International Journal of Mathematics and Statistics Studies
Vol.10, No.4, pp.1-13, 2022
Print ISSN: 2053-2229 (Print),
Online ISSN: 2053-2210 (Online)

Syntactic complexity in Japanese dialects

Wenchao Li
Zhejiang University

Citation: Li W. (2022) Syntactic complexity in Japanese dialects, *International Journal of Mathematics and Statistics Studies*, Vol.10, No.4, pp.1-14

ABSTRACT : *This study attempted the clustering of Japanese dialects at a syntactic level in light of dependency grammar. it is Kagoshima, a southwestern area of Japan, that bears the longest mean dependency distance (1.7851) and Akita of the north-eastern prefecture that shows the shortest mean dependency distance (1.1805). A test of Euclidean distances and clustering based on the MDD brings to light that, the north-eastern and north Kinki areas tend to bear a shorter MDD, indicating a less complex syntax in spoken dialect; the Kantoo and Chuubu, south Kinki, and Shikoku areas tend to present a medium syntactic complexity; the south areas tend to present long MDDs. In a broader sense, the more northward the region, the deeper the syntactic complexity of spoken dialect. The longest DD (3.4) is detected in Gifu prefecture. This has to do with its location, between eastern and western Japan, connected to seven prefectures: Toyama, Ishikawa, Fukui, Shiga, Aichi, Mie and Nagano.*

KEYWORDS: Japanese, word length, dialect, Euclidean distance, distribution

INTRODUCTION

Since the 1960s, language complexity measurement has been intensively studied in linguistic typological works (Yngve 1960; Gibson 2000; Sampson et al. 2009; Kortmann et al. 2012; Baerman et al. 2015; Mufwene et al. 2017). Initiated by Hudson (1990), this boom in studies tackling complexity at the syntactic level raised a crucial question regarding how complexity can be measured. Liu (2008a) proposed a solution of employing dependency distance (DD) as an insightful metric. Among syntactic structures, the verb is the GOVERNOR, and all elements are connected via a 'governor-dependent' relationship. DD refers to the linear distance between a word and its governor (Tesnière 1959; Hudson 2007; Liu 2008). DD reflects (a) morphological complexity (Liu and Xu, 2011); (b) text readability (Liu 2008); (c) linguistic typology (Yan and Liu 2021); (d) language evolution (Liu and Xu 2011); (e) acquisition (Jiang, Quyang and Liu 2019); (f) cognitive difficulty in language production and comprehension (Gibson 1998, 2000; Temperley 2007); and (g) interpretation (Wang

and Liu 2019). Building on these contributions to understanding language complexity, a general law in human languages has been posited: minimizing dependency distance (Liu 2008; Temperley 2007; Futrell et al. 2015), which applies to language acquisition, interpretation, etc. This discovery aligns with Zipf's (1949) principle of least effort—which states that shorter words tend to be more frequently used. Building on these insights, this study delves into syntactic complexity by calculating the mean dependency distance (MDD) of 46 dialects in Japan, covering 42 prefectures, two urban prefectures (Osaka and Kyoto), one metropolis prefecture (Tokyo) and one circuit prefecture (Hokkaido).

Dialectology studies

The study of dialectology begins in 1880s (Boberg et al. 2018; Wolfram et al. 2006). In earlier times, the intention was to draw geographic dialect boundaries and the main methods included fieldwork and questionaries. After over a century's development, the focus has shifted to language planning, language change in a macro way, and pragmatics, gender, discourse in a micro way (Chambers 2008; Fasold 1984, 1990; Labov 2001). Since the 1980s, owing to the development of computer science, a good deal of corpora has been produced; dialectology is being tackled by computational and quantitative approaches, e.g. Labov (2004), Čepaitienė and Bakšienė (2020). The study of Japanese dialect is of particular interest for the following reasons. First, Japanese is morphologically agglutinative and syntactically a relatively word-free language. A three-valance benefactive event "Mary gives a pencil to John" can be encoded into six expressions.

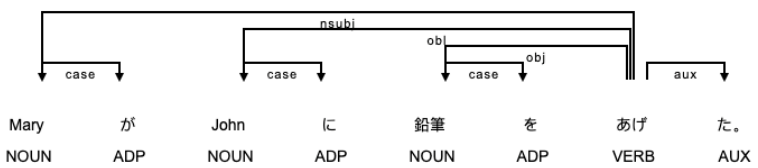
(1) a. Marry ga John ni enpitsu o **ageta**. b. Mary ga enpitsu o John ni **ageta**.

c. John ni Mary ga enpitsu o **ageta**. d. John ni enpitsu o Mary ga **ageta**.

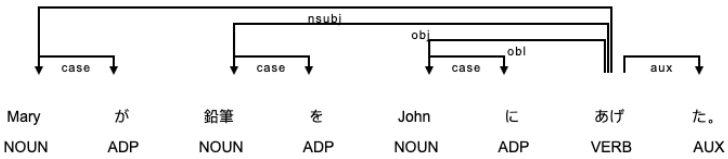
e. Enpitsu o Mary ga John ni **ageta**. f. Enpitsu o John ni Mary ga **ageta**.

The verb *ageta*, as a governor, constantly appear in the end, i.e. [$_{CP}$ XP [$_{C}$] [$_{IP}$ ...V]]. (2) presents the relationships between the governor (verb) and its dependents (subjects, direct and indirect objects).

Syntactic structure of (1a)



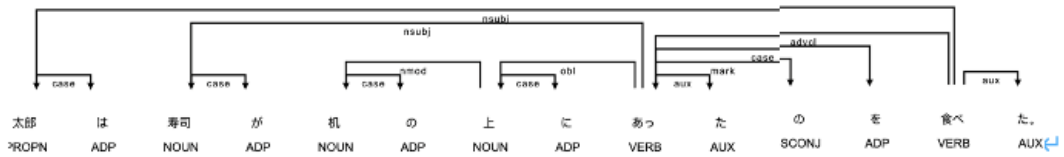
Syntactic structure of (1b)



Japanese DD has been found to be unidirectional: head final [$_{CP}$ XP [$_{C}$] [$_{IP}$ ...V]] (Liu 2010; Kanayama et al. 2018). This finding can be further supported by complex clauses: internally headed relative clause (the head noun is inside a relative clause), cf. (3) and externally headed relative clause (the head noun is outside a relative clause), cf. (4).

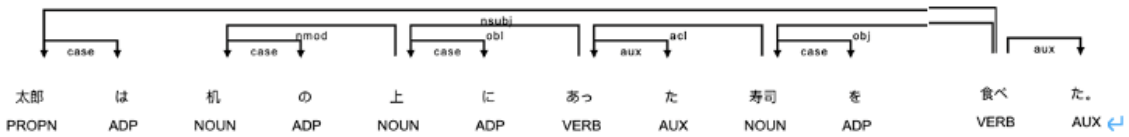
Internally headed relative clause

Taroo-wa [**sushi**-ga tsukue-no ue-ni atta]-no-o tabeta.
Taro-TOP sushi-NOM desk-GEN top-DAT was-one-ACC eat-PAST
'Taroo ate the sushi that is on the desk.'



Externally headed relative clause

Taroo-wa [tsukue-no ue-ni atta]-**sushi**-o tabeta.
Taro-TOP desk-GEN top-DAT was-sushi-ACC eat-PAST
'Taroo ate the sushi that is on the desk.'



The syntactic feature of relatively free word order in standard Japanese inspires us to ponder to what extent the various dialects in Japanese display freedom in word order. The pilot study on Japanese dialect can be traced back to 1775, Koshigaya's work titled "Monoruikoshoo," which addresses accent variations. After Meiji Restoration (1868), Tokyo dialect became the standard language. Dialectological studies at that time mainly served the purpose of expanding Tokyo dialect. For instance, Yamada (1893) made the dictionary "Nohon daijiten," where the vocabulary is glossed by the Tokyo accent. Shinbo and Tsunemi (1932) produced another dictionary that

introduces thr Tokyo accent. In the Shoowa Period (1926-1989), Japanese dialect studies shifted to geographic linguistics, with focus on accent diversities. Hattori carries out fieldwork in different areas and summarises the diversity of dialects' accents in his publication (1931). He arrives at an intriguing finding that Hiroshima and Tokyo share similar accents when it comes to two-syllable nouns. The boundary of Eastern Japanese dialect and Western Japanese dialect is the city named Kuwana in Mie prefecture. Building on this, Hattori achieved a dichotomous classification: (a) Koo rui (type Koo), which includes Kinki and Shikoku dialects; and (b) Otsu rui (type Otsu), which includes Toohoo and Sanyoodoo dialects. Another pathway comes from a diachronic linguistic perspective, e.g., Kindaichi (1937), that discovers the connections between Kyoto dialect and the Heian period (794-1185). Other remarkable works include Hirayama's (1936) investigation of the accent law in southwest of Kyuushuu area, Toojoo's (1954) classification of Japanese dialects, Inoue's (1981) distribution of age in various dialects, etc.

Existing studies seem to focus on the diversities regarding accent, delving into the syntactic aspect unequally. This study adopts the framework "dependency grammar" to tackle the syntactic complexity of dialects. With this in place, we intend to draw a classification of Japanese dialects at a syntactic level. Moreover, we wish to uncover the association between MDDs and their frequencies, and comprehend whether they follow a regularity.

In this article, Section 2 outlines the methodology (including the corpora and calculation of syntactic complexity), Section 3 addresses results and discussions, and Section 4 presents the conclusion.

METHODOLOGY

The data is drawn from Corpus of Japanese Dialects, a dynamic spoken dialect corpus including 4000 hours of recordings of dialects from 46 prefectures in Japan: Hokkaidoo, Aomori, iwate, Miyagi, Akita, Yamagata, Fukushima, Tochigi, Gunma, Saitama, Chiba, Tokyo, Kanakawa, Niigata, Toyama, Ishikawa, Fukui, Yamanashi, Nagano, Gifu, Shizuoka, Aichi, Mie, Shiga, Kyoto, Oosaka, Hyoogo, Nara, Wakayama, Tottori, Shimane, Okayama, Hiroshima, Yamaguchi, Tokushima, Kakkawa, Ehime, Koochi, Fukuoka, Saga, Nagasaki, Kumamoto, Ooita, Miyazaki, Kagoshima, Okinawa. The spoken data are transcripted into texts saved in csv files. To calculate the syntactic complexity, the following procedures were carried out:

Step 1: Draw raw data from the corpora
Step 2: Save the csv data into visual studio code
Step 2: Parse each sentence via the GiNZA v4 Parser
Step 3: Produce a computer programme to calculate the dynamic MDD and the distribution of dependency relationship types from the parsed outputs
Step 4: Produce a computer program to Euclidean distance clustering of the diverse dialects

This study's calculation of DD is aligned with Liu, Hudson, and Feng's (2009) insights. The words in a sentence are assigned in a string (i.e., W_1 ...W_i ...W_n). Regarding any dependency relationship between words W_a and W_b, if W_a is a governor and W_b is its dependent, then the DD between the two words is |governor − dependent| (the absolute value). The MDD of the whole sentence would then be:

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

RESULTS AND DISCUSSION

Drawing on the methodology highlighted above, this section proceeds by analysing the syntactic complexity in modern Japanese. Mean dependency distance is employed as the measuring metric. This section is organised as follows. Section 3.1 considers the dependency distance of each dialect and carries out a clustering analysis. Section 3.2 looks into the probability distribution of dependency distance, searching for the parameters that may indicate the variety of dialects.

MDDs of diverse spoken dialects

Figure 1 and Table 2 show the MDDs of the spoken data of the 46 dialects.

Figure 1. MDDs of Japanese dialects

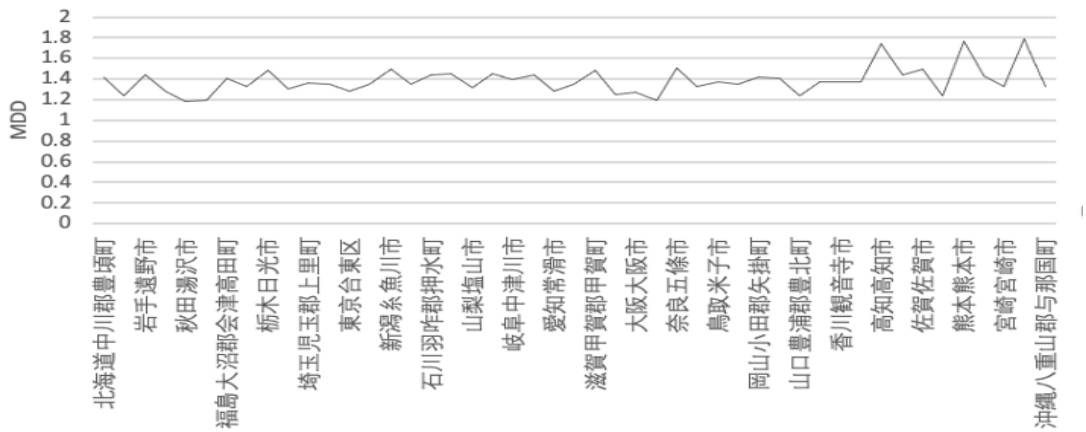


Table 2. MDD of Japanese dialects

Dialects	MDD	Dialects	MDD	Dialects	MDD
Akita	1.1805	Okinawa	1.3309	Ooita	1.4318
Yamagata	1.1924	Chiba	1.3472	Ishikawa	1.4394
Hyoogo	1.1961	Mie	1.3507	Shizuoka	1.4395
Yamaguchi	1.2347	Toyama	1.3529	Fukuoka	1.4421
Aomori	1.2377	Shimane	1.3538	Iwate	1.4447
Nagasaki	1.2415	Kanagawa	1.3547	Fukui	1.4522
Kyooto	1.2485	Saitama	1.3598	Nagano	1.4531
Oosaka	1.2781	Tokushima	1.3714	Tochigi	1.4805
Miyagi	1.2863	Ehime	1.3716	Shiga	1.4903
Tokyo	1.2882	Kagawa	1.3767	Niigata	1.4955
Aichi	1.289	Tottori	1.3769	Saga	1.5019
Gunma	1.3107	Gifu	1.3946	Nara	1.5064
Yamanashi	1.3154	Hiroshima	1.4099	Koochi	1.7440
Miyazaki	1.324	Fukushima	1.4126	Kumamoto	1.7633
Ibaraki	1.3292	Okayama	1.4183	Kagoshima	1.7851
Wakayama	1.3292	Hokkaidoo	1.4198		

The average DD of all dialects is 1.3849, shorter than workplace conversation (2.15) and everyday conversation (1.85)[1]. All spoken dialects have the minimum MDD (0.5) and maximum MDD (2.6), except Gifu, which has one token of MDD: 3.4, and Niigata, which has one token of MDD: 3.1. Gifu dialect bears the longest MDD owing to its location, i.e., between eastern and western Japan, connected to seven prefectures: Toyama, Ishikawa, Fukui, Shiga, Aichi, Mie and Nagano.

[1] The MDDs of everyday conversation and workplace conversation are obtained via the everyday conversation corpus and workplace conversation corpus produced by the National Institute for Japanese Language and Linguistics.

Figure 2. The location of Gifu prefecture

Kindaichi (2010) argues that Gifu dialect is among the inner cycle of Tokyo dialect. In our finding, which is based upon syntactic complexity, no evidence comes to support this idea. Hattori (1931) discovers that Hiroshima and Tokyo share similar accents in two-mora lexicons. In our syntactic data, no similarity between the two areas has been detected.

It is also explicated that, the more southward the region, the longer the MDD turns, i.e., the north-eastern prefecture Akita shows the shortest MDD (1.1805) and the southwestern city Kagoshima presents the longest MDD (1.7851). The Kantoo area (Gunma and Ibaraki), north-eastern area (Fukushima and Miyagi) bear a relatively shorter MDD, with the maximum MDD no bigger than 1.9. On the other hand, Kinki, Chuugoku and Shikoku areas seem to incorporate MDDs that are lightly longer, but do not extend to 2. All these strands support the minimizing dependency distance (Liu 2008; Temperley 2007; Futrell et al. 2015) and Zipf's (1949) principle of least effort.

Clustering of dialects

Hattori (1931) points out that in terms of accent, the boundary of Eastern and Western dialects is Kuwana city (Mie prefecture). A question of where the boundary for the syntax among the dialects is, arises immediately. A test of Euclidean distances and clustering based on the MDD as a unit of syntactic complexity measurement was conducted. The finding is presented in Figure 3.

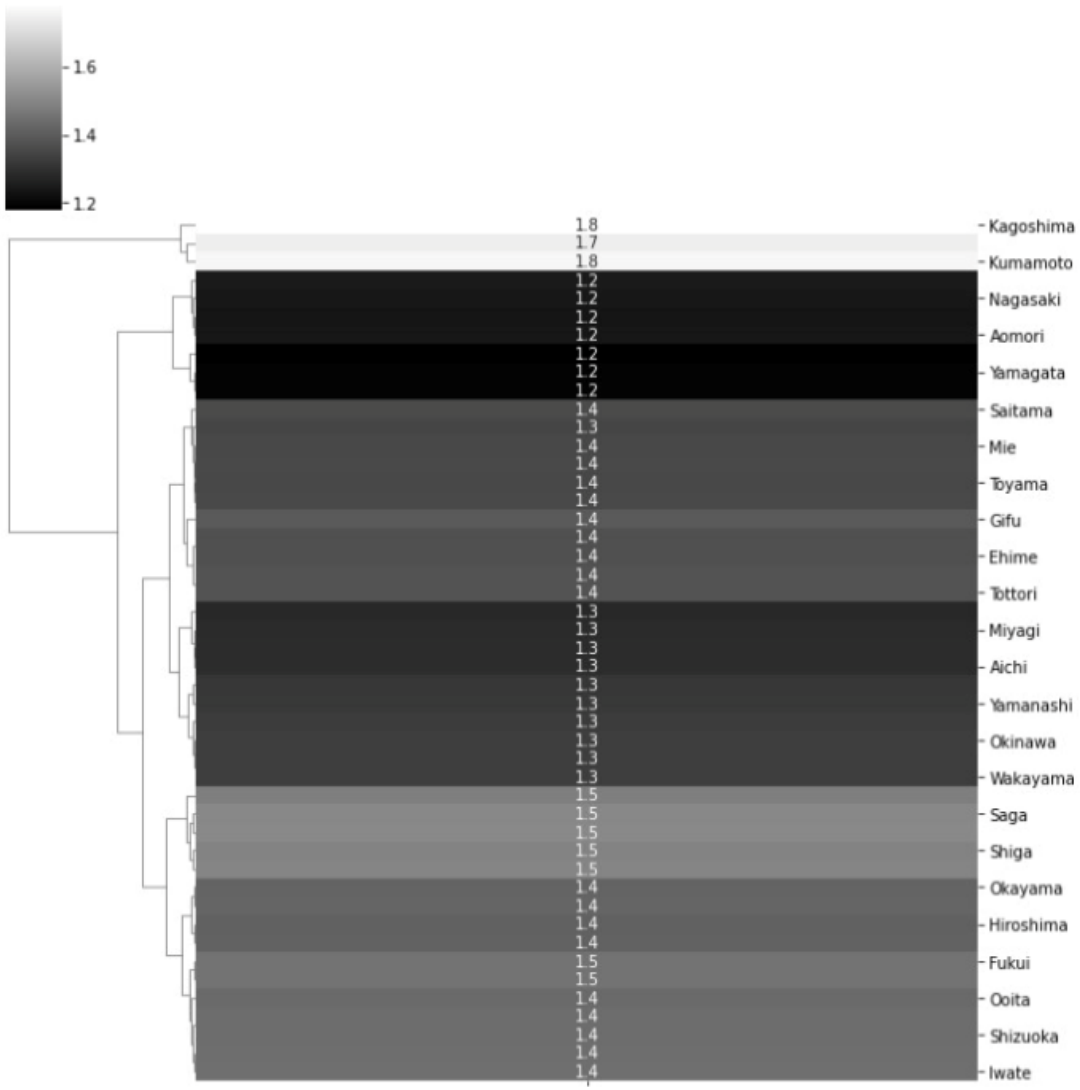


Figure 3. Euclidean distances and clustering based on the MDD

As the clustering tree suggests, the north-eastern areas (Akita, Yamagata, Aomori, Miyagi), the north Kinki areas (Hyoogo, Kyooto, Oosaka), Yamaguchi, Nagasaki, Tokyo, Aichi are among the short-MDD group; most Kantoo and Chuubu areas (Gunma, Kanagawa, Chiba, Ibaraki, Saitama, Yamanashi); the south Kinki areas (Wakayama, Okinawa, Mie); the Shikoku areas (Ehime, Kagawa, Tokushima), Chuugoku areas (Shimane, Tottori), Toyama, Gifu and Miyazaki fall into the middle-MDD group; Hiroshima, Fukushima, Okayama, Hokkaidoo, Ooita, Ishikawa, Shizuoka, Fukuoka, Iwate, Fukui, Nagano, Tochigi, Shiga, Niigata, Saga, Nara, Koochi, Kumamoto, Kagoshima are clustered into the long-MDD group.

Classification of Japanese dialects based on syntactic complexity

a. **Short-MDD group**: Akita, Yamagata, Aomori, Miyagi, Hyoogo, Yamaguchi, Ngasaki, Kyooto, Oosaka, Tokyo, Aichi.

b. **Middle-MDD group**: Gunma, Yamanashi, Miyazaki, Ibaraki, Wakayama, Okinawa, Chiba, Mie, Toyama, Shimane, Kanagawa, Saitama, Tokushima, Ehime, Kagawa, Tottori, Gifu.

c. **Long-MDD group**: Hiroshima, Fukushima, Okayama, Hokkaidoo, Ooita, Ishikawa, Shizuoka, Fukuoka, Iwate, Fukui, Nagano, Tochigi, Shiga, Niigata, Saga, Nara, Koochi, Kumamoto, Kagoshima.

Probability distribution of dialects

We further aim to explore the parameters that may indicate the syntactic complexity of the dialect. To this end, a Python program is produced to fit the power law function ($y=ax^b$). The fittings of the data bring to light that MDD and frequency can be demonstrated by the power law function, with 0.8967 as the lowest value of the determination coefficient R^2 and 1.0000 as the highest ($R^2 > 0.90$, very good; $R^2 > 0.80$, good; $R^2 > 0.75$, acceptable; $R^2 < 0.75$, unacceptable). The finding is summarised in Table 3.

Table 3. Fitting the Power Law function to MDD of different Japanese dialects

Dialects	**MDD**	***a***	***b***	**R^2**	**Fitting results**
Akita	1.1805	27	-37.47	0.9970	$y=27x^{-37.47}$
Yamagata	1.1924	30	-40.02	0.9898	$y=30x^{-40.02}$
Hyoogo	1.1961	30	-38.25	0.9756	$y=30x^{-38.25}$
Yamaguchi	1.2347	20	-2.79	0.9999	$y=20x^{-2.79}$
Aomori	1.2377	33	-39.02	0.9381	$y=33x^{-39.02}$
Nagasaki	1.2415	17	-4.08	1.0	$y=17x^{-4.08}$
Kyooto	1.2485	42	-44.01	0.9809	$y=42x^{-44.01}$
Oosaka	1.2781	28	-1058	1.0	$y=28x^{-1058}$
Miyagi	1.2863	31	-49.01	0.9675	$y=31x^{-49.01}$
Tokyo	1.2882	40	-50.90	0.9756	$y=40x^{-50.90}$
Aichi	1.2890	38	-41.09	0.9715	$y=38x^{-41.09}$
Gunma	1.3107	32	-1058.96	1.0	$y=32x^{-1058.96}$
Yamanashi	1.3154	39	-53.01	0.9941	$y=39x^{-53.01}$

Miyazaki	1.3240	30	-40.02	0.9898	$y = 30x^{-40.02}$
Ibaraki	1.3292	38	-1058	1.0	$y = 38x^{-1058}$
Wakayama	1.3292	45	-1058	1.0	$y = 45x^{-1058}$
Okinawa	1.3309	23	-2.52	1.0	$y = 23x^{-2.52}$
Chiba	1.3472	27	-1058.96	1.0	$y = 27x^{-1058.96}$
Mie	1.3507	33	-44.35	0.9917	$y = 33x^{-44.35}$
Toyama	1.3529	37	-41.24	0.8967	$y = 37x^{-41.24}$
Shimane	1.3538	32	-1058	1.0	$y = 32x^{-1058}$
Kanagawa	1.3547	48	-1058	1.0	$y = 48x^{-1058}$
Saitama	1.3598	38	-53.34	0.9985	$y = 38x^{-53.34}$
Tokushima	1.3714	26	-4.70	1.0	$y = 26x^{-4.70}$
Ehime	1.3716	58	-78.64	0.9940	$y = 58x^{-78.64}$
Kagawa	1.3767	33	-46.13	0.9980	$y = 33x^{-46.13}$
Tottori	1.3769	26	-3.70	1.0	$y = 26x^{-3.70}$
Gifu	1.3946	45	-44.40	0.9919	$y = 45x^{-44.40}$
Hiroshima	1.4099	32	-41.13	0.9788	$y = 32x^{-41.13}$
Fukushima	1.4126	39	-1058.96	1.0	$y = 39x^{-31058.96}$
Okayama	1.4183	29	-3.27	1.0	$y = 29x^{-3.27}$
Hokkaidoo	1.4198	15	-2.90	1.0	$y = 15x^{-2.90}$
Ooita	1.4318	23	-3.09	0.9258	$y = 23x^{-3.09}$
Ishikawa	1.4394	37	-44.79	0.9523	$y = 37x^{-44.79}$
Shizuoka	1.4395	26	-2.37	1.0	$y = 26x^{-2.37}$
Fukuoka	1.4421	31	-36.14	0.9286	$y = 31x^{-36.14}$
Iwate	1.4447	31	-36.14	0.9286	$y = 31x^{-36.14}$
Fukui	1.4522	25	-3.64	1.0	$y = 275x^{-3.64}$
Nagano	1.4531	34	-44.02	0.9814	$y = 34x^{-44.02}$
Tochigi	1.4805	35	-41.91	0.9459	$y = 35x^{-41.91}$
Shiga	1.4903	34	-38.69	0.9121	$y = 34x^{-38.69}$
Niigata	1.4955	25	-3.57	0.9993	$y = 25x^{-3.57}$
Saga	1.5019	18	2.16	1.0	$y = 18x^{-2.16}$
Nara	1.5064	26	-5.32	0.9913	$y = 26x^{-5.32}$
Koochi	1.7440	29	-4.85	1.0	$y = 29x^{-4.85}$
Kumamoto	1.7633	36	-48.68	0.9931	$y = 36x^{-48.68}$
Kagoshima	1.7851	17	-1.76	1.0	$y = 17x^{-1.76}$

CONCLUSION

This study examined the syntactic complexity of Japanese dialects in light of dependency grammar, with the findings supporting the following notions. The average DD of spoken dialects across 46 areas in Japan is 1.3849, shorter than workplace and everyday conversations. This indicates that spoken dialects are aligned with Zipf's principle of least effort. The longest DD (3.4) is detected in Gifu prefecture. This perhaps has to do with its location, between eastern and western Japan, connected to seven prefectures: Toyama, Ishikawa, Fukui, Shiga, Aichi, Mie and Nagano. Having said this, it is Kagoshima, a southwestern area of Japan, that bears the longest mean dependency distance (1.7851) and Akita of the north-eastern prefecture that shows the shortest mean dependency distance (1.1805). A test of Euclidean distances and clustering based on the MDD brings to light that, from a syntactic complexity point of view, the north-eastern and north Kinki areas tend to bear a shorter MDD, indicating a less complex syntax in spoken dialect; the Kantoo and Chuubu, south Kinki, and Shikoku areas tend to present a medium syntactic complexity; the south areas tend to present long MDDs. In a broader sense, the more northward the region, the deeper the syntactic complexity of spoken dialect. Finally, the MDD–frequency relationships of all dialects fit to power law function ($y=ax^b$). This study has attempted the clustering of Japanese dialects at a syntactic level. There is room for further exploration; for instance, lexical complexity measured by word length. Another paper will be devoted to this exploration.

References

Agnė, Čepaitienė and Rima, Bakšienė. 2020. A Quantitative Analysis of the Eastern Aukštaitian Subdialects of the Lithuanian Language. *Scando Slavica*. 66(1):1-37.

Baerman, Matthew, Dunstan Brown, and Greville G. Corbett, eds. 2015. *Understanding and measuring morphological complexity*. Oxford: Oxford Univ. Press.

Boberg, Charles, John Nerbonne and Dominic Watt (eds). 2018. The Handbook of Dialectology. Hoboken NJ, USA: Wiley Blackwell.

Fasold, Ralph W. 1984. *The sociolinguistics of society*. Oxford and New York: Blackwell.

Fasold, Ralph W. 1990. *The sociolinguistics of language*. Cambridge, MA: Blackwell.

Futrell1 R, Mahowald K, Gibson E. 2015. Large-scale evidence of dependency lengthminimization in 37 languages. *PNAS* 112(33): 10336-10341.

Gibson, E. 1998. Linguistic complexity: locality of syntactic dependencies. *Cognition*, 68, 1-76.

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.

Hattori, Shiroo. 1933. Accent and dialect. [Akusento to hoogen]. Meijishoin.

Hirayama, Teruo. 1936. The accent of south Kyuushuu 1, 2. [Minamikyuushuu akusento no kenkyuu]. Dialectology [Hoogen] 6 (4, 5).

Hudson, Richard. 2007. *Language networks: The new word grammar*. Oxford: OxfordUniversity Press.

Hudson, Richard. 1990. *English Word Grammar*. Oxford: Blackwell.

Kindaichi, Haruhiko. 1937. A comparision of Japanese dialects and the accent in Heian Period: a focus upon two-syllable noun. [Gendai shohoogen no hikaku kara mita Heianjidai no akusento: tokuni nionsetsumeeshi ni tsukite].

Kortmann, Bernd, and Benedikt Szmrecsanyi, eds. 2012. *Linguistic complexity. Second language acquisition, indigenization, contact*. Berlin and Boston: De Gruyter.

Labov, William 2001. *Principles of Linguistic Change*. Volume 2: Social factors. Oxford: Blackwell.

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

Liu, Haitao and Xu, Chunshan. 2011. Can syntactic networks indicate morphological complexity of a language. *Europhysics Letters*, 93 (2).

Mufwene, Salikoko S., Christophe Coupé, and Francois Pellegrino, eds. 2017. *Complexity in language: Developmental and evolutionary perspectives*. Cambridge, UK: Cambridge Univ. Press.

Yan Jianwei and Liu, Haitao. 2021. Morphology and word order in Slavic languages: Insights from annotated cor- pora. *Voprosy Jazykoznanija*, 4: 131–159.

Yamada, Bimyoo. 1893. Japanese Dictionary [Nihon daijiten]. Suuzando. Sampson, Geoffrey, David Gil, and Peter Trudgill, eds. 2009. *Language complexity as an evolving variable*. Oxford: Oxford Univ. Press.

Shinbo, Kaku and Tsunemi, Senri. 1932. Japanese pronunciation dictionary [Kokugo hatsu on akuento jiten]. Kooseekaku.

Temperley, D. 2007. Minimization of dependency length in written English. *Cognition* 105, 300-333.

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

Toojoo, Misao. 1954. Josetsu [Preface]. Japanese dialect [Nihon Hoogengaku]. Yoshikawa koobunkan.

Wang, Yawen and Liu, Haitao. 2019. The Effects of Source Languages on Syntactic Structures of Target Languages in the Simultaneous Interpretation: A Quantitative Investigation Based on Dependency Syntactic Treebanks. *Glottometrics* 45: 89-113.

Wolfram, W. and N. Schilling-Estes. 2006. American English. Dialects and Variation. Second Edition. Oxford: Blackwell.

Yngve, V. 1960. A model and a hypothesis for language structure. *Proceedings of the American Philosophical Society*, 104, 444-466.

Zipf, G. K. 1949. *Human Behaviour and the Principle of Least Effort: An Introduction to Human Ecology*. Addison–Wesley, Cambridge, MA.