbal adverbial phrases, the length of its subject noun phrase, or the number of auxiliaries before the main verb. For example, the dependency distance between the main verb and the root in the example sentence is 3, indicating that there are only two words between the main verb and the root. When the dependency distance between the main verb and the root of a sentence is longer than average, then it reflects that its syntactic structure is more complex in terms of the structural characteristics mentioned above.

3.2 Analysis 2: MDDs of each individual essay

The procedure above does not consider individual differences of learners. Therefore, in Analysis 2 we look at the MDD of each essay written by the learners. For each essay of one CEFR level of one country or region (including those written by native speakers of English) in the ICNALE, its MDD is calculated by dividing the sum of all the dependency distances in the essay by the number of the dependency distances in the same essay. This is repeated for all the essays in the same group. The same procedure is also repeated for the essays in all the CEFR levels of all the countries or regions, and the essays written by native speakers in the ICNALE.

3.3 Analysis 3: MDDs of different dependency types

For all the essays in one group in the ICNALE (such as the same CEFR level of one country or region, or those written by native speakers of English), the MDD of one dependency type is calculated by dividing the sum of all dependency distances of that type in the essays in the group by the number of the dependency distances of that type in the same group. This process is repeated for all the essays in all the levels of all countries or regions in the ICNALE, and for the following ten dependency types: *nsubj* (nominal subject; the dependency (d.) between a verb and its nominal subject), *dobj* (direct object; the d. between a verb and its direct object), *nmod* (nominal modification; the d. between a preposition or a possessive and the word which it depends on), *det* (determiner; the d. between a determiner and a noun), *amod* (adjectival modification; the d. between a noun and an adjective), *advmod* (adverbial modification; the d. between an adverb and the word which it depends on), *relcl* (relative clauses; the d. between a noun and the head verb of the relative clause), *advcl* (adverbial clauses; the d. between a verb and the head verb of an adverbial clause), *conj* (conjunct; the d. between conjuncts), and *root* (root; the d. between the main verb of a sentence and its root). Then, Mann-Whitney tests are conducted (using *jstar XR 1.1.8j*) on pairs of the groups in ICNALE, in terms of the MDDs of these dependency types. These comparisons proceed from lower-level groups to higher-level groups among non-native learner groups (A2 and B1_1, B1_1 and B1_2, and B1_2 and B2), and the highest-level group of the non-native learners (B2) and native-speaker groups (ENS1, ENS2, and ENS3).

4 Results

4.1 Analysis 1

Table 1 is the descriptive statistics of the MDDs in the essays in the ICNALE:

Table 1: The descriptive statistics of the MDDs in the ICNALE.

	Number of dependencies	Mean	SD	Median	Mode	Min.	Max
A2	103770	2.815	3.020	2	1	1	71
B1_1	275112	2.865	3.130	2	1	1	73
B1_2	269969	2.908	3.177	2	1	1	69
B2	78515	3.010	3.395	2	1	1	58
ENS1	45090	3.096	3.678	2	1	1	51
ENS2	19809	3.040	3.579	2	1	1	59
ENS3	25351	3.146	3.888	2	1	1	57

The MDDs of the essays in A2 (approx. 2.815) are the shortest among these learner groups, and it increases along with the CEFR levels; 2.865 in B1_1, 2.908 in B1_2, and 3.01 in B2. The MDDs of the essays written by native speakers of English are all above 3, which is longer than those of the essays written by English learners in ICNALE.

In order to investigate whether these differences were statistically significant, Mann-Whitney U tests were conducted, using an application for statistical tests (*js-STAR XR* 1.1.8j). 10 % of the total number of dependencies were randomly chosen from these groups because the total numbers of dependencies in these groups go beyond the limit of the application. The difference in the MDDs between A2 (Mdn=2) and B1_1 (Mdn=2) was statistically significant, $U(N=10376, N=27511,)=440415445.5$, $z=2.43$, $p=0.0151$.

Table 2: The result of the Mann-Whitney U test of the mean dependency distances of all dependency types between A2 and B1_1

A2		B1_1				
Mean rank	**n**	**Mean rank**	**n**	***U***	***z***	***p***
2.78	10376	2.88	27511	440415445.5	2.43	.0151*

** $p < .01$ * $p < .05$

The difference of the MDDs between B1_1 (Mdn = 2) and B1_2 (Mdn = 2) was statistically very significant, $U(N=27511, N=26996,)=465578377.5$, $z=3.21$, $p=0.0013$.

Table 3: The result of the Mann-Whitney U test of the mean dependency distances of all dependency types between B1_1 and B1_2

B1_1		B1_2				
Mean rank	**n**	**Mean rank**	**n**	***U***	***z***	***p***
2.88	27511	2.94	26996	465578377.5	3.21	.0013**

** $p < .01$ * $p < .05$

The difference of the MDDs between B1_2 (Mdn = 2) and B2 (Mdn = 2) was statistically significant, $U(N=26996, N=7851,)=104045243.5$, $z=2.46$, $p=0.0139$.

Table 4: The result of the Mann-Whitney U test of the mean dependency distances of all dependency types between B1_2 and B2

B1_2		B2				
Mean rank	**n**	**Mean rank**	**n**	***U***	***z***	***p***
2.94	26996	3.01	7851	104045243.5	2.46	.0139*

** $p < .01$ * $p < .05$

The difference of the MDDs between B2 (Mdn=2) and all ENS groups (Mdn = 2) was statistically not significant, $U(N=7851, N=9026,)=35299721$, $z=0.42$, $p=0.6745$.

Table 5: The result of the Mann-Whitney U test of the mean dependency distances of all dependency types between B2 and ENS

B2		ENS				
Mean rank	n	Mean rank	n	*U*	*z*	*p*
3.01	7851	3.05	9026	35299721	0.42	.6745

** $p < .01$ * $p < .05$

These results suggest that the higher the CEFR level of learners, the higher the MDD of their written production. Thus, the CEFR level of a learner may serve as an indicator of the MDD of her written production. The difference between the MDD of B2 and that of native speakers suggests that learners of advanced-intermediate levels (or higher, possibly) are closer to native speakers of English than those of lower levels, in terms of the MDD of their written production.

Additionally, the MDDs of essays written by native speakers of English are above 3, which suggests that it may serve as a threshold to distinguish non-native written production and native written production. The scenario suggested by this result would be as follows: The MDDs of learners' essays may approach 3 as their CEFR level goes up, it may go beyond 3 if they reach native-like proficiency, yet it does not exceed 4, which is the threshold found in Liu (2008).

4.2 Analysis 2

Figure 1 is the distribution of MDDs of individual essays in different CEFR levels (along with those written by native speakers of English) in the ICNALE. The box plots are aligned from the left to the right according to the levels of learners (from A2 to B2) or the types of native speakers (from ENS1 to ENS3). The y axis is MDDs. One dot represents one individual essay, and one box represents the distribution of individual essays according to their MDDs. Outliers in each level are not included in these box plots.

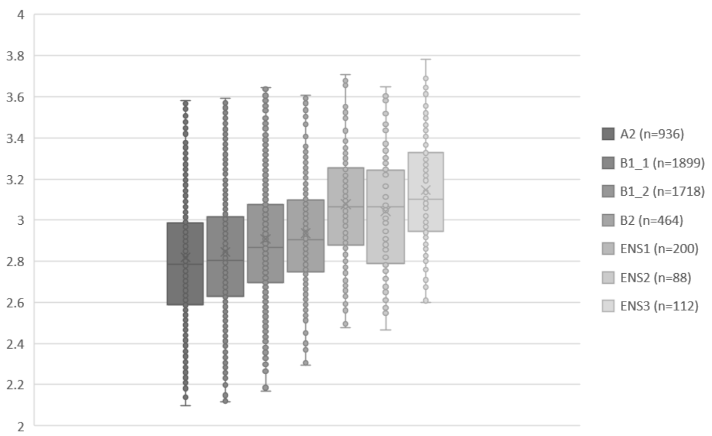

Figure 1: The distributions of MDDs of individual essays in different CEFR levels along with those written by native speakers of English in the ICNALE.

The ratio of essays with MDD over 3 increases from A2 (approx. 23%) to B2 (about 36%), and among native speakers of English, the ratios of essays with MDD over 3 exceed 50% (about 59% in ENS1, 55% in ENS2, and 63% in ENS3). Across the groups of different CEFR levels and of native speakers, the MDDs do not go over the threshold of 4, except for a small number of outliers, which are not included in the box plot above; 10 in A2, 7 in B1_1, 8 in B1_2, zero in B2, and 1 in each of the three groups of ENS.

As for the research question of whether we can use MDD to distinguish between non-native and native speakers, the result of this study seems to be promising—the MDD 3 seems to serve as a threshold to distinguish non-native written production and native written production. Along with this result, most essays in ICNALE do not go over threshold 4, in line with previous research.

It is worth mentioning the small number of outliers that go over the threshold of 4. On one hand, these outliers in lower and intermediate learners may indicate that they are producing sentences with too much effort, exceeding their memory load, resulting in producing unnatural sentences in terms of MDD. The rare instances of outliers in B2+ and ENSs, on the other hand, may indicate that they produce sentences within the limit of memory burden. The existence of such outliers can be one of the research questions in terms of individual differences of learners.

4.3 Analysis 3

Figure 2 shows the MDDs of the ten dependency types across different proficiency levels.

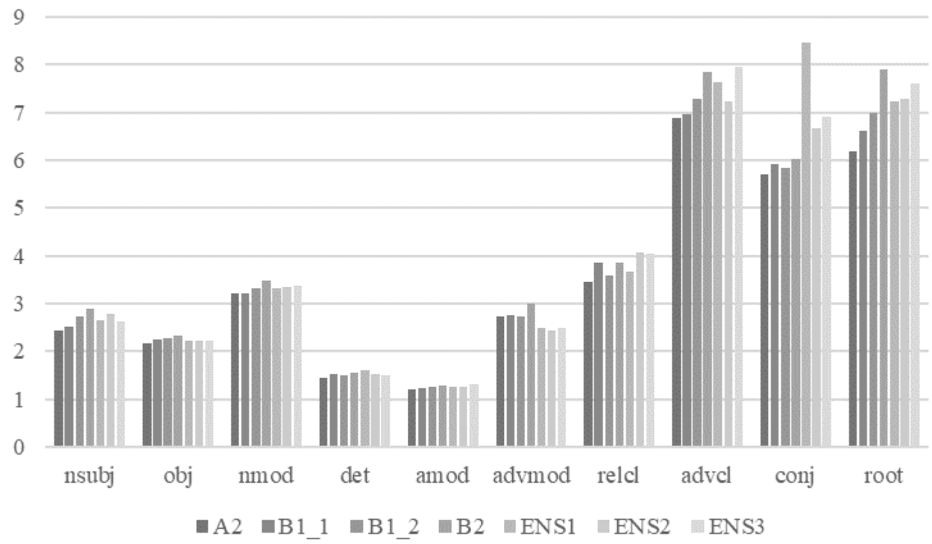

Figure 2: The MDDs of the ten dependency types across different proficiency levels

The majority of the pairs of different proficiency groups (such as A2 and B1_1) have statistically significant differences in MDDs of each dependency type; the descriptive statistics of the MDDs of these ten dependency types, and the results of the Mann-Whitney U tests between different proficiency levels, are all shown in the Appendices.

Different dependency types in learner data show different behaviors across different learner proficiencies, providing us with more detailed insight into how learners develop in terms of dependency distance. The MDDs of these dependency types increase statistically significantly from the lower-level group to the higher-level group, with some exceptions. Still, they are statistically significantly different from those of native speakers, except for the dependency type *relcl*. On the contrary, the MDDs of the dependency type *advmod* do not show a difference across different proficiency groups, yet it is statistically significantly different between B2 and all the three ENS groups. The MDDs of the type *advcl* show inconsistency; they do not show a statistically significant difference between A2 and B1_1, B2 and ENS1/ENS3. They do not support the claim that the MDDs of clause-embedding dependency types increase along with the proficiency of learners. In addition to this, the MDD of the dependency type *conj* in ENS1 is much longer than those of non-native learners, and those of other native speakers.

These facts might be due to the characteristics of these dependency types; specifically, their variation of syntactic settings. For example, the relative clause in English does not show much syntactic variation, always following the noun which it modifies, and one of the factors lengthening the dependency distance *relcl* is the word count of the subject noun when it is an object-extracting relative clause. Adverbs, however, have relatively a wide variety of syntactic settings, allowing many intervening words between the word it modifies and itself, thus lengthening its MDD. The long MDD of the dependency type *conj* in ENS1 may also reflect the wide variety of coordination in native speakers' production, which is absent in the production data of learners. It is expected that these facts can be integrated into the practice of language teaching, contributing learners to produce more naturally. Before giving definite conclusions in this respect, these results should be replicated with different learner data and other dependency types which we did not focus on in this study to further our understanding.

5 Conclusion

This study investigated the differences in the mean dependency distances (MDDs) of the English essays in a learner corpus, along with their different proficiency levels, with the assumption that the MDDs of learners' production increase from lower to higher proficiency levels. Using the ICNALE dataset, the study found that proficiency levels of learners are reflected in the MDDs of their essays. Moreover, it was found that a criterion, expressed as MDDs of individual learners' essays, can distinguish one group of learners from another, along with a threshold 4 on dependency distances of individual learners' essays. In addition to these findings, we also uncovered differences in MDDs for different dependency types across different learner proficiency levels. These findings contribute to extending our research for employing the MDD of an individual learner's production to evaluate his or her (written) production and to identify his or her proficiency level.

This study does not take into consideration the difference in the MDDs of learners from different countries or regions, which can be studied in future research. Additionally, the results suggest that the exact threshold for distinguishing non-native and native written production must be somewhere between 3 and 4, which needs to be identified through further research. More precisely, we need to investigate the profiles of learners whose production has a mean dependency distance of 3, such as their scores on TOEFL, TOEIC, and other English proficiency tests. By reversing the relationship between proficiency level (the independent variable) and the MDD (the dependent variable) of individual learners, we must ascertain whether the MDDs of individual learners (this time, the independent variable) predict their proficiency levels (the dependent variable). These topics can be studied in the future.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 20K00583.

References

De Marneffe, M., MacCartney, B., Manning, C. (2006). Generating Typed Dependency Parses from Phrase Structure Parses. In: Calzolari, N., Choukri, K., Gangemi, A., Maegaard, B., Mariani, J., Odijk, J., Tapias, D. (Eds.), *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*, pp. 449-454. Genoa: European Language Resources Association (ELRA).

De Marneffe, M., Manning, C., Nivre, J., Zeman, D. (2021). Universal Dependencies. *Computational Linguistics*, 47 (2), pp. 255-308. https://doi.org/10.1162/COLI_a_00402.

Fang, Yu., Liu, H. (2018). What factors are associated with dependency distances to ensure easy comprehension? A case study of ba sentences in mandarin Chinese. *Language Sciences*, 67, pp. 33-45. https://doi.org/10.1016/j.langsci.2018.04.005.

Futrell, R., Mahowald, K., Gibson, E. (2015). Large-scale evidence for dependency length minimization in 37 languages. *Proceedings of Natural Academy of Science,* 112(33), pp. 10336-10341. https://doi.org/10.1073/pnas.1502134112.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68(1), pp 1-76. https://doi.org/10.1016/S0010-0277(98)00034-1.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In: Alec P. Marantz, A.P, Miyashita, Y. O'Neil, W. (Eds.), *Image, language, brain: Papers from the first mind articulation project symposium*, pp. 95-126. MIT Press, Massachusetts.

Gildea, D., Temperley, D. (2010). Do grammars minimize dependency length? *Cognitive Science*, 34(2), pp. 286-310. https://doi.org/10.1111/j.1551-6709.2009.01073.x.

Grodner, D., Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science*, 29(2), pp. 261-290. https://doi.org/10.1207/s15516709cog0000_7.

Hudson, R. (1995). Measuring syntactic difficulty [Unpublished paper]. Retrieved March 30th, 2022 from https://dickhudson.com/wp-content/uploads/2013/07/Difficulty.pdf.

Ishikawa, S. (2013). The ICNALE and sophisticated contrastive interlanguage analysis of Asian learners of English. *Learner Corpus Studies in Asia and the World*, 1, pp. 91-118.

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, pp. 93-104. https://doi.org/10.1016/j.langsci.2015.04.002.

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), pp. 172-186. https://doi.org/10.1080/09296174.2020.1754611.

Liu, H. (2007). Probability distribution of dependency distance. *Glottometrics*, 15(1), pp. 1-12.

Liu, H. (2008). Dependency distance as a metric of language comprehension difficulty. *Journal of Cognitive Science*, 9(2), pp. 159-191.

Liu, H., Xu, C., Liang, J. (2017). Dependency distance: A new perspective on syntactic patterns in natural languages. *Physics of Life Reviews*, 21, pp. 171-193. https://doi.org/10.1016/j.plrev.2017.03.002.

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), pp. 295-313. https://doi.org/10.1080/09296174.2017.1373991.

Ouyang J., Jiang, J., Liu, H. (2022). Dependency distance measures in assessing L2 writing proficiency. *Assessing Writing,* 51, pp. 100-603.

Oya, M. (2013). Degree centralities, closeness centralities, and dependency distances of different genres of texts. In: *Selected papers from the 17th Conference of Pan-pacific Association of Applied Linguistics*, pp. 42-53.

Oya, M. (2022). The relevance of dependency distances in the study of L2 production. *The 41st Thailand TESOL International (Virtual) Conference*, January 22nd, 2022.

Popescu, I. I., Best, K. H., Altmann, G. (2014). *Unified modeling of length in language. Studies in Quantitative Linguistics 16*. Lüdenscheid: RAM-Verlag.

Tesnière, L. (1959). *Éléments de syntaxe structural*. Paris: Klincksieck.

Yan, H., Li, Y. (2019). Beyond length: Investigating dependency distance across L2 modalities and proficiency levels. *Open Linguistics*, 5(1), pp. 601-614. https://doi.org/10.1515/opli-2019-0033.

Wang, Y., Liu, H. (2017). The effects of genre on dependency distance and dependency direction. *Language Sciences* 59, pp. 135-147. https://doi.org/10.1016/j.langsci.2016.09.006.

Zeman, D. (2015). Slavic languages in Universal Dependencies. In: *Slovko 2015: Natural Language Processing, Corpus Linguistics, E-learning*, pp. 151-163.

Zeman, D., Nivre, J., Abrams, M., Ackermann, E., Aepli, N., Aghaei, H., Agić, Ž., Ahmadi, A., Ahrenberg, L., Ajede, C.K., et al. (2020). *Universal Dependencies 2.7, LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL)*, Faculty of Mathematics and Physics, Charles University. http://hdl.handle.net/11234/1-3424.

Appendices

Appendix 1.
The descriptive statistics of the dependency distances of the type *nsubj* in the ICNALE

nsubj	**Number of dependencies**	**Mean**	**SD**	**Median**	**Mode**	**Min.**	**Max.**
A2	13223	2.434	2.245	2	1	1	64
B1_1	33173	2.525	2.281	2	1	1	47
B1_2	30029	2.720	2.475	2	1	1	45
B2	8038	2.883	2.610	2	1	1	35
ENS1	4894	2.660	2.431	2	1	1	33
ENS2	2031	2.774	2.952	2	1	1	43
ENS3	2721	2.636	2.825	2	1	1	45

Appendix 2.
The descriptive statistics of the dependency distances of the type *dobj* in the ICNALE

dobj	**Number of dependencies**	**Mean**	**SD**	**Median**	**Mode**	**Min.**	**Max.**
A2	6582	2.167	1.269	2	2	1	32
B1_1	18761	2.239	1.233	2	2	1	31
B1_2	16635	2.265	1.406	2	2	1	42
B2	4535	2.338	1.312	2	2	1	21
ENS1	2533	2.247	1.289	2	2	1	19
ENS2	1141	2.232	1.568	2	2	1	30
ENS3	1357	2.223	1.811	2	2	1	40

Appendix 3.
The descriptive statistics of the dependency distances of the type *nmod* in the ICNALE

nmod	**Number of dependencies**	**Mean**	**SD**	**Median**	**Mode**	**Min.**	**Max.**
A2	10911	3.214	2.311	3	3	1	52
B1_1	29565	3.213	2.406	3	2	1	45
B1_2	29954	3.321	2.441	3	3	1	65
B2	8814	3.485	2.676	3	3	1	47
ENS1	4690	3.322	2.558	3	3	1	45
ENS2	2078	3.336	2.434	3	2	1	33
ENS3	2467	3.366	2.532	3	2	1	41

Appendix 4.
The descriptive statistics of the dependency distances of the type *det* in the ICNALE

det	**Number of dependencies**	**Mean**	**SD**	**Median**	**Mode**	**Min.**	**Max.**
A2	7984	1.458	0.768	1	1	1	20
B1_1	21847	1.520	0.801	1	1	1	22
B1_2	22718	1.495	0.791	1	1	1	19
B2	7033	1.551	0.805	1	1	1	17
ENS1	3736	1.615	0.802	1	1	1	8
ENS2	1565	1.522	0.821	1	1	1	10
ENS3	1927	1.488	0.721	1	1	1	9

Appendix 5.
The descriptive statistics of the dependency distances of the type *amod* in the ICNALE

amod	**Number of dependencies**	**Mean**	**SD**	**Median**	**Mode**	**Min.**	**Max.**
A2	4544	1.211	0.712	1	1	1	19
B1_1	12261	1.229	0.856	1	1	1	38
B1_2	13269	1.258	0.804	1	1	1	21
B2	4085	1.288	0.821	1	1	1	16
ENS1	2265	1.252	1.147	1	1	1	43
ENS2	1201	1.250	0.633	1	1	1	7
ENS3	1392	1.304	1.470	1	1	1	38

Appendix 6.
The descriptive statistics of the dependency distances of the type *advmod* in the ICNALE

advmod	**Number of dependencies**	**Mean**	**SD**	**Median**	**Mode**	**Min.**	**Max.**
A2	5547	2.728	2.916	2	1	1	48
B1_1	14675	2.751	3.094	2	1	1	41
B1_2	15274	2.740	3.060	2	1	1	58
B2	4540	2.985	3.656	2	1	1	41
ENS1	2821	2.481	2.971	1	1	1	33
ENS2	1249	2.440	2.928	1	1	1	42
ENS3	1714	2.487	2.960	1	1	1	33

Appendix 7.
The descriptive statistics of the dependency distances of the type *relcl* in the ICNALE

relcl	**Number of dependencies**	**Mean**	**SD**	**Median**	**Mode**	**Min.**	**Max.**
A2	1122	3.441	1.941	3	2	1	21
B1_1	698	3.848	2.419	3	2	1	27
B1_2	3009	3.601	1.982	3	2	1	19
B2	894	3.855	2.417	3	2	1	38
ENS1	528	3.676	2.132	3	2	1	25
ENS2	203	4.069	2.320	4	2	1	16
ENS3	293	4.034	2.271	3	2	1	18

Appendix 8.
The descriptive statistics of the dependency distances of the type *advcl* in the ICNALE

advcl	**Number of dependencies**	**Mean**	**SD**	**Median**	**Mode**	**Min.**	**Max.**
A2	2478	6.784	4.395	6	4	1	50
B1_1	6656	6.962	4.563	6	4	1	65
B1_2	6386	7.271	4.703	6	4	1	57
B2	1966	7.835	5.276	7	4	1	45
ENS1	1208	7.631	5.135	6	6	1	43
ENS2	563	7.236	5.196	6	3	1	37
ENS3	766	7.965	6.486	6	4	1	56

Appendix 9.
The descriptive statistics of the dependency distances of the type *conj* in the ICNALE

conj	**Number of dependencies**	**Mean**	**SD**	**Median**	**Mode**	**Min.**	**Max.**
A2	2840	5.702	5.143	4	2	1	58
B1_1	7766	5.921	5.207	4	2	1	46
B1_2	8220	5.828	5.322	4	2	1	48
B2	2305	6.026	5.612	4	2	1	48
ENS1	1577	8.471	7.476	6	2	1	51
ENS2	801	6.657	6.487	4	2	1	59
ENS3	1106	6.910	6.275	5	2	1	44

Appendix 10.
The descriptive statistics of the dependency distances of the type *root* in the ICNALE

root	**Number of dependencies**	**Mean**	**SD**	**Median**	**Mode**	**Min.**	**Max.**
A2	7416	6.197	4.904	5	2	1	71
B1_1	18127	6.619	5.289	5	3	1	68
B1_2	16300	6.993	5.457	5	4	1	69
B2	4121	7.902	6.117	6	4	1	52
ENS1	1811	7.230	6.115	5	2	1	46
ENS2	872	7.273	6.301	5	2	1	44
ENS3	872	7.608	6.839	5	2	1	47

Appendix 11.
The result of the Mann-Whitney U tests of mean dependency distances of different dependency types between A2 and B1_1

	A2		B1_1				
	Mean rank	**n**	**Mean rank**	**n**	***U***	*z*	***p***
nsubj	2.43	13223	2.53	33173	14197193	3.94	<.0001**
dobj	2.17	6582	2.24	18761	59205113	4.97	<.0001**
nmod	3.21	10911	3.21	29565	1.59E+08	2.43	.0151*
det	1.46	7984	1.52	21847	83766118	5.23	<.0001**
amod	1.21	4544	1.23	12261	27638720	0.78	.4354
advmod	2.73	5547	2.75	14675	40199775	1.35	.177
relcl	3.44	1122	3.85	698	349189	3.89	.0001**
advcl	6.87	2478	6.96	6656	8217106	0.26	.7949
conj	5.7	2840	5.92	7766	10745100	2.02	.0434*
root	6.2	7416	6.62	18127	64322647	5.41	<.0001**

** $p < .01$ * $p < .05$

Appendix 11.
The result of the Mann-Whitney U tests of mean dependency distances of different dependency types between B1_1 and B1_2

	B1_1		B1_2				
	Mean rank	**n**	**Mean rank**	**n**	***U***	*z*	***p***
nsubj	2.53	33173	2.72	30029	468643813	12.85	<.0001**
dobj	2.24	18761	2.27	16635	156007029	0.04	.9681
nmod	3.21	29565	3.32	29954	427197366	7.44	<.0001**
det	1.52	21847	1.49	22718	244566060	2.65	.008**
amod	1.23	12261	1.26	13269	79560736.5	3.03	.0024**
advmod	2.75	14675	2.74	15274	111056640	1.36	0.1738
relcl	3.85	698	3.6	3009	983469	2.62	.0088**
advcl	6.96	6656	7.27	6386	20350580	4.2	<.0001**
conj	5.92	7766	5.83	8220	31065706	2.92	.0035**
root	6.62	18127	6.99	16300	40444985	7.93	<.0001**

** $p < .01$ * $p < .05$

Appendix 12.
The result of the Mann-Whitney U tests of mean dependency distances of different dependency types between B1_2 and B2

	B1_2		B2				
	Mean rank	**n**	**Mean rank**	**n**	***U***	*z*	***p***
nsubj	2.72	30029	2.88	8038	115726746	5.67	<.0001**
dobj	2.27	16635	2.34	4635	36996496	4.21	<.0001**
nmod	3.32	29954	3.49	8814	26772001.5	5.67	<.0001**
det	1.49	22718	1.55	7033	76836559	4.85	<.0001**
amod	1.26	13269	1.29	4085	26393823.5	2.53	.0114*
advmod	2.74	15274	2.99	4540	34301125.5	1.1	.2713
relcl	3.6	3009	3.85	894	1252141	3.14	.0017**
advcl	7.27	6386	7.84	1966	5983251.5	3.15	.0016**
conj	5.83	8220	6.03	2305	9430295.5	0.34	.7339
root	6.99	16300	7.9	4121	30667467	8.63	<.0001**

** $p < .01$ * $p < .05$

Appendix 13.
The result of the Mann-Whitney U tests of mean dependency distances of different dependency types between B2 and ENS1

	B2		ENS1				
	Mean rank	**n**	**Mean rank**	**n**	***U***	***z***	***p***
nsubj	2.88	8038	2.66	4894	8683145.5	4.79	<.0001**
dobj	2.34	4635	2.25	2599	5723943	3.51	.0004**
nmod	3.49	8814	3.32	4690	19782175	4.11	<.0001**
det	1.55	7033	1.62	3736	12517403.5	4.04	.0001**
amod	1.29	4085	1.25	2265	4481805	2.06	.0394*
advmod	2.99	4540	2.48	2821	5907562.5	5.6	<.0001**
relcl	3.85	894	3.68	528	224597.5	1.53	.126
advcl	7.84	1966	7.63	1208	1162895	0.98	.3271
conj	6.03	2305	8.47	1577	1425897	11.42	<.0001**
root	7.9	4121	7.23	1811	3382567.5	5.75	<.0001**

** $p < .01$ * $p < .05$

Appendix 14.
The result of the Mann-Whitney U tests of mean dependency distances of different dependency types between B2 and ENS2

	B2		ENS2				
	Mean rank	**n**	**Mean rank**	**n**	***U***	***z***	***p***
nsubj	2.88	8038	2.77	2031	7795488	3.14	.0017**
dobj	2.34	4635	2.23	1141	2466749.5	3.52	.0004**
nmod	3.49	8814	3.34	2078	8851968.5	2.37	.0178*
det	1.55	7033	1.52	1565	5409454.5	1.06	.2891
amod	1.29	4085	1.25	1201	2413576	0.85	.3953
advmod	2.99	4540	2.44	1249	2616402	4.18	<.0001**
relcl	3.85	894	4.07	203	85754	1.22	.2225
advcl	7.84	1966	7.24	563	503755.5	3.25	.0011**
conj	6.03	2305	6.66	801	869495.5	2.45	.0143*
root	7.9	4121	7.27	872	1623061	4.49	<.0001**

** $p < .01$ * $p < .05$

Appendix 15.
The result of the Mann-Whitney U tests of mean dependency distances of different dependency types between B2 and ENS3

	B2		ENS3				
	Mean rank	**n**	**Mean rank**	**n**	***U***	***z***	***p***
nsubj	2.88	8038	2.64	2721	10086545	6.06	<.0001**
dobj	2.34	4635	2.22	1357	2854308	5.18	<.0001**
nmod	3.49	8814	3.37	2467	10574515.5	2.08	.0375*
det	1.55	7033	1.49	1927	6605265.5	1.7	.0891
amod	1.29	4085	1.3	1392	2786468.5	1.11	.267
advmod	2.99	4540	2.49	1714	3582912	4.83	<.0001**
relcl	3.85	894	4.03	293	124312	1.31	.1902
advcl	7.84	1966	7.96	766	725387	1.49	.1392
conj	6.03	2305	6.91	1106	1148279.5	4.69	<.0001**
root	7.9	4121	7.63	1034	1915408	5.03	<.0001**

** $p < .01$ * $p < .05$

Appendix 16.
The result of the Mann-Whitney U tests of mean dependency distances of different dependency types between ENS1 and ENS2

	ENS1		ENS2				
	Mean rank	**n**	**Mean rank**	**n**	***U***	***z***	***p***
nsubj	2.66	4894	2.77	2031	4948135	0.29	.7718
dobj	2.23	2647	2.23	1141	1499670.5	0.34	.7339
nmod	3.32	4690	3.34	2078	4828371.5	0.6	.5485
det	1.62	3736	1.52	1565	2728586	3.83	.0001*
amod	1.25	2265	1.25	1201	1339562	0.73	.4654
advmod	2.48	2821	2.44	1249	1760573	0.03	.9761
relcl	3.68	528	4.07	203	48137	2.13	.0332*
advcl	7.63	1208	7.24	563	314902.5	2.51	.0121*
conj	8.47	1577	6.66	801	530406	6.39	<.0001**
root	7.23	1811	7.27	872	786634.5	0.16	.8729

** $p < .01$ * $p < .05$

Appendix 17.
The result of the Mann-Whitney U tests of mean dependency distances of different dependency types between ENS1 and ENS3

	ENS1		ENS3				
	Mean rank	**n**	**Mean rank**	**n**	***U***	***z***	***p***
nsubj	2.66	4894	2.64	2721	6468165.5	2.07	.0384*
dobj	2.23	2647	2.22	1357	1739583	1.63	.0131*
nmod	3.32	4690	3.37	2467	5694069	1.1	.2713
det	1.62	3736	1.49	1927	3328488.5	4.65	<.0001**
amod	1.25	2265	1.3	1392	1558516	0.58	.5619
advmod	2.48	2821	2.49	1714	2412478.5	0.12	.9045
relcl	3.68	528	4.03	293	69794.5	2.32	.0203*
advcl	7.63	1208	7.96	766	454061	0.7	.4839
conj	8.47	1577	6.91	1106	765271	5.41	<.0001**
root	7.23	1811	7.63	1034	922426	0.66	.5092

** $p < .01$ * $p < .05$

Appendix 18.
The result of the Mann-Whitney U tests of mean dependency distances of different dependency types between ENS2 and ENS3

	ENS2		ENS3				
	Mean rank	**n**	**Mean rank**	**n**	***U***	***z***	***p***
nsubj	2.77	2031	2.64	2721	2674233	1.9	.0574
dobj	2.23	1141	2.22	1357	753032	1.18	.2380
nmod	3.34	2078	3.37	2467	2546246	0.39	.6965
det	1.52	1565	1.49	1927	1496078.5	0.4	.6892
amod	1.25	1201	1.3	1392	832732.5	0.17	.8650
advmod	2.44	1249	2.49	1714	1067363.5	0.13	.8966
relcl	4.07	203	4.03	293	29614.5	0.08	.9362
advcl	7.24	563	7.96	766	204499	1.61	.1074
conj	6.66	801	6.91	1106	424908	1.52	.1285
root	7.27	872	7.63	1034	446130	0.39	.6965

** $p < .01$ * $p < .05$