# Degree centrality as a measure of robustness of dependency structures of the sentences in a large-scale learner corpus of English

**Masanori Oya**
School of Global Japanese Studies,
Meiji University
masanori_oya2019@meiji.ac.jp

## Abstract

This paper examines the differences in the robustness of syntactic dependency structures in written English produced by learners of varying proficiency levels and by native English speakers. The robustness of these dependency structures is represented by their degree centralities, and corpus-based investigation revealed that learners with higher proficiency levels tend to produce sentences with lower degree centralities. This means that they produce more robust, and more embedded sentences. It is also revealed that the sentences produced by native speakers of English tend to produce more embedded sentences than non-native speakers.

## 1 Introduction

The aim of this paper is to examine the differences in the robustness of syntactic dependency structures in written English produced by learners of varying proficiency levels and by native English speakers.

Structural properties of sentences have been explored in the field of second language acquisition (SLA) using a variety of metrics such as word per sentence or type-token ratio with the cover term of *sentence complexity* (e.g., Bardovi-Harlig 1992, Brown 1973, Ellis and Yuan 2005, Hunt 1965, Michel et al. 2007, Norris and Ortega 2009, Ortega 2003, Robinson 2007, Scarborough 1990, Scott 1988, Skehan and Foster, 2005, Wolf-Quintero et al. 1998). The basic tenet behind them is that the proficiency levels of learners can be represented by these metrics. In other words, it is expected that these metrics increase in proportion to the advancement of learners' proficiency levels. For example, Wolfe-Quintero et al. (1998) pointed out that depth of clauses in the sentences produced by leaners of English increases in proportion to their proficiency levels, hence depth of clauses can function as a measure of sentence complexity.

Sentence complexity should not be regarded as a single independent variable, but as a dependent variable that can be represented by multiple variables (depth of clauses is one of them). In this context, it is essential to address these variables related to sentence complexity individually rather than treating them collectively and indiscriminately. By focusing on each factor in turn, we can understand the structural characteristics of the sentences produced by speakers/writers with certain attributes (e.g., native/non-native, beginners/intermediate/advanced, non-native with different backgrounds) more objectively.

This paper introduces the robustness of dependency structures as one of these variables related to sentence complexity. Specifically, I adopt *degree centrality* of the dependency structure of a sentence as a metric to measure its robustness. By modeling the dependency structures of English sentences in a corpus—organized by learners' proficiency levels—as graphs, I compute their degree centralities and investigate whether the distribution of these values reflects the learners' proficiency levels.

This paper is organized as follows: Section 2 summarizes the idea of dependency structures as graphs and their degree centralities, and explains the relationship between degree centralities of dependency structures and their robustness. In Section 3, previous studies are briefly reviewed to point out their drawbacks. Section 4 describes this study of degree centralities of English sentences in a large-scale learner corpus, which is followed by discussions in Section 5. Section 6 concludes this paper.



*Proceedings of the Third Workshop on Quantitative Syntax (QUASY, SyntaxFest 2025)*, pages 9–16

## 2 Dependency trees as graphs

In network analysis, a graph consists of a collection of nodes and a set of edges linking these nodes (Freeman, 1978; Wasserman & Faust, 1994). In this context, the degree of a node is determined by the number of edges connected to it. Previous research (Oya, 2010, 2013, and 2014) has posited that the dependency tree (or structure) of a sentence can be conceptualized as a graph. More specifically, within a dependency tree, words function as nodes, while their dependency relationships are represented as edges, and the degree of a word is the number of other words depending on it and the word which it depends on. For example, an English sentence "I have written this article" has the dependency structure in the format of *Universal Dependencies* (de Marneffe et al. 2021, Zeman et al. 2017) in Figure 1. The degree of the word "written" is four, because it depends on "root," and three words depend on it.

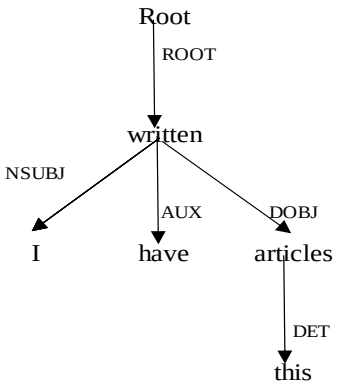


Figure 1: The dependency structure of the sentence "I have written this article."

Graph theory establishes a variety of metrics that quantify the structural characteristics of graphs. If the dependency tree of a sentence is considered a specific type of graph, these metrics can be utilized to analyze its structural properties systematically. This approach enables a more objective and scientifically rigorous examination of its structure, as opposed to relying solely on intuitive interpretations. Based on this premise, Oya (2010) applied *degree centrality* (Freeman 1978; Wasserman & Faust 1994) as a metric to assess the complexity of dependency trees in English sentences (yet the use of the word "complexity" is rather problematic; discussed later).

Degree centrality is a type of index that indicates the significance of a given node within a specific graph. The degree centrality of a graph $C_D$ which contains $g$ nodes is calculated by the following formula (Freeman 1978, Wasserman & Faust 1994):

$$C_D = \frac{\sum_{i=1}^{g}[C_D(n^*) - C_D(n_i)]}{max\sum_{i=1}^{g}[C_D(n^*) - C_D(n_i)]} \quad (1)$$

$C_D(n^*)$ is the largest degree in the given graph, and $C_D(n_i)$ is the degree of a node. The enumerator represents the sum of the largest degree minus the degrees of all the other nodes. The denominator represents the maximal possible sum of the largest degree minus the degrees of all the other nodes. For a graph which contains $g$ nodes, the largest possible degree of its node is $g$-1.

In principle, degree centrality ranges from 0 to 1.

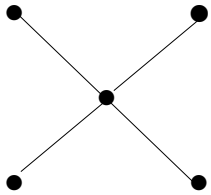

Figure 2: A star graph.

If a graph has a degree centrality of 1, this signifies that a single node within the graph is connected to all other nodes, forming a *star graph*, as illustrated in Figure 2.

The largest degree in the star graph in Figure 2 is 4, which is the largest possible degree of a graph which contains 5 nodes. The degree of all the other nodes is 1. The enumerator and the denominator are the same, as indicated by the following formula, hence the degree centrality of the star graph is 1:

$$C_D = \frac{(4-1)+(4-1)+(4-4)+(4-1)+(4-1)}{(4-1)+(4-1)+(4-4)+(4-1)+(4-1)} = 1 \quad (2)$$

If the graph representing the dependency tree of a sentence has a degree centrality of one, this indicates that a single word serves as the dependency head for all other words in the sentence. In other words, the dependency structure of the sentence is entirely flat.

Degree centrality values decrease as the structure of graphs becomes more linear, meaning that no single node holds greater significance than the others. In Figure 3, the five nodes are arranged in a linear sequence (forming a *line graph*).

The largest degree in the line graph in Figure 3 is 2, and 3 of its nodes have that degree. The other 2 nodes have the degree 1. Hence, the enumerator is 2. The denominator is the same as that of the star graph in Figure 2. Therefore, as indicated by the formula (3), the degree centrality of the line graph in Figure 3 is approximately 0.1667:

$$C_D = \frac{(2-1)+(2-2)+(2-2)+(2-2)+(2-1)}{(4-1)+(4-1)+(4-4)+(4-1)+(4-1)} = \frac{2}{12} \approx 0.1667 \quad (3)$$

The degree centrality of the example sentence "I have written this article" is calculated as follows: its dependency structure contains 6 nodes (including Root). Its largest observed degree is 4 (with *written*). One node has the degree 2 (with *article*), and all the other nodes have the degree one. The largest possible degree of a graph with 6 nodes is 5. Therefore, as indicated by the formula (4), the degree centrality of the dependency structure of the example sentence is 0.7:

$$C_D = \frac{(4-1)+(4-4)+(4-2)+(4-1)+(4-1)+(4-1)}{(5-1)+(5-1)+(5-5)+(5-1)+(5-1)+(5-1)} = \frac{14}{20} = 0.7 \quad (4)$$

If the graph representing the dependency tree of a sentence exhibits a low degree centrality, this indicates that one word depends on another, which in turn depends on yet another, and so forth, resulting in a more embedded dependency structure.

The degree centrality of a network (and a dependency structure of a sentence) is concerned with its robustness. Sentences with larger degree centralities (with flatter dependency structure) contain a certain core word (or words) on which many of the other words in the same sentence depend. If the core word is deleted or overlooked, then the whole structure falls apart into unrelated words, and fails to be interpreted appropriately. On the other hand, sentences with lower degree centralities (with more embedded dependency structure) have no such core, or more than one core, and therefore, even if one of the words is deleted, there will be some fragments of structure which can be interpreted, though not completely. Hence, sentences with lower degree centralities are more robust than those with higher degree centralities.

It should be noted that the robustness of dependency structures as discussed here differs conceptually from syntactic robustness.

## 3 Previous studies

Figure 3: A line graph.

Some previous studies assume that the degree centrality values of sentences within a corpus can function as an indicator of their syntactic complexity, in which the word "complexity" is used as something represented by degree centrality. Oya (2010) observed that the degree centrality values of English essays written by Japanese learners tend to be higher than those of academic journal abstracts, suggesting that the former exhibit flatter and less embedded syntactic structures compared to the latter. Oya (2013) conducted corpus-based research of degree centrality as a syntactic complexity measure. He revealed that sentences in different genres show different distributions of degree centralities, more specifically, sentences in fictions tend to have higher degree centralities than those in journals, meaning that the former have flatter syntactic structure than the latter. Oya (2014) applied the idea of using degree centrality as a syntactic complexity measure into Japanese, based on an English-Japanese small-scale parallel corpus, and it is found that Japanese sentences tend to have higher degree centralities than their English translations, meaning that Japanese sentences are flatter than their English translations.

The previous studies on the degree centralities of sentences contain the following two drawbacks: First, it is assumed that the degree centrality of a sentence can be used as a measure of its complexity without explicit explanation on why it can be. It is certain that structures with lower degree centralities are more robust, and it is found that the robustness is one of the characteristics of complex systems (e.g., Artime, Grassia, De Domenico et al. 2024), yet it is not certain that more robust sentences are more complex. These previous

studies should have used the word “robustness” instead of “complexity” of syntactic structure. Second, there has been no study of degree centrality as a measure of structural robustness of English sentences generated by learners of English as a second language (L2) at different proficiency levels, let alone comparing and contrasting the degree centralities of English sentences generated by non-native speakers of English (non-ENS) and those generated by native speakers of English (ENS). In this context, this study is the first attempt to examine whether the degree centralities of sentences generated by non-ENS in different proficiency levels show distributions which are different across these different proficiency levels, and from those by ENS. If any difference between them is found, that will give us a new insight into the difference between non-ENS and ENS in terms of the robustness of the sentences they generate, based on the theoretical background of graph theory.

## 4 This study

The research question of this study is as follows:
(1) Do degree centralities of the sentences generated by non-ENS at different proficiency levels show different distributions across these levels?
(2) Do degree centralities of the sentences generated by non-ENS show distributions which are different from those generated by ENS?

### 4.1 Data

The production data examined in this study are the written essay section of the International Corpus Network of Asian Learners of English (ICNALE; Ishikawa 2013, 2023), a learner corpus of English, production data of English from college-level students with a variety of backgrounds across Asia (China, Hong Kong, Indonesia, Japan, Korea, Pakistan, the Philippines, Singapore/Malaysia, Taiwan, and Thailand), along with production data from native speakers of English. In the written essay section of ICNALE, the topics of the essays are as follows:

Topic A: College students need to have a part-time job.

Topic B: Smoking should be completely banned at all restaurants in the country.

A notable characteristic of the ICNALE is its systematic classification of production data based on learners’ proficiency levels, as defined by the Common European Framework of Reference for Languages (CEFR). These levels include A2, B1_1 (B1 low), B1_2 (B1 high), and B2+. In the written essay section of the ICNALE, each learner is assigned a proficiency level according to their scores on various English proficiency tests, and their essays are then categorized accordingly within these CEFR levels.

### 4.2 Procedure

The average degree centrality of the sentences of the essays of Topic A and Topic B in each CEFR category in the ICNALE is calculated by a Python script which was coded by the 1st author, then these average degree centralities are compared across these CEFR categories. Also, the distributions of degree centralities of individual sentences are compared across these CEFR categories, in terms of the percentage of the degree centralities falling within particular subranges of the interval of 0.1. Since degree centralities fall within the range from 0 to 1, it is further divided into those subranges. If it is found that the degree centralities of sentences produced by learners who are categorized into one particular CEFR category, say, B1_1, fall within a certain subrange, such as that from 0.2 to 0.3, significantly more often than other subranges, then it indicates the structural characteristics of the sentences produced by learners of that CEFR category, suggesting that learners who belong to the CEFR category B1_1 tend to produce sentences whose degree centralities often fall within that subrange.

This procedure was conducted for each Topic individually, so that we can examine whether there is any difference of degree centralities due to the difference of topics: Topic A is related to college life, and therefore it must be more familiar to the learners than Topic B, which is related to one of the social issues. As Oya (2013) pointed out that sentences of different genres show different distributions of degree centralities, it is expected that the difference of topics in the ICNALE would result in different distributions of degree centralities.

### 4.3 Results

| | N | ADC | SD |
|---|---|---|---|
| A2 | 7287 | 0.36 | 0.19 |
| B1_1 | 14369 | 0.34 | 0.17 |
| B1_2 | 12967 | 0.31 | 0.17 |
| B2 | 6244 | 0.3 | 0.15 |
| ENS | 1779 | 0.23 | 0.14 |

Table 1: The average degree centralities of the sentences in the essays about Topic A (part-time job). ADC: average degree centralities

| | N | ADC | SD |
|---|---|---|---|
| A2 | 7460 | 0.38 | 0.2 |
| B1_1 | 14678 | 0.37 | 0.18 |
| B1_2 | 13440 | 0.34 | 0.17 |
| B2 | 3249 | 0.33 | 0.17 |
| ENS | 1981 | 0.27 | 0.16 |

Table 2: The average degree centralities of the sentences in the essays about Topic B (Ban on smoking). ADC: average degree centralities

Table 1 shows the average degree centralities of the sentences in the essays about Topic A (part-time job), and Table 2 shows those in the essays about Topic B (smoking ban on local restaurants):

In both groups, the average degree centralities decrease from A2 at the largest among them to the ENS at the lowest.

A one-way between subjects ANOVA was conducted to compare the average degree centralities across the categories for each topic group. For Topic A, there was a significant effect of categories on average degree centralities at the $p<.01$ level [F(4, 4264) = 276.87]. Post hoc comparisons using the Tukey HSD test indicated that the mean scores for A2 (M = 0.36, SD = 0.19), B1_1 (M = 0.34, SD = 0.17), B1_2 (M = 0.31, SD = 0.17), B2 (M = 0.3, SD = 0.15) and ENS (M = 0.23, SD = 0.14) are all different from each other. For Topic B, there was also a significant effect of categories on average degree centralities at the $p<.01$ level [F(4, 4080) = 211.32]. Post hoc comparisons using the Tukey HSD test indicated that the mean scores for A2 (M = 0.33, SD = 0.2), B1_1 (M = 0.37, SD = 0.18), B1_2 (M = 0.34, SD = 0.17), B2 (M = 0.33, SD = 0.17) and ENS (M = 0.27, SD = 0.16) are all different from each other, except for the pair of B1_2 and B2. These results suggest that degree centralities of sentences on average decrease in negative proportion to the proficiency of the learners, and yet they are still larger than those produced by ENSs. The scenario can be summarized roughly as follows: Leaners at lower proficiency levels write English sentences which contain flatter structure (with larger degree centralities), and as their proficiency level gets higher, they come to produce sentences with more embedded structure (with smaller degree centralities).

It is also interesting to note that the average degree centrality of Topic A is smaller than that of Topic B regardless of the CEFR category. This may

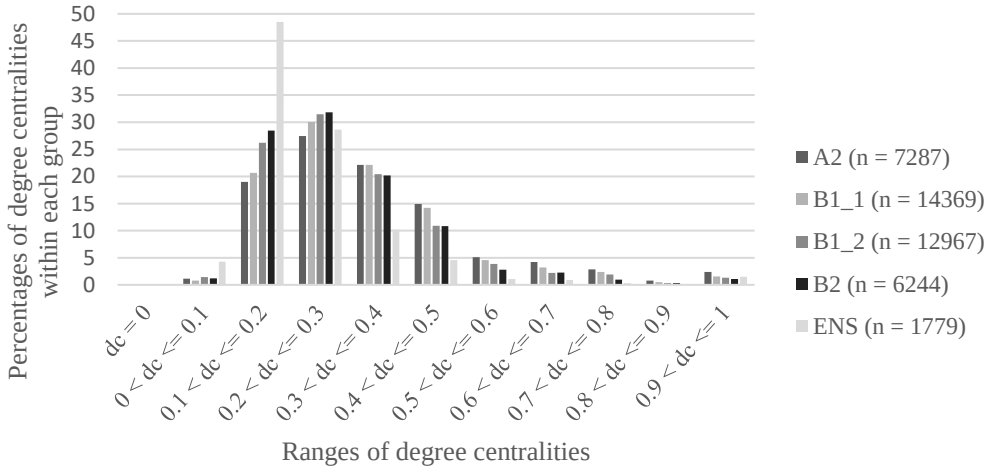


Figure 3: Percentages of degree centralities of the sentences in each group

(Topic A: Part-time job)

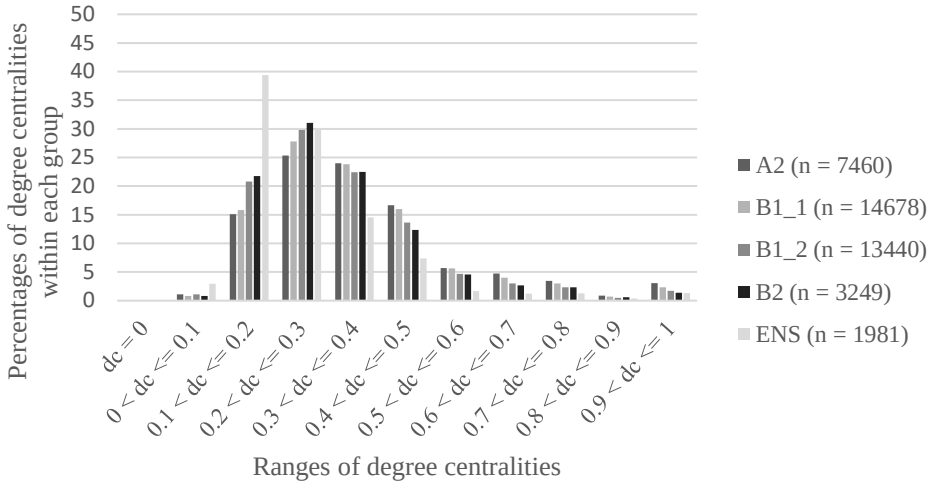


Figure 4: Percentages of degree centralities of the sentences in each group

(Topic B: Ban on smoking)

suggest that the topic of an essay influences the degree centralities of its sentences. As far as Topic A and Topic B in the ICNALE are concerned, Topic A is related to their own life and therefore they may find it easier to describe their own experiences and opinions, adding into their sentences more phrases describing the details, resulting in more embedded sentences with smaller degree centralities. Topic B, on the other hand, needs to establish their own argument on a topic which they may be less familiar with, hence their description stays simple and this can be reflected on less embedded sentences.

Figure 3 and 4 show the distributions of degree centralities of individual sentences in all the CEFR categories, represented by the percentages of degree centralities which fall within each subrange. About 50% of the sentences the degree centralities of the category ENS for Topic A (part-time job) fall within the range of degree centralities which are 0.2 or smaller and larger than 0.1, while in the same range the degree centralities of the sentences generated by non-ENS groups are less than 30% of those in each group. Similar distributions of degree centralities can be found for Topic B (Ban on smoking); About 40% of the sentences the degree centralities of the category ENS for this group fall within the range of degree centralities which are 0.2 or smaller and larger than 0.1, while in the same range the degree centralities of the sentences generated by non-ENS groups are less than 25% of those in each group.

## 5. Discussion

The results described above seem to answer the research questions positively: First, in the ICNALE, average degree centralities of the sentences generated by non-ENSs at different proficiency levels show different distributions across these levels, decreasing from lower levels to higher levels. Second, average degree centralities of the sentences generated by non-ENS show distributions which are different from those generated by ENSs. These findings must be put into the context of linguistic description along with the explanation on why it is the case. There are at least three questions to be addressed: (1) Why do non-ENSs at lower proficiency levels tend to produce sentences with flatter structure? (2) Why do they come to produce more embedded sentences as they become more proficient? And (3) Why do ENSs produce more embedded sentences than non-ENSs? If degree centralities of sentences (or their robustness) on average decrease as the proficiency of learners gets higher, it can be explained that it gets more robust than before. These issues need to be addressed in future research, for better understanding of degree centralities of sentence structure, from the viewpoint of (1) investigating the relationship between degree centralities and other sentence complexity measures, such as type-token ratio, word per sentence, and (2) formulating the theory which explains how and why the robustness of sentences increases along with the development of learners' proficiency.

An anonymous reviewer noted that these issues have been addressed in previous research from the perspective of sentence length. Ferrer-i-Cancho and Gómez-Rodríguez (2019) argue that shorter sentences often conflict with the principle of dependency distance minimization (DDM) (cf. Liu, 2008; Futrell et al., 2015). According to DDM,

language users tend to prefer shorter dependency distances, as longer distances increase the cognitive burden on working memory. Ferrer-i-Cancho and Gómez-Rodríguez (2019) suggest that the apparent violation of DDM in short sentences arises from their characteristically flat, star-like dependency structures. A plausible scenario can thus be outlined: beginner-level language learners, who typically produce shorter sentences, tend to generate structures with higher degree centralities. As their proficiency improves, they begin to construct longer, more syntactically complex sentences, which in turn exhibit less star-like configurations—potentially as a strategy to mitigate the cognitive load imposed by long dependency distances within such structures. This hypothesis warrants further empirical investigation in future research.

## 6. Conclusion

This paper examined the differences in the robustness of syntactic dependency structures in written English produced by learners of varying proficiency levels and by native English speakers. The robustness of these dependency structures is represented by their degree centralities, and corpus-based investigation revealed that learners with higher proficiency levels tend to produce sentences with lower degree centralities, meaning they produce more robust, and more embedded sentences, yet the sentences produced by native speakers of English tend to produce more embedded sentences than non-native speakers. The results of this study lead us to further exploration of degree centralities of dependency structures as a measure of their robustness.

## References


Oriol Artime, Marco Grassia, Manlio De Domenico, James P. Gleeson, Hernán A. Makse, Giuseppe Mangioni, Matjaž Perc and Filippo Radicchi. 2024. Robustness and resilience of complex networks. *National Review of Physics* 6, 114–131. https://doi.org/10.1038/s42254-023-00676-y

Kathleen Bardovi-Harlig. 1992. A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly*, 26, 390-395.

Roger Brown. 1973. *A First Language: The Early Stages*. Harvard University Press.

Marie-Catherine de Marneffe, Christopher D. Manning, Joakim Nivre, Daniel Zeman. 2021. Universal Dependencies. *Computational Linguistics*, 47(2).

Rod Ellis and Fanguan Yuan. 2005. 'The effects of careful within-task planning on oral and written task performance' in R. Ellis (ed.): *Planning and Task Performance in a Second Language*, 167-192. John Benjamins.

Ramon Ferrer-i-Cancho and Carlos Gómez-Rodríguez. 2019. Anti dependency distance minimization in short sequences. A graph theoretic approach. *Journal of Quantitative Linguistics*, 28(1), 50–76. https://doi.org/10.1080/09296174.2019.1645547

Linton Freeman. 1978. Centrality in social networks. *Social Networks* vol.1, 215-239.

Richard Futrell, Kyle Mahowald, and Edward Gibson. Large-scale evidence of dependency length minimization in 37 languages. *The Proceedings of the National Academy of Sciences (PNAS)* 112 (33), 10336-10341. https://doi.org/10.1073/pnas.1502134112

Kellog W. Hunt. 1965. Grammatical structures written at three grade levels. *NCTE Research Report* No. 3. Champaign, IL, USA: NCTE.

Shin-ichiro Ishikawa. 2013. The ICNALE and sophisticated contrastive interlanguage analysis of Asian learners of English. *Learner corpus studies in Asia and the world*, 1, 91-118.

Shin-ichiro Ishikawa. 2023. *The ICNALE Guide: An Introduction to a Learner Corpus Study on Asian Learners' L2 English*. Routledge.

Haitao Liu. 2008. Dependency distance as a metric of language comprehension difficulty. *Journal of Cognitive Science*, 9(2), 159-191. http://dx.doi.org/10.17791/jcs.2008.9.2.159

Marije C. Michel, Folkert Kuiken and Ineke Vedder. 2007. The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics in Language Teaching*, 45, 241-259.

John M. Norris and Lourdes Ortega. 2009. Towards an organic approach to investigating CAF in instructed SLA: The case of complexity, *Applied Linguistics* 30(4), 555-578, https://doi.org/10.1093/applin/amp044

Lourdes Ortega. 2003. Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics* 24(4), 492–518, https://doi.org/10.1093/applin/24.4.492

Masanori Oya. 2010. Directed acyclic graph representation of grammatical knowledge and its application for calculating sentence complexity. *Proceedings of the 15th International Conference of Pan-Pacific Association of Applied Linguistics*, 393-400.

Masanori Oya. 2013. Degree centralities, closeness centralities, and dependency distances of different genres of texts. *Selected Papers from the 17th Conference of Pan-Pacific Applied Linguistics*, 42-53.

Masanori Oya. 2014. *A Study of Syntactic Typed-Dependency Trees for English and Japanese and Graph-centrality Measures*. Ph.D. dissertation, Waseda University.

Peter Robinson. 2007. Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics*, 45, 237-257.

Hollis S. Scarborough 1990. Index of productive syntax. *Applied Psycholinguistics* 11, 1-22.

Cherry M. Scott. 1988. Spoken and written syntax. In M. Nippold (ed.): *Later Language Development: Ages Nine through Nineteen*. Little, Brown.

Peter Skehan and Pauline Foster. 2005. Strategic and on-line planning: The influence of surprise information and task time on second language performance. In R. Ellis (ed.): *Planning and Task Performance in a Second Language*. John Benjamins.

Stanley Wasserman and Katherine Faust. 1994. *Social Network Analysis*. Cambridge: Cambridge University Press.

Kate Wolfe-Quintero, Shunji Inagaki, and Hae-Yong Kim. 1998. *Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity*. University of Hawaii Press.

Zeman, D., Popel, M., Straka, M., Hajič, J., Nivre, J., Ginter, F., Luotolahti, J., Pyysalo, S., Petrov, S., Potthast, M., Tyers, F., Badmaeva, E., Gökırmak, M., Nedoluzhko, A., Cinková, S., Hajič jr., J., Hlaváčová, J., Kettnerová, V., Urešová, Z., Kanerva, J., Ojala, S., Missilä, A., Manning, C., Schuster, S., Reddy, S., Taji, D., Habash, N., Leung, H., de Marneffe, M.C., Sanguinetti, M., Simi, M., Kanayama, H., de Paiva, V., Droganova, K., Alonso, H.M., Çöltekin, Ç., Sulubacak, U., Uszkoreit, H., Macketanz, V., Burchardt, A., Harris, K., Marheinecke, K., Rehm, G., Kayadelen, T., Attia, M., Elkahky, A., Yu, Z., Pitler, E., Lertpradit, S., Mandl, M., Kirchner, J., Alcalde, H.F., Strnadová, J., Banerjee, E., Manurung, R., Stella, A., Shimada, A., Kwak, S., Mendonça, G., Lando, T., Nitisaroj, R., and Li, J. (2017). CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies. *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, 1–19, Vancouver.