

Multifractal Analysis of the Distribution of Three Grammatical Constructions in English Texts

Rosmawati & Wander Lowie

To cite this article: Rosmawati & Wander Lowie (2024) Multifractal Analysis of the Distribution of Three Grammatical Constructions in English Texts, *Journal of Quantitative Linguistics*, 31:1, 19-42, DOI: [10.1080/09296174.2024.2302674](https://doi.org/10.1080/09296174.2024.2302674)

To link to this article: <https://doi.org/10.1080/09296174.2024.2302674>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 08 Feb 2024.



Submit your article to this journal [↗](#)



Article views: 636



View related articles [↗](#)



View Crossmark data [↗](#)

Multifractal Analysis of the Distribution of Three Grammatical Constructions in English Texts

Rosmawati ^a and Wander Lowie ^b

^aCentre for Professional Communication, Singapore Institute of Technology, Singapore;

^bCenter for Language and Cognition / Applied Linguistics, University of Groningen, Groningen, the Netherlands

ABSTRACT

Both the Menzerath-Altmann law and the Zipf-Mandelbrot law note that language is a fractal structure and, like any other fractals, follows power laws. Studies on fractal linguistics demonstrated that the relation between the scaling measures and the number of constituents in language indeed follow the power law probability distribution. However, most of these studies looked at languages from the structural perspective only, often ignoring the grammatical aspects of language structures. In this study, we ventured into English grammar and used a multifractal analysis to explore the nature of multifractality in three grammatical structures in English texts – i.e. the Finite Verb Phrases, Noun Phrases, and Head Nouns. In this paper, we present the evidence of multifractality in the distribution of these constructs and discuss how the parameters of multifractality align with the current understanding of register variation in different text types.

ARTICLE HISTORY Received 06 May 2023; Revised 02 January 2024; Accepted 03 January 2024

Introduction

What do cauliflowers, snowflakes, and the coast of Britain have in common? Not much, except that they are all fractal structures. Fractal is a mathematical concept that originally referred to geometric figures that can be split, infinitely, into self-similar yet reduced-size copies of the original, such as the Sierpinski triangle and the Koch snowflakes.¹ This definition, however, has been extended to include non-geometric figures and abstract constructs that are self-affine statistically, such as the coastlines of Britain and language constructs. The complexity of a fractal is described by a measure called the Fractal Dimension (FD, or in some cases, just D), which is operationalized as the negative ratio between $\log N$ and $\log \varepsilon$ (where N is the number of copies

CONTACT Wander Lowie  w.m.lowie@rug.nl

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

in each split, and ε is the scaling ratio). Such a log–log relationship is also referred to as the power law relationship and is generally considered as evidence of fractality, including fractality in language constructs.

Recent studies in fractal linguistics, such as Ausloos (2012) and Chatzigeorgiou et al. (2017), highlighted the multifractal property of language. The current paper builds on the concept of multifractality and explores the multifractal characteristics of the distribution of three grammatical constructs in English (i.e. Finite Verb Phrase, Noun Phrase, and Head Noun). To do this, we will employ the multifractal analysis to test our hypothesis that the distribution of these constructs manifests on the characteristics of multifractality. The outcomes of this study will advance our understanding of the nature of multifractality in language and open a channel for future discussions on the possibilities for universality or comparability among fractal structures across languages. It is hoped that this study will constitute an important step in bringing the field of linguistics and its sub-disciplines closer to the frontier of fractal studies as have been reached in other disciplines. Furthermore, this study will contribute to demonstrating the use of multifractal analysis, which is a relatively new and uncommon approach to exploring the distribution of grammatical constructs in texts, in understanding the fractal nature of language.

Literature Review

Fractals in Languages

In the study of fractal linguistics, the Zipf-Mandelbrot Law and the Menzerath-Altmann Law are perhaps the two most important laws observed in fractal structures in languages. Both the Zipf-Mandelbrot Law and the Menzerath-Altmann Law highlight the self-similarity property of language structures and note the power law relationship that exists in many levels of language. This power law relationship is considered the evidence of fractality.

The Zipf-Mandelbrot Law was proposed as a generalization of Zipf's law and combined the work of these two great scholars – Benoit Mandelbrot's concept of fractal (Mandelbrot, 1967, 1983) and George Kingsley Zipf's power law. It is operationalized as a mathematical model that describes the inverse proportional relationship between the frequency rank of words and their number of occurrences. Such a relationship is observed in many natural languages. For example, Jayaram and Vidya (2008) demonstrated how this law applied to four Indian languages while Abney et al. (2014) noted power law clustering in their dyadic conversational data. A more recent study by Cong (2021) confirmed that the log-log rank-frequency distribution of words in an English corpora and words in a Mandarin Chinese corpus followed this law. Najafi et al. (2015) went further to suggest that this characteristic can be

effectively utilized for automatic extraction of keywords in a text as the word distribution follows this law.

Such an inverse proportional relationship is also noted in the Menzerath-Altmann Law, which was based on Paul Menzerath's work on the negative correlation between the size of a linguistic unit and its complexity (Menzerath, 1954) and was later operationalized by Gabriel Altmann into a mathematical formula (Altmann, 1980). According to the Menzerath-Altmann Law, the longer a language construct is, the smaller its constituents are. For example, the longer a sentence is (as measured in the number of clauses), the shorter the clauses are (as measured in the number of words). When plotted onto the logarithmic axes, the resulting graph is a linear line, confirming this relationship to be that of the power law – hence, fractality. This relationship has been observed, through empirical investigations, in natural languages, e.g. Japanese (Sanada, 2016), Chinese (Chen, 2018), as well as in translated texts (Jiang & Ma, 2022), interpreted texts (Jiang & Jiang, 2022), and in sign language (Andres et al., 2021), showing that the Menzerath-Altmann Law applies rather universally to human languages and across different levels of linguistic constructs. This also aligns with Hřebíček's conjecture that such relationship is not dependent on 'the unit of measurement used' (1998, p. 236).

More recently, with the introduction and integration of Complex Dynamic Systems Theory (CDST) into the fields of linguistics and language studies, the notion of fractality in language is again emphasized and revisited through the lens of CDST. Within the framework of CDST, there are a small handful of studies that investigated fractality in language and put forward evidence to show how this property manifests in language. Lowie et al.'s (2014) study, for example, demonstrated that the pattern of lexical processing in second language use showed characteristics of pink noise, which is a self-similar (fractal) pattern of variability. Their findings corroborated Kello's et al. (2008) study which reported evidence of pink noise in the intrinsic fluctuation of a spoken word – showing a fractal pattern in the acoustic measurements of word repetition. Similarly, Evans (2020) found that the frequency/density plot of clausal use by his participant across 30 weeks followed the power-law distribution and that the amount of variance in the data set was about the same across the three different temporal scales chosen for analysis (i.e. from 30 weeks, to 30 conversational turns in 1 week, to 28 AS-units in one conversational turn), hence evidence of self-similarity in the developmental trajectory of complex syntax. The findings of this study lent strong support to the CDST's notion of fractality in language.

Regardless of the underpinning theoretical perspective, be it the Menzerath-Altmann Law or the Zipf-Mandelbrot Law or CDST, all these studies demonstrated one main characteristic of fractal structures that is observed in language, i.e. the power law relationship, and hence supported the notion of language as a fractal structure. However, more in-depth

exploration into the fractal structure in languages revealed the possibility of multifractality in human languages.

Multifractal Structures and Multifractality in Languages

Multifractals are stochastic fractals, in which each of the smaller copies may not be in the exact geometrical shape as the original, and may, therefore, have a different fractal dimension from the other copies. Each level of iteration is formed with the same mechanism but with different random probability. The resulting structures are not geometrically self-similar in each level but are instead statistically similar while manifesting a random looking appearance, such as the human's vein structure. Real life fractals, like urban morphology (Ariza-Villaverde et al., 2013b; Chen & Huang, 2018), heartbeat and brain activity (Lin & Sharif, 2010), temperature anomaly (de Souza et al., 2013), air flow and pollutant dispersion (Mouzourides et al., 2021), river network (Ariza-Villaverde et al., 2013a), drainage system (Ariza-Villaverde et al., 2015), structure of blood vessels (Li et al., 2021) are shown to be of multifractal nature. These real-life multifractals do not manifest perfect self-similarity as do the mathematical fractals. In other words, they are not geometrically neat and are rather random-looking shapes. Common to all multifractals is that they are not possible to be explained by a simple rule or function nor described by a single fractal dimension – they are to be described by the combination of rules and chances.

Turning to the multifractals in the linguistic world, when we think about how language is produced, it is not difficult to notice that each language output is subject to the grammatical rules of the language and the stylistic choice, among other things, of the language users. In other words, there are so many possible options that we can use to express our thoughts. Each of these options has a probability to be selected as the output (i.e. what we eventually choose to produce). This probability is subject (largely) to the grammar of the language we are using and our own style of using that language. What we eventually produce is therefore the combined outcome of grammatical rules and probability (which is then finalized through our selection), which makes up the output itself. Following this line of thought, it is highly probable that fractals in language are multifractal. Moreover, according to CDST, language is a nonlinear dynamic system, in which chaos and order, attractors and repellers, probability and self-organization, among others, guide its manifestation (Larsen-Freeman, 2012; Lowie, 2013). Given these many factors in play, and particularly since all of them are stochastically contingent, it is likely that the fractal structures in language are multifractal.

Indeed, evidence of multifractality in human languages has been put forward by several studies. Carrizales-Velazquez et al. (2022), for example,

reported the presence of multifractality in their study of word length sequences in 22 English texts regardless of the methods used for detection, be it the Multi-Fractal Higuchi Dimension Analysis or the Multi-Fractal Detrended Fluctuation Analysis. Multifractality is also detected in 10 European languages, including English, in Chatzigeorgiou et al. (2017) who studied part of the Europarl parallel corpus. By converting texts into word-length series, they measured the multifractal spectrum of the resulting series and found evidence of multifractality in all of the 10 languages they focussed on. Similarly, Ausloos (2012) also found evidence of multifractality in both natural and human-made languages (such as Esperanto). By converting three texts (two English fantasy novels and one translation in Esperanto) into frequency time series as well as length time series and subsequently applying multifractal analysis to the resulting series, Ausloo found the existence of multifractal spectrum in all three texts examined. All these studies lend support to the hypothesis that fractals in language are of multifractal nature.

Therefore, in this paper, we will test this hypothesis by investigating the nature of fractality in the distribution of three grammatical constructs – i.e. Finite Verb Phrases (FVP), Noun Phrases (NP), and Head Noun (HN), respectively – in three types of written English texts: literary texts, children's literary texts, and academic texts (See the This Study section for the rationale of this selection). In this study, we will use multifractal analysis to detect and measure the multifractal spectrum of these constructs.

Multifractal Analysis

Multifractals cannot be described by a single fractal dimension. Each of the smaller copies of a multifractal may have a different fractal dimension from the other copies. This is because multifractals are stochastic fractals, where there is an element of randomness in their formation. Therefore, when measuring multifractals, probability needs to be taken into consideration. Analysis of mono-fractals, however, does not take into consideration the role of probability, understandably, because each level of iterations in mono-fractals is guided by the same rule with exact (non-randomized) probability. When exploring real-life fractals, which are mostly multifractals, it is therefore important to use a more appropriate analysis method that considers the role of stochastic probability. A multifractal analysis is one appropriate method to do this.

Multifractal Analysis is a method of analysing and characterizing multifractal structures through measurements and descriptions of global parameters and local parameters. The global parameters include the generalized correlation dimensions and the mass exponent, while the local parameters include the Lipschitz – Hölder (singularity) exponent and the fractal

dimensions of the fractal subsets. These parameters are obtained through the following formulae:

$$D_q = -\lim_{\varepsilon \rightarrow 0} \frac{I_q(\varepsilon)}{\ln \varepsilon} = \frac{1}{q-1} \lim_{\varepsilon \rightarrow 0} \frac{\ln \sum_{i=1}^{N(\varepsilon)} P_i(\varepsilon)^q}{\ln \varepsilon} \quad (1)$$

$$\tau(q) = (q-1)D_q \quad (2)$$

$$\alpha(q) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\ln \varepsilon} \sum_i \mu_i(\varepsilon) \ln P_i(\varepsilon) \quad (3)$$

$$\text{where } \mu_i(\varepsilon) = \frac{P_i(\varepsilon)^q}{\sum_i P_i(\varepsilon)^q}$$

$$f(\alpha) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\ln \varepsilon} \sum_i \mu_i(\varepsilon) \ln \mu_i(\varepsilon) \quad (4)$$

Equation (1) returns a series of values for the generalized correlation dimensions and equation (2) a series of values for the mass exponent. Equation (3) returns a series of values for the Lipschitz – Hölder (singularity) exponent and equation (4) a series of fractal dimensions of the fractal subsets (Interested readers are referred to Long et al. (2021) and Chen and Wang (2013) for more detailed explanation of multifractal analysis).

Not only does this analysis integrate the role of probability (through the term $P_i(\varepsilon)$ in the equations), but it also considers the possibilities that different subsets might have different weighing effects (through the power term q) on the overall pattern. Moreover, this analysis has a main benefit on the practical level, i.e. it will return different results for mono-fractals and for multifractals, making this analysis a good tool to use as a preliminary check on whether a structure is mono-fractal or multifractal.

As can be seen from the formulae, multifractal analysis does not produce a single measure of fractal dimension, but instead a set of 4 parameters, which form the spectrum of dimensions of the fractal structures in the overall pattern, including the global multifractal spectrum, i.e. **D_q -q spectrum**, and local parameter spectrum, i.e. **$f(\alpha)$ - α spectrum**. These four parameters and their associated spectra form a set of criteria for testing whether a structure is multifractal or not. These criteria include: (1) The inequality $D_0 > D_1 > D_2$, (2) An inverse S-shape for the D_q -q spectrum, (3) A non-linear plot of τ_q with respect to q , and (4) A unimodal curve for the $f(\alpha)$ - α spectrum. These criteria will be used in the present study to look for evidence of multifractality in language. In this study, only when all 4 criteria are satisfied will we contend that multifractality is evidenced.

Register Variation

Large-scale corpus linguistic studies and register studies, such as Biber and Gray (2016), Goulart et al. (2020), have pointed out that language manifests differently in different genres and modalities. For example, the oral and narrative discourse in English is more clausal, while the written non-narrative literate discourse (such as English academic prose) is more phrasal (Biber, 2014). Looking at registration variation through these textual-differences linguistically, Biber (2019) noted that modern academic prose in English, particularly the non-narrative content-focussed literate discourse, is heavily characterized by phrasal features, such as abstract nominalization and complex noun phrases. Literary texts, on the hand, are shown to have a preference for reporting verbs, adverbials of time, participle adjectives, action-describing features, clausal structures, verbal structures, past-tense verbs, etc (Biber & Conrad, 2009; Egbert & Mahlberg, 2020; Mazdayasna & Firouzi, 2013). As such, the widely held understanding regarding the register variation between academic and literary texts is that academic texts are generally less verbal, more nominal, and employ more phrasal elaboration, and that literary texts are more narrative in nature and employ more clausal elaboration, with many grammatical features that are more common in the conversational contexts than in the written discourse (Biber & Conrad, 2019).

Similarly, children's literature (such as fairy tales and picture books) is shown to have the features that are similar or relatively close to those of literary texts. Relative clauses, for example, were found to be common in this register and their usage seemed to increase with the intended age of their readership (Montag & MacDonald, 2015). This grammatical feature was found to co-occur frequently with noun animacy and pronouns in children's literary texts (Hsiao et al., 2022). This aligns with the characteristics of child-directed language being generally characterized with rather frequent word repetition and lower word density (Hindman et al., 2021), which is usually reflected in children fiction through high use of lexical verbs and direct speech (Thompson & Sealey, 2007). Perhaps also due to its educational purposes, children's literary texts were frequently found to contain prevalent uses of canonical utterances (i.e. the Subject-Verb-Object/Complement constructions) (Cameron-Faulkner & Noble, 2013), as well as compound sentences (e.g. sentences joined with 'and') (Rahman, 2012). These, however, are not as ubiquitous in the non-children-directed literary texts.

We were therefore interested to see, for example, how the dimensions of multifractality of the distribution of FVP in academic texts compare to those of the same construct in literary texts, given the less verbal and more nominal characteristics of academic prose. We were also interested to see if these parameters are also affected by the level of complexity of the texts

themselves – for example, how the dimensions of multifractality of the distribution of NP in novels compare to those of the same construct in fairy tales. We hypothesized that while multifractality is a universal trait, its dimensions, however, vary across different types of texts.

This Study

In this study, we tested this hypothesis by examining three grammatical structures in English: Finite Verb Phrases (FVP), Noun Phrases (NP), and Head Nouns (HN). Our research questions were as follows:

- (1) a. Do the distributions of Finite Verb Phrases (FVP), Noun Phrases (NP), and Head Nouns (HN) in English manifest fractal structures?
b. If yes, are they mono-fractal or multifractal?
- (2) How does the generalized correlation dimension of these multifractal distributions align with or differ from the field's current understanding about the register characteristics of English texts?

In this study, we chose to include three types of texts for investigation: literary texts (novels written in English), children's literary texts (fairy tales), and academic texts (journal articles) – see Text Selection and Coding subsection for more information on these texts. The rationale for choosing to include these three types of texts was twofold: (1) to see if the characteristics of multifractality could be universally found across different types of texts, and more importantly (2) to leverage on the different characteristics of texts across these three different types to see whether the information unveiled by the property of multifractality in language aligns with (or differs from) the current understanding of text characteristics in different registers.

In this paper, a total of 12 sets of analysis were conducted. Of these, nine sets were conducted on texts, while the other three were done on a rectangle, a triangle, and a Cantor ternary set, respectively. The results of the analysis on texts are the main focus of the study, while the analyses on a rectangle, a triangle, and a Cantor ternary set were conducted as our internal proof of the validity of this Multifractal Analysis method in differentiating non-fractal, mono-fractal, and multifractal structures.²

Text Selection and Coding

We selected three texts per category for inclusion into the analysis in this paper. The details about these texts are summarized in [Table 1](#).

Table 1. Descriptions of texts selected for MFA in this study.

No	Category	Text name	Sourced from	Section excerpted	Word count
1	Literary texts (Sourced from published novels)	Literary text 1	The Maze Runner (2009, by James Dashner)	Chapter 2	1139
2		Literary text 2	The Fault in Our Stars (2012, by John Green)	Chapter 2	1101
3		Literary text 3	Gone Girl (2012, by Gillian Flynn)	Chapter 1	1147
4	Children's literary texts (Sourced from fairy tales that are used as reading materials in schools in the UK)	Children text 1	Little Red Riding Hood	All	247
5		Children text 2	Three Little Pigs	All	619
6		Children text 3	Hansel & Gretel	All	569
7	Academic texts (Sourced from published journal articles)	Academic Text 1	Text-linguistic approaches to register variation (2019, by Douglas Biber, <i>Register Studies</i>)	Section 4 Discussion (1 paragraph) + Section 5 Discussion/Conclusion (whole section)	1195
8		Academic Text 2	Technical disruption in foreign language teaching: The rise of simultaneous machine translation (2018, by Scott Crossley, <i>Language Teaching</i>)	Counter argument section (1 paragraph) + Implication and conclusion (whole section)	1116
9		Academic Text 3	Investigating variability in L2 development: Extending a complexity theory perspective on L2 writing studies and authorial voice (2020, by Gary Fogal, <i>Applied Linguistics</i>)	Discussion (whole section) + Conclusion (paragraph 1).	1128

As can be seen from Table 1, all the texts excerpted from the sources (except those in the children's literary text category) were of around 1100–1200 words long. The reason for this length is operational (see below for more details) in order to ensure that there were at least 3–4 words in each of the subsets in the last level of analysis. However, for children's literary texts, the length of the texts did not reach this range because of practical reasons—children's literary texts are generally not long. As such, the levels of analysis were adjusted accordingly, to ensure that at least 3–4 words in each of the subsets in the last level of analysis as well.

In excerpting the literary texts, we took the first 1100–1200 words from the sections as indicated in Table 1. In determining the end of the excerpt, we included as many complete paragraphs as possible. When this was not possible, we would end the excerpt at the last sentence where the word count drew closest to the upper margin of our target word count. In excerpting the academic texts, we intentionally chose the sections with the

least amount of in-text citation, paraphrasing, referencing. We then cleaned the texts to remove any of citation brackets, footnotes/endnotes, etc., and the remaining first 1100–1200 words were included in our analysis. Again, we included as many complete paragraphs as possible. Where not possible, we would end the excerpt at the last sentence where the word count was the closest to our upper margin.

Once the texts were excerpted, they were then coded for three types of grammatical constructions, namely finite verb phrases (FVP), noun phrases (NP), and head nouns (HN). We followed Verspoor and Sauter's (2000) definition in operationalizing these three constructs in this study and did not distinguish between a single-word FVP and a multi-word FVP, as well as between a single-word NP and a multi-word NP, as the focus of this study was not on the structural distinction between them.

Once the coding phase was completed, the coded texts were then submitted to the following steps:

Step 1: each text was considered as a whole at this first level of analysis ($k = 1$; $\varepsilon = 1$), and the occurrences of the target constructions were tallied. When analysing the FVP constructions, then occurrences of all FVPs were counted in words. Similarly, when analysing the NP constructions and the HN constructions, their occurrences were also counted in words. For example, in the following text of Little Red Riding Hood (text no. 4 in this study), the FVP constructions were identified (underlined here for the purpose of illustration):

Once upon a time there was a girl called Little Red Riding Hood. She lived with her mother in a village near a forest. One day, Little Red Riding Hood went to visit her grandmother. She took a basket of food with her. On her way, Little Red Riding Hood met a wolf. "Hello," said the wolf. "Where are you going?" "I'm going to visit my grandmother who lives in the forest," explained Little Red Riding Hood. The wolf ran to Grandmother's house. He went inside, put on a nightgown and got into her bed. A little later, Little Red Riding Hood came to the house. She knocked on the door, then went inside. Little Red Riding Hood went over to Grandmother's bed. "Oh, Granny, what big ears you have," she said. "All the better to hear you with," answered the wolf. "Oh, Granny, what big eyes you have," said Little Red Riding Hood. "All the better to see you with," replied the wolf. "Oh, Granny, what big teeth you have," said Little Red Riding Hood. "All the better to eat you with," said the wolf. "Help!" shouted Little Red Riding Hood as she realized that there was a wolf in her grandmother's bed. She ran out of the house. A woodcutter was nearby. He heard Little Red Riding Hood's scream and ran to the house. The woodcutter hit the wolf over the head. The wolf ran away, and Little Red Riding Hood never saw him again.

Figure 1. FVP constructs in the Little Red Riding Hood.

The number of words that were underlined (i.e. identified as FVP constructs) was tallied (amounting to 40 words) and recorded in the first line in Table 2.

Table 2. An example of tally sheet for FVP constructions in the Little Red Riding Hood.

Level of iteration	Measurement box size	Box number	Tally	Probability $P_i(\varepsilon)$
k = 1	$\varepsilon = 1$	1	40	1
k = 2	$\varepsilon = 1/2$	1	20	0.5
		2	20	0.5
k = 3	$\varepsilon = 1/4$	1	9	0.225
		2	11	0.275
		3	9	0.225
		4	11	0.275
k = 4	$\varepsilon = 1/8$	1	3	0.075
		2	6	0.15
		3	7	0.075
		4	4	0.1
		5	5	0.125
		6	4	0.1
		7	7	0.175
		8	4	0.1

Repeat the process until k = 9; $\varepsilon = 1/256$

Box 1

Once upon a time there was a girl called Little Red Riding Hood. She lived with her mother in a village near a forest. One day, Little Red Riding Hood went to visit her grandmother. She took a basket of food with her. On her way, Little Red Riding Hood met a wolf. “Hello,” said the wolf. “Where are you going?” “I’m going to visit my grandmother who lives in the forest,” explained Little Red Riding Hood. The wolf ran to Grandmother’s house. He went inside, put on a nightgown and got into her bed. A little later, Little Red Riding Hood came to the house. She knocked on the door, then went inside. Little Red Riding Hood went over to Grandmother’s bed. “Oh,

Box 2

Granny, what big ears you have,” she said. “All the better to hear you with,” answered the wolf. “Oh, Granny, what big eyes you have,” said Little Red Riding Hood. “All the better to see you with,” replied the wolf. “Oh, Granny, what big teeth you have,” said Little Red Riding Hood. “All the better to eat you with,” said the wolf. “Help!” shouted Little Red Riding Hood as she realized that there was a wolf in her grandmother’s bed. She ran out of the house. A woodcutter was nearby. He heard Little Red Riding Hood’s scream and ran to the house. The woodcutter hit the wolf over the head. The wolf ran away, and Little Red Riding Hood never saw him again.

Figure 2. The text split into two.

Step 2: each text was split into two halves in the next level of iteration ($k = 2$; $\varepsilon = \frac{1}{2}$). In each half, the occurrences of the target constructions were tallied (see Figure 2). For example,

The total number of words in the FVP constructions in box 1 and box 2 was 20 and 20, respectively, and recorded in the next two lines in Table 2.

Step 3: This process was repeated until the subsets became too small to be meaningful. In this study, this was reached at $k = 9$ ($\varepsilon = 1/256$), in which the text of approximately 1200 words were split into 256 subsets of about 4–5 words each.

Step 4: Once all the occurrences of the target constructions in each subset in each iteration were noted, they were then converted to probability. To make sure that the results were comparable, this study opted for normalized probability, which was calculated by dividing the number of occurrences at each level by total number of occurrences in the text. Hence, in the conversion for probability, the denominator was always the number of occurrences of the target construction at level 1.

Step 5: The probability was then converted to Renyi's information entropy through the following formula:

$$I_q(\varepsilon) = \frac{1}{1-q} \ln \sum_{i=1}^{N(\varepsilon)} P_i(\varepsilon)^q \quad (5)$$

Step 6: The results were then fitted to the four formulae (see formulae 1–4) to obtain the global parameters (D_q , τ_q) and the local parameters (α , f). The partition function was defined as:

$$\Gamma(q, \tau) = \sum_{i=1}^{N(\varepsilon)} \mu_i^q(\tau) \quad (6)$$

In this paper, we reported on the results of this multifractal analysis to answer our first research question and explored more deeply on the generalized fractal dimension parameter to answer our second research question.

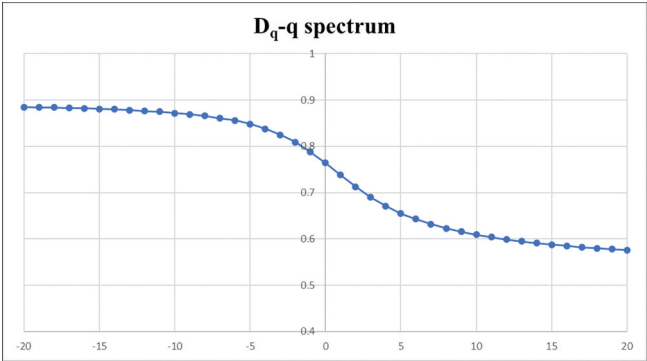
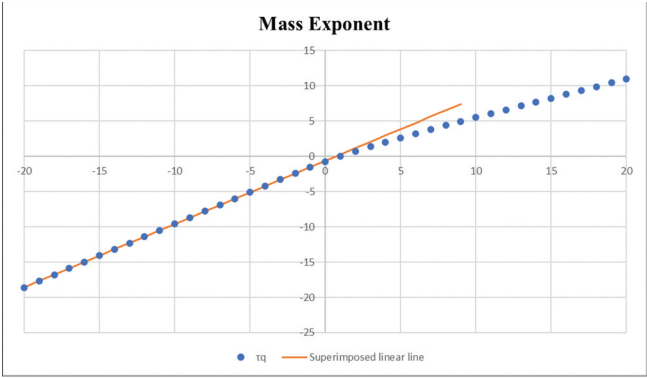
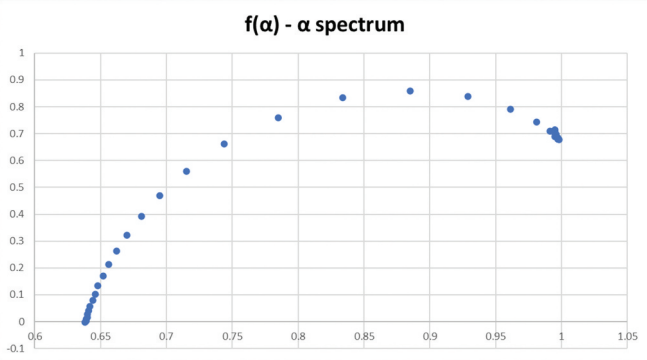
Results

The results of the analyses showed that the distributions of all the three grammatical constructions (i.e. FVP, NP, and HN), across the three categories of texts explored in this study, were multifractal. They satisfied all four criteria of multifractality, i.e. (1) The inequality $D_0 > D_1 > D_2$, (2) An inverse S-shape for the D_q - q spectrum, (3) A non-linear plot of τ_q with respect to q , and (4) A unimodal curve for the $f(\alpha)$ - α spectrum.

Table 3 shows an example of the results of MFA on the FVP constructs in an academic text (text no. 7 in Table 1).

Therefore, the answer to Research Question no.1 in this study is positive. The distributions of Finite Verb Phrases (FVP), Noun Phrases (NP), and Head Nouns (HN) in English manifest multifractal structures.

Table 3. Satisfying the criteria of multifractality.

Criterion	Finite verb phrases		
$D_0 > D_1 > D_2$	D_0 0.764	D_1 0.738	D_2 0.713
R^2	0.962	0.964	0.964
An inverse S-shape for the D_q -q spectrum	D_q-q spectrum 		
A non-linear plot of τ_q with respect to q (as evident in the bent after point 0)	Mass Exponent 		
A unimodal curve for the $f(\alpha)$ - α spectrum	$f(\alpha)$ - α spectrum 		

In regard to research question no. 2, the generalized correlation dimension (D_q) offers interesting findings. As can be seen in Table 4, all the D_q values reported here had a very good goodness-of-fit value of about 90% or above. There are 3 important findings highlighted by the D_q values reported in Table 4:

Table 4. The generalised correlation Dimensions of FVP, NP, and HN in all 9 texts.

No	Text	Generalised Correlation Dimension														
		FVP					NP					HN				
		q = -20	q = 0	q = 1	q = 2	q = 20	q = -20	q = 0	q = 1	q = 2	q = 20	q = -20	q = 0	q = 1	q = 2	q = 20
1	Literary Text 1	D_g 0.967	0.828	0.803	0.782	0.673	1.12	0.899	0.883	0.871	0.813	1.02	0.893	0.879	0.866	0.745
2	Literary Text 2	R^2	0.936	0.973	0.974	0.975	0.955	0.975	0.975	0.975	0.974	0.965	0.975	0.975	0.975	0.969
		D_g	0.971	0.845	0.826	0.809	0.7	1.12	0.902	0.887	0.877	1.03	0.891	0.877	0.865	0.773
		R^2	0.94	0.973	0.974	0.974	0.973	0.952	0.974	0.975	0.975	0.963	0.975	0.975	0.975	0.973
3	Literary Text 3	D_g	0.942	0.815	0.792	0.77	0.637	1.06	0.902	0.889	0.88	1.02	0.894	0.882	0.871	0.785
4	Children Text 1	R^2	0.942	0.973	0.974	0.974	0.967	0.966	0.974	0.975	0.975	0.968	0.975	0.975	0.975	0.971
		D_g	0.826	0.738	0.721	0.702	0.574	1.02	0.845	0.832	0.821	1.01	0.838	0.813	0.79	0.681
		R^2	0.895	0.942	0.944	0.945	0.941	0.943	0.947	0.947	0.942	0.932	0.948	0.948	0.949	0.938
5	Children Text 2	D_g	0.939	0.812	0.779	0.746	0.617	1.05	0.878	0.864	0.854	1.02	0.874	0.859	0.845	0.756
6	Children Text 3	R^2	0.934	0.963	0.964	0.964	0.946	0.957	0.964	0.964	0.96	0.957	0.964	0.965	0.965	0.962
		D_g	0.93	0.824	0.805	0.786	0.651	1.07	0.878	0.86	0.846	1	0.871	0.855	0.848	0.767
		R^2	0.928	0.963	0.964	0.965	0.957	0.958	0.964	0.965	0.964	0.95	0.965	0.962	0.965	0.966
7	Academic Text 1	D_g	0.885	0.764	0.738	0.713	0.576	1.1	0.901	0.891	0.884	0.99	0.893	0.88	0.868	0.754
8	Academic Text 2	R^2	0.904	0.962	0.964	0.964	0.959	0.959	0.974	0.975	0.975	0.966	0.975	0.975	R0.975	0.971
		D_g	0.894	0.778	0.756	0.736	0.61	1.1	0.901	0.89	0.882	0.981	0.89	0.878	0.867	0.785
		R^2	0.939	0.966	0.968	0.969	0.96	0.965	0.974	0.975	0.975	0.965	0.975	0.975	0.975	0.975
9	Academic Text 3	D_g	0.867	0.794	0.776	0.754	0.59	1.07	0.903	0.896	0.891	1.01	0.895	0.882	0.87	0.795
		R^2	0.939	0.965	0.966	0.967	0.949	0.961	0.974	0.975	0.975	0.966	0.975	0.975	0.975	0.973

- (1) For FVPs, the D_q values for all the three academic texts were the lowest. This aligns with the characteristics of academic texts being not verbal.
- (2) For NPs, the D_q values for academic texts and literary texts were comparable, while the D_q values for children's literary texts were slightly lower. This is expected of both text types: academic texts are generally compact and heavily nominalized, while for literary texts, nouns are generally accompanied by other constituents (premodifiers, postmodifiers, etc) to give a rich description of the story, the characters, the place, etc. In children's literary texts, noun phrases are generally, and relatively, shorter to align with children's cognitive capacity in processing them.
- (3) For HNs, there was no significant difference among the 3 types of texts.

Discussion

While the distributions of the three grammatical constructs (i.e. FVP, NP, and HN) showed the characteristics of multifractality across each of the 9 texts analysed in this study, the parameters of their distributions differed across types of grammatical constructs and across categories of texts. In this section, we will discuss in detail how these parameters differed and what the differences might mean.

Finite Verb Phrases

The generalized correlation dimensions of the distributions of the FVP constructs showed a clear descending pattern of $D_0 > D_1 > D_2$ in all texts. While such a descending order met the first criterion of multifractality, the values in each of these dimensions offered information about the said distributions. The value of D_0 , for example, indicates the space filling capacity, which offers information on the extent on which this construct fills the space of the text. From Table 4, it is clear that the D_0 values were the lowest in the academic texts, and highest in the children's literary texts. This is in line with our common understanding of these three registers: academic texts are less verbal than the other two registers, and children's literary texts are mainly verbal (Biber & Conrad, 2019).

Equally interesting is the value of D_1 which bears the information on the uniformity of the distribution of the multifractals. In other words, the D_1 value offers information on how uniformly distributed a certain multifractal construct is. In our study, this value can be interpreted as how many fractal elements (i.e. FVP constructs) appear at a place (i.e. the text). Our results showed that, this value was again lowest in the academic texts. Interestingly, this value was higher in the literary texts than in the children's literary texts, suggesting that the FVP constructs were more uniformly distributed in the literary texts than in the

children's literary texts, most likely due to the lower word density in child-directed language (Hindman et al., 2021).

The global multifractal spectrum, which was obtained by plotting the values of the generalized correlation dimensions against the q values, offered a visual comparison of the multifractality of the distribution of FVP constructs in the 9 texts (see Figure 3).

As can be seen in Figure 3, the highest tails are of the literary texts and the children's literary texts.³ This is not unexpected as literary texts are generally verbally rich (Thompson & Sealey, 2007). The lowest tails are of the academic texts, as this register is possibly the least verbal type of text. Also, as a higher right tail suggests a more compact structure, this observation therefore converges with the above discussion on the space filling capacity and the uniformity of distribution of the FVP constructs in these texts.

Noun Phrases

Looking at the generalized correlation dimensions of the multifractal distributions of the NP constructs (see Table 4), it is clear that the D_0 values were quite comparable between the literary texts and the academic texts, and relatively lower in the children's literary texts. While one would expect that these values for the academic texts would be higher than those of the other 2 registers as academic texts are known to be highly nominalized (and hence will be more saturated with NPs), the results showed that there was not much difference in the values of D_0 between the academic texts and the literary texts.

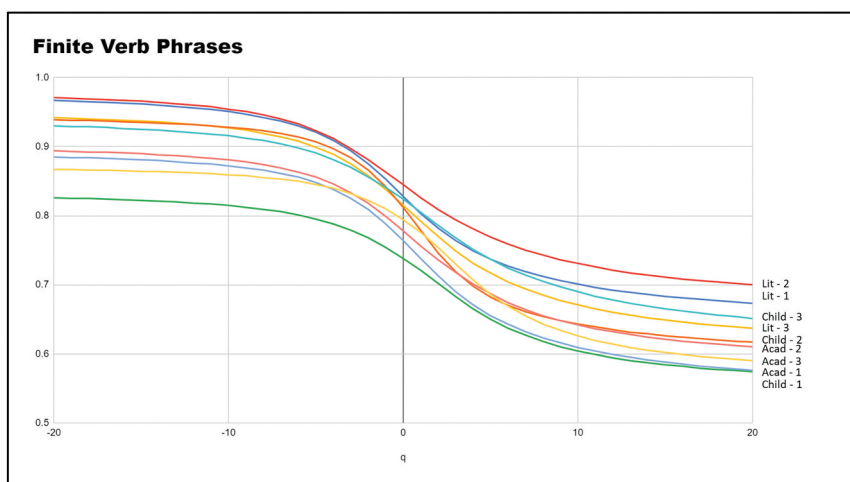


Figure 3. Global spectrums of FVP distribution in all 9 texts.

However, considering the syntactic constructions in literary works, one would notice that nouns are more than often modified with adjectives (and sometimes, a string of adjectives) for description of places, characters, events, etc. in literary works, as noted in findings from the corpus linguistics field, that many literary texts are of narrative nature with high occurrences of adjectives (Biber & Conrad, 2019; Thompson & Sealey, 2007). Therefore, it is perhaps not surprising to see that the distributions of NP constructs in literary texts were quite comparable to those in the academic texts.

During the coding process in this study, the authors noticed that there were many NPs (Adj + N constructions) in the literary texts, and they were distributed rather uniformly across the body of the texts, as can be seen in the values of D_1 . Interestingly, the D_1 and D_2 values were also comparable between the literary texts and the academic texts, and slightly lower in the children's literary texts. This suggests that in the literary texts and the academic texts (more than in the children's literary texts), there was a more uniform distribution of NP constructs, as well as a higher chance of finding other NP constructs in neighbouring segments of the texts. In other words, if a text was cut into many segments of (roughly) similar length, the chance of finding an NP in a segment and then yet another NP in the next segment appeared to be higher in the literary texts and the academic texts than in the children's literary texts as this construct seemed to be more uniformly distributed in literary texts and academic texts than in children's literary texts – as seen in the D_1 and D_2 values obtained in this study.

The global multifractal spectra of the multifractal distributions of NP constructs in the 9 texts in this study are presented in Figure 4. There are two main points of observation that can be made from these spectrums:

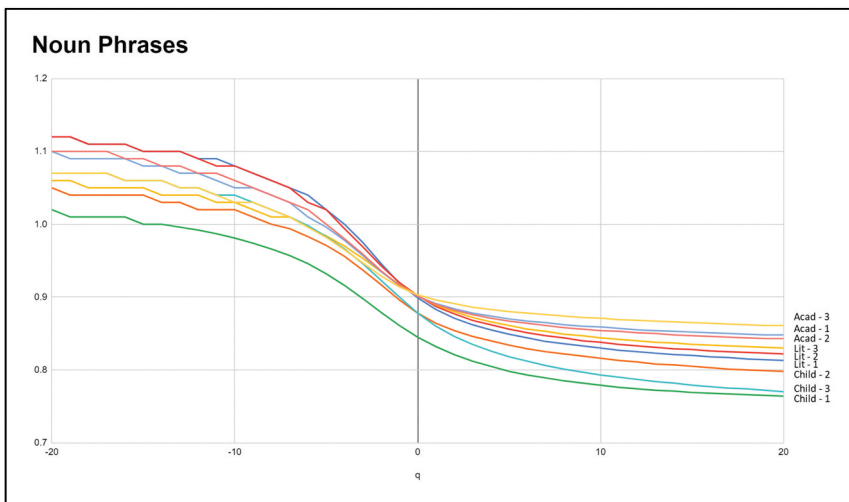


Figure 4. Global spectrums of NP distribution in all 9 texts.

- (1) The academic texts had higher right tails than the literary texts and the children's literary texts. This suggests that the academic texts were more nominal than the literary texts which were in turn also more nominal than the children's literary texts.
- (2) The curly left tails show the abnormal range of D_q values that exceeded the upper limit of this parameter (i.e. its topological dimension $D_t = 1$). This suggests that the range of q values chosen in these rounds of analysis was wider than necessary.

Head Nouns

The last construct analysed in this study is the HN constructs. Like the other two constructs in this study, the distributions of HN constructs also manifested the characteristics of multifractality.

Table 4 shows the values of D_0 , D_1 , D_2 in the HN constructs across the 9 texts analysed in this study. Very similar to those of the NP constructs, the values of D_0 , D_1 , D_2 of the HN constructs were comparable between the literary texts and the academic texts, followed by lower values in the children's literary texts. This is, again, not unexpected, as the HN constructs are part of the NP constructs and therefore likely to share similar characteristics.

As can be seen in Figure 5, the D_q spectrums are all in an inverse S-shape, again confirming the multifractal characteristics of the constructs under investigation. However, unlike in the NP constructs, there is less obvious clustering in the right tails here and that they converge quite well together, except for children text no. 1 (which has a significantly low word count). This means that the distribution of this construct in all the texts was relatively comparable. Such a

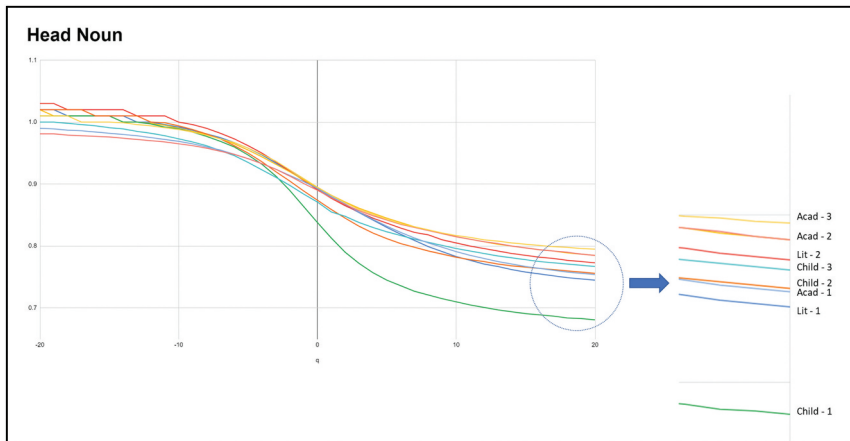


Figure 5. Global spectrums of HN distribution in all 9 texts.

finding aligns with Thompson and Sealey's (2007) study that reported comparable frequency of nouns in the children and the adult fiction corpora.

Contribution and Limitations of the Study

This study constitutes an explorative step into understanding and measuring the multifractal nature of language, using an approach relatively new and uncommon to linguists – the multifractal spectrum analysis. It adds to the demonstration of using this analysis to identify the multifractal patterns in texts. Moreover, this study lends support to Hřebíček's conjecture that fractality is not dependent on the unit of measurement used. Unlike Ausloos (2012) and Chatzigeorgiou et al. (2017) whose measurement unit was word length (operationalized as the number of letter in each word), the unit of measurement in this study was phrase length (operationalized as the number of words in each phrase). Regardless of the measurement units used, evidence of multifractality was found. This could perhaps be an aspect to consider when teaching and assessing language. For example, when teaching the noun phrase constructs, it might benefit learners to show them the (multi-)fractal structure of these phrases and their distribution in English texts. When it comes to language assessment, it might merit consideration to re-evaluate the currently available measurements of syntactic complexity (which are generally based on linear measurements) in light of this (multi-)fractal characteristic of syntactic constructs.

Our results also lend support to the CDST's hypothesis about fractality in language (Evans, 2020; Larsen-Freeman & Cameron, 2008). Our analysis of the distribution of FVP, NP and HN constructs in English across three types of texts confirmed this property. Last but not least, this study contributes to expanding our understanding of fractal linguistics on the phrasal level.

However, as an explorative study, we do not have many references to guide us in terms of choosing appropriate benchmarks, or parameter settings, or direction settings. Therefore, many of the choices we made in this first and explorative study were arbitrary. For example, in deciding the number of texts in each category, we were limited by the time and resource constraints, and decided to go with 3 texts per category – an arbitrary number we chose for convenience. In selecting which texts to be included, we resorted to convenient sampling and chose from texts that were readily available in the public domains. Furthermore, the texts we selected for the academic category came from closely related cognate disciplines. Given that texts are written according to the disciplinary conventions, our selection might have limited the extent of the evidential support for our hypothesis.

Also, during the analysis, we made several discretionary decisions, due to the lack of extensive referential comparisons in the field. For example, when deciding the range of the epsilon values, we went with $k = 9$. This value was decided upon checking the application of this analysis in other fields (e.g. urban

morphology study such as Long et al. (2021)) and we were aware that there is a possibility that it might not be the most appropriate value for textual analysis, particularly due to the limited number of data points when estimating the limit values through regression analysis. Mathematically speaking, the more data points we have, the better the estimate will be. However, considering our own limited resources, we went with this value ($k = 9$) as per commonly used in other fields. Similarly, when deciding the width of the moment of order band, we went with $q = \pm 20$, for the same reason – see for example Saeedimoghaddam et al. (2020). The results of this study showed that while this arbitrary q value worked well for the FVP constructs, it was not the case with the NP and the HN constructs. A further customization of q values will be needed for each type of constructs under investigation.

Conclusion

The results of this study show that fractality in the distribution of FVP, NP, and HN constructs in English texts is indeed multifractal. The parameters of their multifractal property offer information to describe their distribution in English texts. Through examining the parameters of these multifractal constructs, this study shows that, (1) academic texts are less verbal compared to the other two categories, (2) while academic texts and literary texts are relatively comparable in terms of the use of NP, and the distribution of NP is more compact in the academic texts, (3) there is not much difference in the distribution of the HN constructs in these 3 categories of texts. All of these align with the general understanding about the characteristics of academic texts, literary texts, and children's literary texts (Biber & Conrad, 2019).

Based on the results of our study and other studies on multifractality in language, we hypothesize that multifractal is likely an inherent property of language and would be found in other registers as well. As the texts analysed in this study were all composed by experienced L1 writers, our findings open a world of possibility for further research, for example, to investigate multifractality in texts written by less experienced and/or L2 writers, or multifractality in language produced in other modalities. We would therefore like to invite passionate colleagues and fellow fractal linguists to try and apply this analysis for other registers/genres/categories of texts and utterances. Together, we are able to test if multifractality is indeed a universal property of language.

Notes

1. See https://www.google.com/search?q=fractal+shapes&sxsrf=APwXEddSkpkCCpeKz_-gV8sbExHNWxw1Rw:1681097634329&source=lnms&tbm=isch&sa=X&ved=2ahUKEwiSuc7asJ7-AhVzyjgGHZyhA-IQ_AUoAXoECAEQAw

&biw=1024&bih=481&dpr=1.88 for visual representations of fractal structures.

2. In this study, we fitted this multifractal analysis onto regular Euclidean shapes and a mono-fractal. The results differed greatly from those of multi-fractals. The results can be made available upon request.
3. With the exception of Children text 1, which may present itself as an outlier due to its extremely low word count.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Rosmawati  <http://orcid.org/0000-0002-1697-5705>

Wander Lowie  <http://orcid.org/0000-0002-2241-0276>

References

- Abney, D., Paxton, A., Dale, R., & Kello, C. (2014). Complexity matching in dyadic conversation. *Journal of Experimental Psychology: General*, 143(6), 2304–2315. <https://doi.org/10.1037/xge0000021>
- Altmann, G. (1980). Prolegomena to Menzerath's law. In R. Grotjahn (Ed.). *Glottometrika*. (vol. 2, pp.1–10). Brockmeyer.
- Andres, J., Benešová, M., & Langer, J. (2021). Towards a fractal analysis of the sign language. *Journal of Quantitative Linguistics*, 28(1), 77–94. <https://doi.org/10.1080/09296174.2019.1656149>
- Ariza-Villaverde, A., Jiménez-Hornero, F., & Gutiérrez de Ravé, E. (2013a). Multifractal analysis applied to the study of the accuracy of DEM-based stream derivation. *Geomorphology*, 197, 85–95. <https://doi.org/10.1016/j.geomorph.2013.04.040>
- Ariza-Villaverde, A., Jiménez-Hornero, F., & Gutiérrez de Ravé, E. (2013b). Multifractal analysis of axial maps applied to the study of urban morphology. *Computers, Environment and Urban Systems*, 38, 1–10. <https://doi.org/10.1016/j.compenvurbsys.2012.11.001>
- Ariza-Villaverde, A., Jiménez-Hornero, F., & Gutiérrez de Ravé, E. (2015). Influence of DEM resolution on drainage network extraction: A multifractal analysis. *Geomorphology*, 241, 243–254. <https://doi.org/10.1016/j.geomorph.2015.03.040>
- Ausloos, M. (2012). Measuring complexity with multifractals in texts. Translation effects. *Chaos, Solitons & Fractals*, 45(11), 1349–1357. <https://doi.org/10.1016/j.chaos.2012.06.016>
- Biber, D. (2014). Using multi-dimensional analysis to explore cross-linguistic universals of register variation. *Languages in Contrast*, 14(1), 7–34. <https://doi.org/10.1075/lic.14.1.02bib>
- Biber, D. (2019). Text-linguistic approaches to register variation. *Register Studies*, 1(1), 42–75. <https://doi.org/10.1075/rs.18007.bib>
- Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511814358>

- Biber, D., & Conrad, S. (2019). *Register, genre, and style* (2 ed.). Cambridge University Press. <https://doi.org/10.1017/9781108686136>
- Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511920776>
- Cameron-Faulkner, T., & Noble, C. (2013). A comparison of book text and child directed speech. *First Language*, 33(3), 268–279. <https://doi.org/10.1177/0142723713487613>
- Carrizales-Velazquez, C., Donner, R. V., & Guzmán-Vargas, L. (2022). Generalization of Higuchi's fractal dimension for multifractal analysis of time series with limited length. *Nonlinear Dynamics*, 108(1), 417–431. <https://doi.org/10.1007/s11071-022-07202-2>
- Chatzigeorgiou, M., Constantoudis, V., Diakonou, F., Karamanos, K., Papadimitriou, C., Kalimeri, M., & Papageorgiou, H. (2017). Multifractal correlations in natural language written texts: Effects of language family and long word statistics. *Physica A: Statistical Mechanics and Its Applications*, 469, 173–182. <https://doi.org/10.1016/j.physa.2016.11.028>
- Chen, H. (2018). Testing the Menzerath-Altmann Law in the sentence level of written Chinese. *Open Access Library Journal*, 5, e4747. <https://doi.org/10.4236/oalib.1104747>
- Chen, Y., & Huang, L. (2018). Spatial measures of urban systems: From entropy to fractal dimension. *Entropy*, 20(12), 991. <https://doi.org/10.3390/e20120991>
- Chen, Y., & Wang, J. (2013). Multifractal characterization of urban form and growth: The case of Beijing. *Environment and Planning B: Planning and Design*, 40(5), 884–904. <https://doi.org/10.1068/b36155>
- Cong, J. (2021). A Zipfian approach to words in contexts: The cases of Modern English and Chinese. *Journal of Quantitative Linguistics*, 29(4), 465–484. <https://doi.org/10.1080/09296174.2021.1926110>
- de Souza, J., Duarte Queirós, S. M., & Grimm, A. M. (2013). Components of multifractality in the central England temperature anomaly series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 23(2), 023130. <https://doi.org/10.1063/1.4811546>
- Egbert, J., & Mahlberg, M. (2020). Fiction – one register or two?: Speech and narration in novels. *Register Studies*, 2(1), 72–101. <https://doi.org/10.1075/rs.19006.egb>
- Evans, D. (2020). On the fractal nature of complex syntax and the timescale problem. *Studies in Second Language Learning and Teaching*, 10(4), 697–721. <https://doi.org/10.14746/ssllt.2020.10.4.3>
- Goulart, L., Gray, B., Staples, S., Black, A., Shelton, A., Biber, D., Egbert, J., & Wizner, S. (2020). Linguistic perspectives on register. *Annual Review of Linguistics*, 6(1), 435–455. <https://doi.org/10.1146/annurev-linguistics-011718-012644>
- Hindman, A., Farrow, J., Anderson, K., Wasik, B., & Snyder, P. (2021). Understanding child-directed speech around book reading in toddler classrooms: Evidence from early head start programs. *Frontiers in Psychology*, 12, 719783. <https://doi.org/10.3389/fpsyg.2021.719783>
- Hřebíček, L. (1998). Language fractals and measurement in texts. *Archív Orientální*, 66(3), 234–242.
- Hsiao, Y., Dawson, N. J., Banerji, N., & Nation, K. (2022). The nature and frequency of relative clauses in the language children hear and the language children read:

- A developmental cross-corpus analysis of English complex grammar. *Journal of Child Language*, 50(3), 555–580. <https://doi.org/10.1017/S0305000921000957>
- Jayaram, B., & Vidya, M. (2008). Zipf's law for Indian languages. *Journal of Quantitative Linguistics*, 15(4), 293–317. <https://doi.org/10.1080/09296170802326640>
- Jiang, X., & Jiang, Y. (2022). Menzerath–Altmann Law in consecutive and simultaneous interpreting: Insights into varied cognitive processes and load. *Journal of Quantitative Linguistics*, 29(4), 541–559. <https://doi.org/10.1080/09296174.2022.2027657>
- Jiang, Y., & Ma, R. (2022). Does Menzerath–Altmann Law hold true for translational language: Evidence from translated English literary texts. *Journal of Quantitative Linguistics*, 29(1), 37–61. <https://doi.org/10.1080/09296174.2020.1766335>
- Kello, C. T., Anderson, G. G., Holden, J. G., & Van Orden, G. C. (2008). The pervasiveness of 1/f scaling in speech reflects the metastable basis of cognition. *Cognitive Science*, 32(7), 1217–1231. <https://doi.org/10.1080/03640210801944898>
- Larsen-Freeman, D. (2012). Complex, dynamic systems: A new transdisciplinary theme for applied linguistics. *Language Teaching*, 45(2), 202–214. <https://doi.org/10.1017/S0261444811000061>
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford University Press.
- Lin, D. C., & Sharif, A. (2010). Common multifractality in the heart rate variability and brain activity of healthy humans. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 20(2), 023121. <https://doi.org/10.1063/1.3427639>
- Li, P., Pan, Q., Jiang, S., Yan, M., Yan, J., & Ning, G. (2021). Development of novel fractal method for characterizing the distribution of blood flow in multi-scale vascular tree. *Frontiers in Physiology*, 12, 12. <https://doi.org/10.3389/fphys.2021.711247>
- Long, Y., Chen, Y., & Zheng, N. (2021). Multifractal scaling analyses of urban street network structure: The cases of twelve megacities in China. *PLoS ONE*, 16(2), e0246925. <https://doi.org/10.1371/journal.pone.0246925>
- Lowie, W. (2013). Dynamic Systems Theory approaches to second language acquisition. In C. Chapelle (Ed.), *The Wiley-Blackwell Encyclopedia of Applied Linguistics*. Blackwell Publishing Ltd. <https://doi.org/10.1002/97811405198431.wbeal0346>
- Lowie, W., Plat, R., & de Bot, K. (2014). Pink noise in language production: A nonlinear approach to the multilingual lexicon. *Ecological Psychology*, 26(3), 216–228. <https://doi.org/10.1080/10407413.2014.929479>
- Mandelbrot, B. (1967). How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science*, 156(3775), 636–638. <https://doi.org/10.1126/science.156.3775.636>
- Mandelbrot, B. (1983). *The fractal geometry of nature*. W.H. Freeman.
- Mazdayasna, G., & Firouzi, M. (2013). A corpus based study of adjectives in literary and technical texts. *Journal of Foreign Language Teaching and Translation Studies*, 2(1), 73–88. <https://doi.org/10.22034/efl.2013.79182>
- Menzerath, P. (1954). *Die architektonik des deutschen wortschatzes*. Dümmler.
- Montag, J. L., & MacDonald, M. C. (2015). Text exposure predicts spoken production of complex sentences in 8- and 12-year-old children and adults. *Journal of Experimental Psychology General*, 144(2), 447–468. <https://doi.org/10.1037/xge0000054>
- Mouzourides, P., Kyprianou, A., & Neophytou, M. K. A. (2021). Exploring the multi-fractal nature of the air flow and pollutant dispersion in a turbulent urban

- atmosphere and its implications for long range pollutant transport. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 31(1), 013110. <https://doi.org/10.1063/1.5123918>
- Najafi, E., Darooneh, A., & Esteban, F. J. (2015). The fractal patterns of words in a text: A method for automatic keyword extraction. *PLoS ONE*, 10(6), e0130617. <https://doi.org/10.1371/journal.pone.0130617>
- Rahman, R. (2012). *Semantico-syntactic features of oscar wilde's fairy tales* [Doctoral thesis, University of Peshawar]. Pakistan. <http://pr.hec.gov.pk/jspui/handle/123456789/1324>
- Saeedimoghaddam, M., Stepinski, T., & Dmowska, A. (2020). Rényi's spectra of urban form for different modalities of input data. *Chaos, Solitons & Fractals*, 139, 109995. <https://doi.org/10.1016/j.chaos.2020.109995>
- Sanada, H. (2016). The Menzerath-Altmann Law and sentence structure. *Journal of Quantitative Linguistics*, 23(3), 256–277. <https://doi.org/10.1080/09296174.2016.1169850>
- Thompson, P., & Sealey, A. (2007). Through children's eyes?: Corpus evidence of the features of children's literature. *International Journal of Corpus Linguistics*, 12(1), 1–23. <https://doi.org/10.1075/ijcl.12.1.03tho>
- Verspoor, M. & Sauter, K.(2000). *English sentence analysis: An introductory course*. John Benjamins Publishing Company.