https://doi.org/10.51483/IJPAMR.5.2.2025.1-28 ISSN: 2789-9160

SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Pure and Applied Mathematics Research

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Equivalent Processors Modelling the Short-Term Memory

Emilio Matricciani[1*]

[1]Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano, Milan, Italy. E-mail: e-milio.matricciani@polimi.it

Article Info

Volume 5, Issue 2, October 2025
Received : 18 June 2025
Accepted : 03 October 2025
Published : 20 October 2025
doi: 10.51483/IJPAMR.5.2.2025.1-28

Abstract

In the present article, we have further developed the model on the input-output equivalent characteristics of the short-term memory—evaluated form alphabetical texts—by including a processor that memorizes syllables to produce a word. The final model is made of three equivalent processors in series, which independently process: (1) syllables to make a word; (2) words to make a word interval; (3) word intervals to make a sentence. This is a simple but useful approach because the multiple processing of the brain regarding speech/texts is not yet fully understood but syllables, characters, words can be studied in any alphabetical language. We have considered modern translations of the New Testament because these texts address general audiences accustomed to reading common words, not specialized literature. The deep-language parameters—linked to syllables, characters, words and interpunctions—of indistinct human readers/writers are specifically defined with probability distribution functions that provide useful ranges on the communication digital codes that humans can further invent. Their application to a universal readability formula can provide also an estimate of the distribution of readers and texts that these readers can read with a given "built-in" difficulty, as a function of their schooling. Finally, the theory can inspire new research lines in cognitive science studies.

Keywords: *Balto-Slavic languages, Language processing, Deep-language, Germanic languages, Greek, Latin, New testament, Romance languages, Short-term-memory, Translation, Uralic languages*

1. Introduction

1.1. Can the Short-Term Memory be Modelled by Three Equivalent Processors?

Humans can communicate and extract meaning both from spoken and written language. Whereas the sensory processing pathways for listening and reading are distinct, listeners and readers appear to extract very similar information about the meaning of a narrative story heard or read, because the brain assimilates a written text like the corresponding spoken/heard text (Deniz *et al.*, 2019). In the following, therefore, we consider the processing of reading a text—and also writing that text because a writer is also a reader of his/her own text—due to the same brain activity. In other words, the human brain represents semantic information in a modal form, independently of input modality.

* *Corresponding author: Emilio Matricciani, Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano, Milan, Italy. E-mail: e-milio.matricciani@polimi.it*

How the human brain analyzes the parts of a sentence (parsing) and describes their syntactic roles is still a major question in cognitive neuroscience. In References (Matricciani, 2024a and 2024b), we proposed that a sentence is elaborated by the Short-Term Memory (STM) with two independent processing units in series (equivalent processors), with similar size. The clues for conjecturing this input-output model emerged by considering many novels belonging to the Italian and English literatures. In Reference (Matricciani, 2024b), we showed that there are no significant mathematical/statistical differences between the two literary corpora, according to the so-called surface deep-language parameters, suitably defined. In other words, the mathematical structure of alphabetical languages—communication digital codes created by the human mind—seems to be deeply rooted in humans, independently of the particular language used or historical epoch (Matricciani, 2019; Matricciani, 2022; Matricciani, 2023a, 2023b and 2023c; Matricciani, 2024; Matricciani, 2025) and they can give clues to model the input-output characteristics of a complex and partially inaccessible mental process still largely unknown.

The first processor was linked to the number of words between two contiguous interpunctions, variable indicated by I_P—termed word interval (Appendix A lists the mathematical symbols used in the present article)—approximately ranging in Miller's 7 ± 2 law range (Flesch, 1948; Miller, 1955; Crowder, 1993; Lisman and Idiart, 1995; Cowan, 2000; Bachelder, 2001; Saaty and Ozdemir, 2003; Burgess and Hitch, 2006; Richardson, 2007; Mathy and Feldman, 2012; Gignac, 2015; Trauzettel-Klosinski and Dietz, 2012; Melton, 1963; Atkinson and Shiffrin, 1971; Murdock, 1972; Baddeley *et al.*, 1975; Case *et al.*, 1982; Grondin, 2000; Pothos and Joula, 2000; Conway *et al.*, 2002; Jonides *et al.*, 2008; Barrouillest and Camos, 2012; Potter, 2012; Jones and Macken, 2015; Chekaf *et al.*, 2016; Norris, 2017; Houdt *et al.*, 2020; Islam *et al.*, 2023; Rosenzweig *et al.*, 1993; Kaminski, 2017). The second processor was linked to the number M_F of I_P's contained in a sentence, referred to as the extended Short-Term Memory (STM), ranging approximately from 1 to 6. These two units can process a sentence containing approximately a number of words from 8.3 to 61.2, values that can be converted into time by assuming a reading speed. This conversion gives 2.6 ~19.5 seconds for a fast-reading reader, and 5.3 ~ 30.1 seconds for a reader of novels, values well supported by experiments.

The E-STM must not be confused with the intermediate memory (Rosenzweig *et al.*, 1993; Kaminski, 2017). It is not modelled by studying neuronal activity—recalled below in Section 2—but only by studying surface aspects of human communication, such as words and interpunctions, whose effects writers and readers experience since the invention of writing. In other words, the processing model proposed in References (Matricciani, 2024a and 2024b) describes the "input-output" characteristics of the STM by studying the digital codes that humans have invented to communicate.

The modeling of the STM processing by two units in series has never been considered in the literature before References (Matricciani, 2024a and 2024b). Now, the literature on the STM and its various aspects is immense and multidisciplinary, but nobody—as far as we know—has ever considered the connections we found and discussed in References (Matricciani, 2019; Matricciani, 2022; Matricciani, 2023a, 2023b and 2023c; Matricciani, 2024; Matricciani, 2025). Moreover, a sentence conveys meaning, therefore the theory we are further developing in the present article might be one of the necessary starting points to arrive at *the* Information Theory that will finally include meaning.

Today, many scholars are trying to arrive at a "semantic communication" theory or "semantic information" theory, but the results are still, in our opinion, in their infancy (Strinati and Barbarossa, 2021; Shi *et al.*, 2021; Xie *et al.*, 2021; Luo *et al.*, 2022; Wanting *et al.*, 2023; Xie *et al.*, 2021; Bellegarda, 2000; D'Alfonso, 2011; Zhong, 2017). These theories, as those concerning the STM, have not considered the main "ingredients" of our theory, namely I_P and M_F, parameters that anybody can easily calculate and understand in any alphabetical language, as a starting point for including meaning, still a very open issue.

The aim of the present article is to develop further the model on the equivalent input-output characteristics of the STM by including another processor, with a small capacity (from 1 to about 3 ~ 4 cells), set in series before the two processors already mentioned. This first processor memorizes syllables therefore its output is a word.

In other terms, in the present article we propose what we think is a more complete apparent input-output model of the STM made with three equivalent processors in series, which independently process: (1) syllables to make a word; (2) words and interpunctions to make a word interval; (3) word intervals to make a sentence.

Figure 1 sketches the signal flow chart of the proposed three processors. Syllables S_1, S_2, … S_k are stored in the first processor—from 1 to about 3 ~ 4 items, as shown in the following—until a space (a pause in speaking or reading) or an

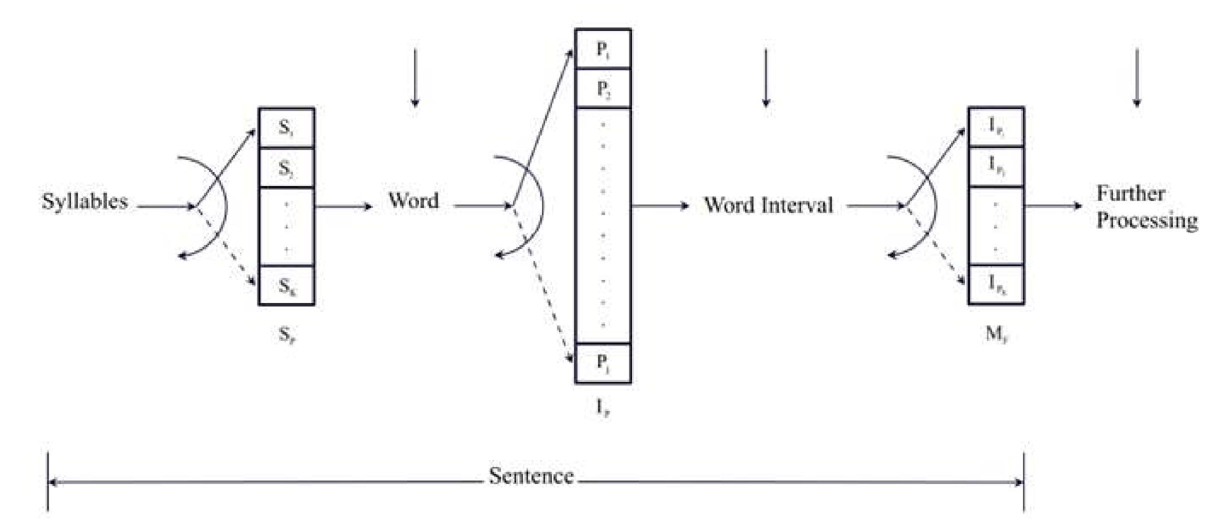


Figure 1: Flow-Chart of Three Processing Units (Processors) of Human STM. Syllables S_1, S_2, … S_k are Stored in the First Processor, from 1 to about 3 ~ 4 Items, Until a Space or an Interpunction is Introduced (Vertical Arrow) to Fix the Length of the Word. Words P_1, P_2, … P_j are Stored in the Second Processor—Approximately in Miller's Range—Until an Interpunction (Vertical Arrow) is Introduced to Fix the Length of I_P. The Word Interval I_P is then Stored in the Third Processor, from about 1 to 6 Items, Until the Sentence Ends with a Full Stop, a Question Mark or an Exclamation Mark (Vertical Arrow). The Process is then Repeated for the Next Sentence.

interpunction is introduced (vertical arrow) to fix the length of the word. Words P_1, P_2, … P_j are stored in the second processor—approximately in Miller's range—until an interpunction or a pause (vertical arrow) is introduced to fix the length of I_P. The word interval I_P is then stored in the third processor—from about 1 to 6 items—until the sentence ends with a full stop (or a pause), a question mark or an exclamation mark (vertical arrow). The process is then repeated for the next sentence.

Since the second processor (modelled by I_P) and the third processor (modelled by M_F) are discussed in detail in References (Matricciani, 2024a and 2024b), the purpose of the present article is mainly to study and define the statistical characteristics of the newly added first processor and to confirm the presence of the second and third processors.

However, since syllables are fundamental to conjecture and to dimension the capacity of the new processor, Section 2 summarizes what is mostly known about the brain processing of syllables according to cognitive neuroscience.

The article will then proceed as follows. Section 3 recalls and defines the deep-language parameter. Section 4 reports statistics of syllables and their connection with characters and words. Section 5 shows the statistical independence of S_P, I_P and M_F. Section 6 defines the details of universal input-output STM model drawn in Figure 1. Section 7 examine the human difficulty in reading texts, measured by a universal readability formula and finally Section 8 summarizes the main conclusion of the article and proposes new investigations. Appendixes A-F report useful material and further data.

2. Syllables Brain Processing

Humans can produce a remarkably wide array of word sounds to convey specific meanings. To produce fluent speech, however, a structured succession of processes is involved in planning the arrangement and structure of phonemes in individual words (Levelt *et al.*, 1999; Kazanina *et al.*, 2018; Khanna *et al.*, 2024). These processes are thought to occur rapidly during natural speech in the language network involved in word planning (Bohland and Guenther, 2006; Basilakos *et al.*, 2017; Lee *et al.*, 2018; Glanz *et al.*, 2021; Fedorenko *et al.*, 2016) and sentence construction (Nelson *et al.*, 2017; Walenski *et al.*, 2019; Rosen, 1992).

Scholars of cognitive neuroscience try to understand speech-processing units in the brain through non–invasive means. They have reached the conclusion that speech processing works on multiple hierarchical levels, which involves the coexistence of various higher-order processing windows, all of which are important.

Speech information is conveyed at least in two specific timescales involving syllabic and phonemic information. Syllabic information extends over relatively long interval (~200 milliseconds), whereas phonemic information extends in shorter intervals (~50 milliseconds) (Poeppel, 2003; Saberi and Perrott, 1999). These two timescales define the upper and lower boundaries of speech intelligibility [62]. To understand speech, the brain must extract and combine information at both timescales (Saberi and Perrott, 1999; Ghitza, 2011; Doelling *et al.*, 2014). Whereas the neural dynamics underlying the extraction of acoustic information at the syllabic time scale has received extensive attention, the precise neural mechanisms underlying the parsing of information at these two timescales are not fully understood (Peelle *et al.*, 2013; Luo and Poeppel, 2007; Park *et al.*, 2023; Oganian and Chang, 2019; Giroud *et al.*, 2023; Poeppel and Assaneo, 2020).

The investigation of the syllabic timescale and its functional role in speech perception is an active area of research. The main acoustic temporal modulation in the speech signal approximates the syllabic rate of the speech stream (Giroud *et al.*, 2023; Poeppel and Assaneo, 2020; Keitel *et al.*, 2018; Ding and Simon, 2014) and occurs in the so-called (neural) theta range ~3 to ~10 Hz, i.e., 3 to 10 syllables per second. During speech listening, neural activity tracks the speech rhythm flexibly. This tracking phenomenon seems to contribute to the segmentation of speech into linguistic units at the syllabic timescale (Ghitza, 2011; Lubinus *et al.*, 2023; Carreiras *et al.*, 1993). The syllabic rate is the strongest linguistic determinant of speech comprehension and, as long as the syllabic speech rhythm falls approximately within the theta range, speech comprehension is possible.

Finally, syllables are placed at an intermediate stage between letter perception and whole-word recognition. In alphabetic languages, syllables can be defined as the phonological building blocks of words. This suggests that syllables are functional units of speech and text perception. Research on visual word recognition carried out in French, Spanish and English, show that syllables are relevant processing units (Ferrand *et al.*, 1996; Ferrand *et al.*, 1997; Lundmark and Erickson, 2022).

In conclusion, the multiple processing of the brain regarding speech/texts is not yet fully understood but syllables, the basic processed constituents of words, can be easily observed and studied in any alphabetical language by reading, for example, literary texts of any language and epoch. This conclusion supports our simple input-output modelling of three processors in series, whose sizes are estimated from a large sample of alphabetical texts, as discussed in the next section.

3. Data Base of Literary Texts

To obtain clues on the first processor modelling the transformation of syllables into words—and also the confirmation of the other two processors already studied, now referred to as the second and third processor—we consider a large selection of the New Testament (NT) books—namely the Gospels according to *Matthew*, *Mark*, *Luke*, *John*, the Book of *Acts*, the *Epistle to the Romans*, the Book of *Revelation* (*Apocalypse*)—in the original Greek versions and in the translations into Latin and 35 modern languages. The sample size is large enough to give reliable statistical results, already studied in (Matricciani, 2020; Matricciani, 2023a and 2023b) to which the reader is referred for more details. Table 1 lists the languages of translation. The statistics reported are discussed in Section 3.

The rationale for studying NT texts is twofold. First, these texts are of great importance for many scholars of multiple disciplines. Although the translations are never verbatim, they strictly respect the subdivision in chapters and verses of the Greek original texts as they are fixed today (Parkes, 2016) discusses how interpunctions were introduced in the Greek original *scriptio continua*. Therefore, these texts can be studied at least at these two different levels (chapters and verses), by comparing how a deep"language variable changes from translation to translation.

Notice that "translation" means also "language" because we deal only with one translation per language—it is curious to notice that there are tens of different translations of the NT in English or in Spanish, but notice that language plays only one of the roles in translation, being the addressed linguistic culture of the audience another one. A "real translation"—the one we always read—is never "ideal", i.e., it never maintains all deep-language mathematical characteristics of the original text, because "domestication" of the foreign text (Greek in our case) prevails over foreignization (Matricciani, 2025a and 2025b).

Secondly, these translations address general audiences, with no particular or specialized linguistic culture, therefore the terms used in any language are common words so that the findings and conclusion we can reach refer to general readers, not to those used to reading specialized literature, like essays or academic articles. Therefore, our conclusions will refer to an indistinct human reader.

Table 1: Mean Value (Left Number of Column, Indicated by the Symbol <>) and Standard Deviation (Right Number, Indicated by *s*) of the Surface Deep-Language Parameters in the Indicated Language of the New Testament Books (*Matthew*, *Mark*, *Luke*, *John*, *Acts*, *Epistle to the Romans*, *Apocalypse*), Calculated from 155 Samples in Each Language. For Example, in Greek $< P_F >$ = 23.07 with Standard Deviation 6.65. The List Concerning the Genealogy of Jesus of Nazareth Reported in *Matthew* 1.1-1.17 and in *Luke* 3.23-3.38 was Deleted for Not Biasing the Statistics of Linguistic Variables. The Source of the Digital Texts Considered is Reported in (Matricciani, 2020). Languages with "*": the Statistics of S_P and C_S were Calculated for the Total Number of Samples (155 Chapters per Language). Languages with "°" the Statistics of S_P and C_S were Calculated only for Matthew (28 Chapters per Language)

Language	Language Family	P_F	I_P	C_P	M_F	S_P	C_S
		$< P_F >$ s	$< I_P >$ s	$< C_P >$ s	$< M_F >$ s	$< S_P >$ s	$< S_P >$ s
Greek	Hellenic	23.07 6.65	7.47 1.09	4.86 0.25	3.08 0.73	—	—
Latin	Italic	18.28 4.77	5.07 0.68	5.16 0.28	3.60 0.77	—	—
Esperanto	Constructed	21.83 5.22	5.05 0.57	4.43 0.20	4.30 0.76	—	—
French*	Romance	18.73 2.51	7.54 0.85	4.20 0.16	2.50 0.32	1.47 0.05	2.85 0.05
Italian*	Romance	18.33 3.27	6.38 0.95	4.48 0.19	2.89 0.40	1.90 0.09	2.35 0.04
Portuguese*	Romance	16.18 3.25	5.54 0.59	4.43 0.20	2.93 0.56	1.85 0.09	2.40 0.05
Romanian°	Romance	18.00 4.19	6.49 0.74	4.34 0.19	2.78 0.65	1.82 0.05	2.35 0.04
Spanish	Romance	19.07 3.79	6.55 0.82	4.30 0.19	2.91 0.47	1.90 0.16	2.27 0.16
Danish°	Germanic	15.38 2.15	5.97 0.64	4.14 0.16	2.59 0.33	1.43 0.04	2.89 0.03
English*	Germanic	19.32 3.20	7.51 0.93	4.24 0.17	2.58 0.39	1.29 0.06	3.29 0.12
Finnish°	Germanic	17.44 4.09	4.94 0.56	5.90 0.31	3.54 0.75	2.27 0.05	2.57 0.03
German*	Germanic	17.23 2.77	5.89 0.60	4.68 0.19	2.94 0.45	1.51 0.07	3.10 0.07
Icelandic	Germanic	15.72 2.58	5.69 0.67	4.34 0.18	2.77 0.39	—	—
Norwegian°	Germanic	15.21 1.43	7.75 0.84	4.08 0.13	1.98 0.22	1.41 0.03	2.89 0.04
Swedish°	Germanic	15.95 2.17	8.06 1.35	4.23 0.18	2.01 0.31	1.50 0.04	2.80 0.05
Bulgarian	Balto−Slavic	14.97 2.61	5.64 0.64	4.41 0.19	2.67 0.43	—	—
Czech°	Balto−Slavic	13.20 3.10	4.89 0.65	4.51 0.21	2.71 0.61	1.80 0.05	1.80 0.05
Croatian°	Balto−Slavic	15.32 3.54	5.62 0.75	4.39 0.22	2.72 0.49	1.87 0.06	2.34 0.04
Polish°	Balto−Slavic	12.34 1.93	4.65 0.43	5.10 0.22	2.67 0.40	1.95 0.06	2.60 0.05
Russian	Balto−Slavic	17.90 4.46	4.28 0.46	4.67 0.27	4.18 0.92	—	—
Serbian	Balto−Slavic	14.46 2.42	5.81 0.69	4.24 0.20	2.50 0.39	—	—
Slovak	Balto−Slavic	12.95 2.10	5.18 0.61	4.65 0.23	2.51 0.36	—	—
Ukrainian	Balto−Slavic	13.81 2.18	4.72 0.41	4.56 0.26	2.95 0.58	—	—
Estonian°	Uralic	17.09 3.89	5.45 0.66	4.89 0.24	3.14 0.64	1.86 0.05	2.61 0.03
Hungarian°	Uralic	17.37 4.54	4.25 0.45	5.31 0.29	4.09 0.93	2.15 0.07	2.47 0.03
Albanian	Albanian	22.72 4.86	6.52 0.78	4.07 0.22	3.48 0.61	—	—
Armenian	Armenian	16.09 3.07	5.63 0.52	4.75 0.40	2.86 0.47	—	—
Welsh	Celtic	24.27 4.75	5.84 0.44	4.04 0.15	4.16 0.76	—	—
Basque	Isolate	18.09 4.31	4.99 0.52	6.22 0.27	3.63 0.81	—	—
Hebrew	Semitic	12.17 2.04	5.65 0.59	4.22 0.17	2.16 0.33	—	—
Cebuano	Austronesian	16.15 1.71	8.82 1.01	4.65 0.10	1.85 0.22	—	—
Tagalog	Austronesian	16.98 3.24	7.92 0.82	4.83 0.17	2.16 0.44	—	—
Chichewa	Niger−Congo	12.89 1.79	6.18 0.87	6.08 0.18	2.10 0.25	—	—
Luganda	Niger−Congo	13.65 2.78	5.74 0.82	6.23 0.23	2.39 0.40	—	—
Somali	Afro−Asiatic	19.57 5.50	6.37 1.01	5.32 0.16	3.06 0.65	—	—
Haitian	French Creole	14.87 1.83	6.55 0.71	3.37 0.10	2.28 0.26	—	—
Nahuatl	Uto−Aztecan	13.36 1.70	6.47 0.91	6.71 0.24	2.08 0.24	—	—

For our analysis, as done in Reference (Matricciani, 2020), we have chosen the chapter level because the amount of text is sufficiently large (155 chapters per language) to assess reliable statistics. Therefore, for each translation/language we have a database of 155 x 37 = 5735 samples, sufficiently large to give reliable statistical results on deep-language parameters, which we briefly recall, or define, in the next section for reader's benefit.

4. Deep-Language Parameters

Let n_C, n_W, n_S, n_I and n_{I_P} be respectively the number of characters, words, sentences, interpunctions and word intervals per chapter. The surface deep-language parameters, defined in Ref. (Matricciani, 2024a), are here recalled.

Number of characters per word, C_P:

$$\boldsymbol{C_P} = \frac{n_C}{n_W} \qquad ...(1)$$

Number of words per sentence, P_F:

$$\boldsymbol{P_F} = \frac{n_W}{n_S} \qquad ...(2)$$

Number of interpunctions per word, referred to as the word interval, I_P:

$$\boldsymbol{I_P} = \frac{n_I}{n_W} \qquad ...(3)$$

Number of word intervals per sentence, M_F:

$$\boldsymbol{M_F} = \frac{n_{I_P}}{n_S} \qquad ...(4)$$

Defined the number of syllables n_{Sy} per chapter, to the previous parameters we add the new ones. Number of syllables per words, S_P:

$$\boldsymbol{S_P} = \frac{n_{Sy}}{n_W} \qquad ...(5)$$

Number of characters per syllables, C_S:

$$\boldsymbol{C_S} = \frac{n_C}{n_{Sy}} \qquad ...(6)$$

Notice that $S_P \times C_S = C_P$.

Table 1 reports mean value and standard deviation of these parameters. We have calculated them from the digital texts (WinWord files) in the following manner: for each chapter we have counted the number of characters, words, sentences, interpunctions and syllables. Before doing so, however, we have deleted the titles, footnotes and other extraneous material present in the digital texts, which, for our analysis, can be considered as "noise".

The count is very simple, although time-consuming and no understanding of the language is required. For each text block, WinWord directly provides the number of characters, words and sentences. The number of sentences, however, was first calculated by replacing periods with periods (full stops): of course, this action does not change the text, but it gives the number of these substitutions, therefore the number of periods. The same procedure was done for question marks and exclamation marks. The sum of the three totals gives the total number of sentences of the text block. The same procedure gives the total number of commas, colons and semicolons. The sum of these latter values with the total number of sentences gives the total number of interpunctions.

As for the syllables, we have calculated their number per chapter by using the website wordcount.com (last access June 18, 2025). Unfortunately, this website does not provide the count for all languages listed in Table 1, therefore we had to limit our investigation on syllables only to the languages available, signaled in Table 1. Moreover, the count was done for the total number of chapters of the database (5735 samples per language) only for some languages (see Table 1). For the other languages, to save processing time, we have considered only *Matthew* (28 samples) because we have noticed that *Matthew* alone can give reliable results on syllable statistics, as Table 2 shows. From Table 2 we can notice that the error is of the order of 1% for means; for standard deviations, the error is larger but it has very little impact in the following. In conclusion, we can merge the two databases for syllables studies.

Table 2: Comparison between Mean Values of $< S_P >$ and $< C_S >$ for *Matthew* (28 Samples per Language) and the NT Database (155 Samples per Language) in the Indicated Languages. The Standard Deviation is Reported in Parentheses

Language	$< S_P >$		$< C_S >$	
	Matthew	NT	*Matthew*	NT
French	1.46 (0.04)	1.47 (0.05)	2.86 (0.05)	2.85 (0.05)
Italian	1.89 (0.05)	1.90 (0.08)	2.27 (0.04)	2.35 (0.04)
Spanish	2.19 (0.06)	1.90 (0.16)	1.94 (0.02)	2.27 (0.16)
Portuguese	1.84 (0.07)	1.85 (0.08)	2.42 (0.04)	2.40 (0.05)
English	1.27 (0.04)	1.29 (0.06)	3.29 (0.06)	3.29 (0.11)
German	1.50 (0.04)	1.51 (0.07)	3.10 (0.06)	3.10 (0.07)

As observed in Reference (Matricciani, 2023), the languages belonging to the Romance family (mostly derived from Latin) show very similar values, with the exception of Portuguese; the same observation can be done for the Balto-Slavic, Niger-Congo and Austronesian languages. Greek is largely diverse from all other languages.

English and French, although attributed to different families, almost coincide. This coincidence can be partially explained by the fact that many English words and several sentence structure come French and/or from Latin—English contains up to 65% of Latinisms, i.e., words of Latin and Old French origin—a language from which romance languages derive.

In the next section, we will explore the statistical relationships of syllables with other linguistic variables. This investigation is useful to arrive at defining the features of the first processor.

5. Exploratory Data Analysis of Syllables

In this section we explore the relationships between syllables, words and characters—for the relationships involving the other deep-language parameters, see Refs. (Matricciani, 2019; Matricciani, 2022; Matricciani, 2023a, 2023b and 2023c; Matricciani, 2024; Matricciani, 2025; Matricciani, 2020)—for French, Italian, Portuguese, English and German, namely the languages for which we consider the full data banks (NT). Notice, however, that the conclusion drawn for these languages are valid also for the others, because similar results can be reported for *Matthew* only, not shown for brevity.

Figure 2 shows the scatterplot between words and syllables (Figure 2a), and between characters and syllables (Figure 2b) for English. We can see a very tight relationship, measured by a very high correlation coefficient, as Table 3 shows.

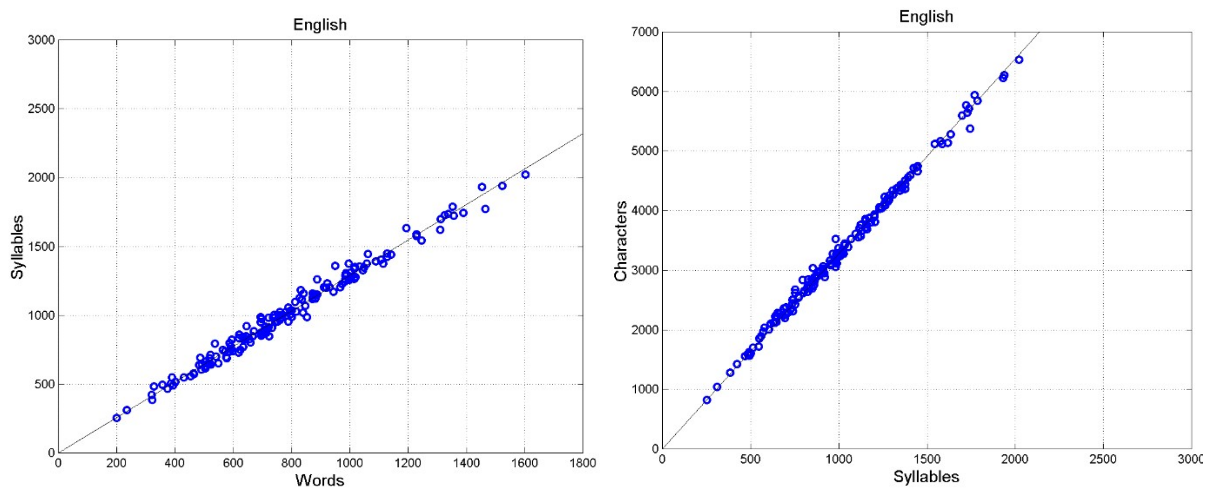


Figure 2: English: (a) Scatterplot between Words and Syllables; (b) Scatterplot between Characters and Syllables (155 Samples)

Table 3: Slope a and Correlation Coefficient r of the Linear Relationship $y = ax$, for the Indicated Language. The Correlation Coefficients are Reported with Four Digits Because they are All Similar

Language	Syllables versus Words		Characters versus Syllables	
	Slope a	Correlation Coefficient r	Slope a	Correlation Coefficient r
French	1.470	0.9949	2.849	0.9990
Italian	1.905	0.9912	2.347	0.9990
Spanish	1.910	0.9703	2.242	0.9729
Portuguese	1.850	0.9909	2.401	0.9987
English	1.289	0.9930	3.275	0.9968
German	1.509	0.9914	3.087	0.9984

Figure 3 shows the scatterplot between S_P and words (i.e., the ratio between ordinate and abscissa of the samples shown in Figure 2a), and between C_S and characters (the ratio between ordinate and abscissa of the samples shown in Figure 2b). The spread around the mean value is very small. The same results are found for the other languages, as it can be seen in Appendix B. Table 3 reports slope and correlation coefficient for the indicated languages, which clearly indicate that the linear relationship between syllables and words, or between characters and syllables, is very tight.

In fact notice that $C_P \approx aS_P$, with a constant, and very large correlation coefficient, hence very small spread (see Table 3). For example, in English $a = 3.275$, $r = 0.9968$, therefore the coefficient of variation $r^2 = 0.9968^2 = 0.9936$, i.e., 99.36% of the variance of the characters is due to the linear relationship and only 0.64% to scattering ("noise" compared to the deterministic linear relationship).

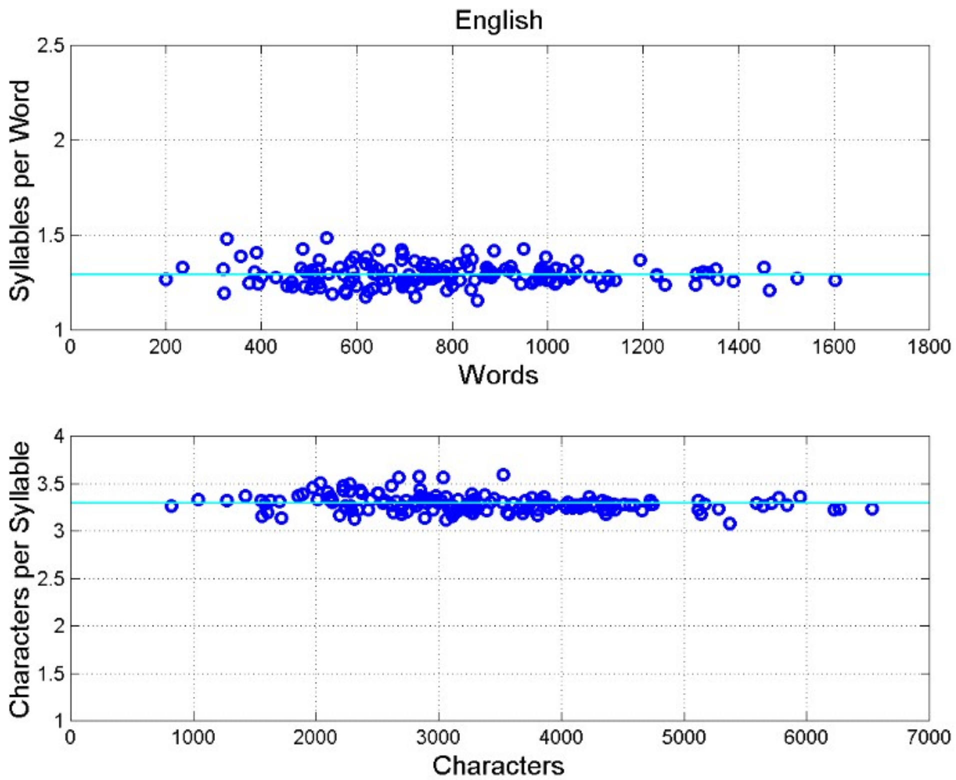


Figure 3: English: Upper Panel: Scatterplot between S_P and Words; Lower Panel: Scatterplot between C_S and Characters (155 Samples). The Mean Value is Drawn with the Cyan Line

In other words, syllables are so tightly correlated with characters that we can use these latter to assess statistical properties. Figure 4 shows a modified first processor that considers now the relationship between syllables and characters: a syllable is coded with characters C_1, C_2, …, C_n—from 1 to about 3 ~ 4 items, see Table 3, until a space or an interpunction is found (vertical line).

For example, the scatterplot of S_P versus other linguistic variables must have approximately the same statistical properties of the scatterplot of C_P versus these variables. Therefore, we can estimate correlation coefficients involving S_P also for the languages for which syllables count is not available (Table 1), by considering C_P.

It is useful to notice that since syllables and characters are extremely correlated, S_P is rarely considered as an ingredient of readability formulae. The historical founder of studies on readability formulae—Rudolf Flesch (1911-1986)—he first did consider also syllables for measuring the difficulty of reading a text, but he ended up in replacing S_P with C_P in his famous readability formula, Flesch's formula for English (Flesch, 1974). Also the universal readability formula proposed in Matricciani (2023) does use C_P, not S_P.

In the next section, we model the deep-language variables to confirm the modelling of the second and third processors defined in References (Matricciani, 2024a and 2024b)—established by studying English and Italian Literatures—and for defining the first processor, by studying the NT texts in the languages listed in Table 1.

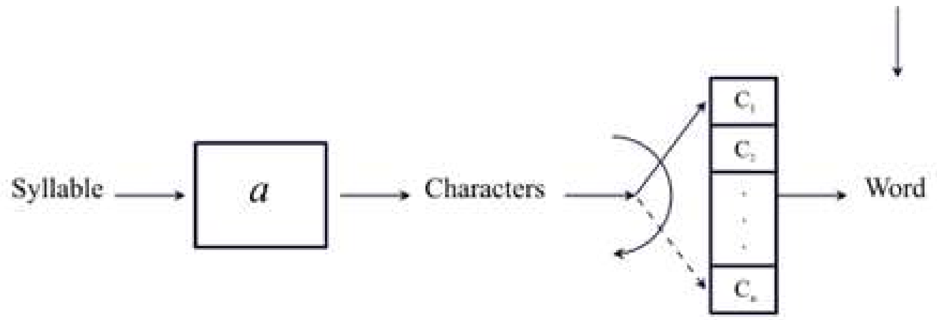


Figure 4: Flow-Chart of the First STM Processor Linking Syllables to Characters. The Multiplying Factor is Given in Table 3 for Some Languages

6. Statistical independence of S_P, I_P and M_F

From the mean values and standard deviations reported in Table 1 (referred to as conditional values) we can obtain the mean value and standard deviation (unconditional values) of the entire data base of a deep-language parameter. According to standard statistical theory, the unconditional mean is given by the mean of means; the unconditional variance is given by the variance of the means added to the mean of the conditional variances (see Appendix C). Notice that in adding variances, since the conditional variances are much smaller than the variances due to means (see Table 1) these latter practically determine the value of the reported standard deviation.

Now, the deep-language parameters can be modelled with a three-log-normal probability density function (Papoulis, 1990). This modelling applies also to the newly defined S_P and C_S. For example, Figure 5 shows the histogram of the pooled data of Italian and Portuguese and its modelling with a three-log-normal probability density function (whose values can be calculated from Table 1, see Appendix D). C_S can be modelled similarly.

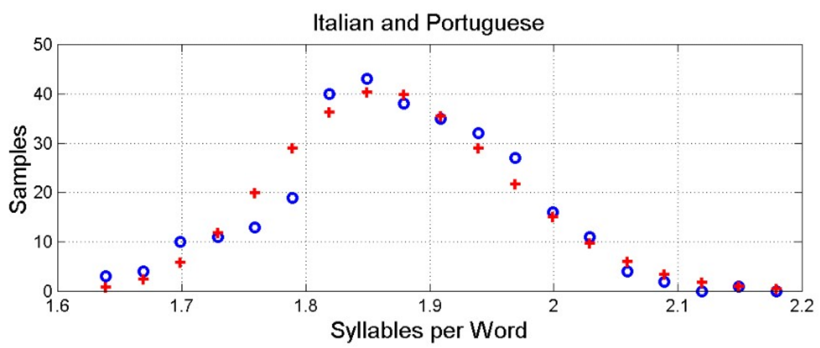


Figure 5: Histogram of S_P and Three–Parameter Lognormal Model, Pooled Data of Italian and Portuguese (2 x 155 = 310 Samples). Blue Circles Indicate the Number of Experimental Samples, Red Crosses Indicate the Number of Samples According to the Lognormal Model of the Pooled Data

Now a good model of the joint (bivariate) probability density function between two deep-language parameters is also lognormal, at least for the central part of the joint distribution. In this case, if the correlation coefficient between the logarithms of two variables is zero, then these variables not only are uncorrelated but they are also independent (Papoulis, 1990). We hypothesized this joint probability density function in References (Matricciani, 2025a and 2025b) to model the STM processing of a sentence with two equivalent and independent processors in series, the first memorizing I_P, the second memorizing M_F. Now, we show that the same modelling can be applied to the texts of any couple of the languages listed in Table 1. For this purpose it suffices to show that S_P (or C_P first processor), I_P (second processor) and M_F (third processor) are approximately uncorrelated.

Figure 6 shows, for English, the scatterplots between S_P, I_P and M_F. The correlation coefficients between the logarithms of the variables are 0.0858, –0.2650 and 0.3320 respectively. It is evident that any couple of variables are practically uncorrelated (correlation coefficients for all languages are reported in Appendix E), therefore they can be considered also practically independent because the bivariate density is modelled as log-normal. Scatterplots concerning other languages, reported in Appendix F, show the same spread.

Figure 7 shows, for the original Greek texts for which we have no syllables count, the scatterplot between C_P, I_P and M_F. Now, C_P stands for S_P because in general $C_P \approx aS_P$ so that the correlation coefficients involving C_P, are good estimates of the values that would be found if syllables counts were available. The correlation coefficients between the logarithms of the variables are 0.3176, –0.0988 and 0.0686 respectively.

In synthesis, we can reliably assume that S_P, I_P and M_F are independent stochastic variables in modeling the equivalent input-output processors of the STM, independently of language and epoch.

Now, this multiple independence seems reasonable because it says that: (a) the number of words making the word interval I_P is not related to the number of syllables making any word contained in the word interval; (b) the number of word intervals, i.e., M_F, making a sentence is not strictly related to the number of words making any word interval. The latter two results were shown, as recalled, for the Italian and English Literature.

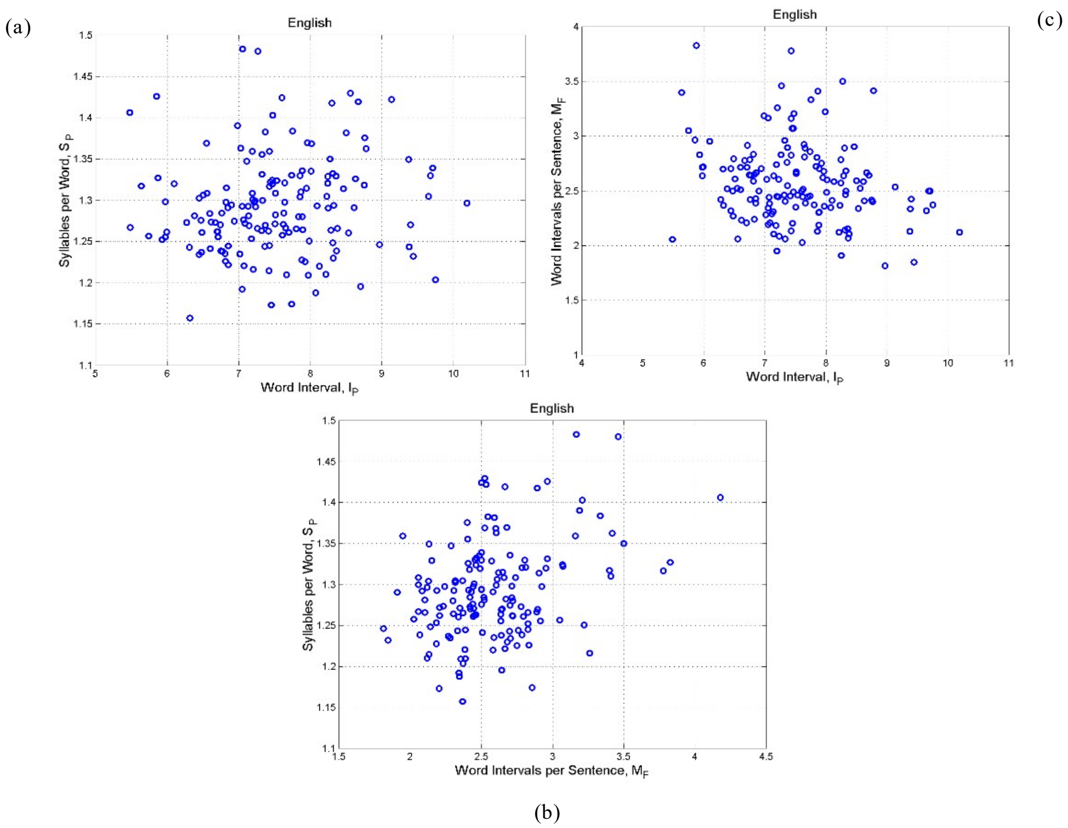


Figure 6: English: (a) Scatterplot between S_P and I_P, Correlation Coefficient between the Logarithms of the Variables is 0.0858; (b) Scatterplot between M_F and I_P, –0.2650; (c) Scatterplot between S_P and M_F, 0.3320

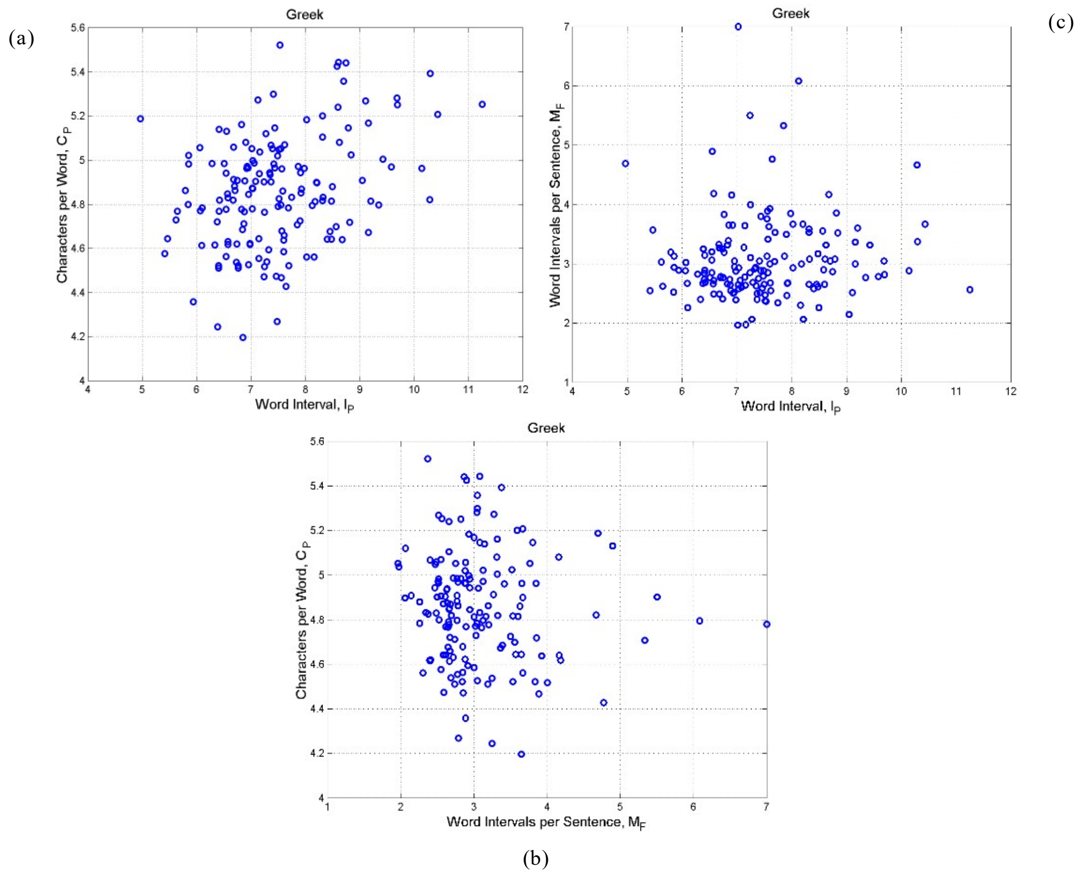


Figure 7: Greek: (a) Scatterplot between and , the Correlation Coefficients between the Logarithms of the Variables is 0.3176; (b) Scatterplot between M_F and I_p, –0.0988; (c) Scatterplot between C_p and M_F, 0.0686

In conclusion, based on the codes invented by humans (at least for the languages listed in Table 1) we can estimate the input-output equivalent STM processors shown in Figures 1, 4 for any language. For example, in Italian, on the average the first processor—syllables to a word—should memorize 1.90 ± 0.09 syllables (or 2.35 ± 0.04 characters, Table 1); the second processor—words and interpunctions to make a word interval—should memorize 6.38 ± 0.95 words; the third processor—word intervals to make a sentence—should memorize 2.89 ± 0.40 sequences of words.

The reader of the present article can assess for his/her language of interest the size of the three processors by looking at Table 1, but it is more interesting to define the STM input-output processors of humans, independently of language. This exercise is done in the next section.

7. Universal Input-Output Model of STM

From the mean values and standard deviations (conditional values) concerning the languages listed in Table 1, we can calculate the unconditional log-normal probability distributions of the linguistic variables, as discussed in Section 6 and Appendix D. This exercise is useful to define the statistical characteristics of the "universal" human STM memory, according, of course, to the simple input-output model shown in Figures 1 and 4. As recalled above, we use the term "universal" because the readers of the NT texts can be any person.

Table 4 reports fundamental statistics of the deep-language parameters and Figures 8-10 show their log-normal probability densities. Table 5 reports probability ranges corresponding to ±1, ±2, ±3 standard deviations of the log-normal models (Table 4).

From Table 5, we expect 99.73% of the samples to fall in the range: (a) 1.23 ~ 3.14 syllables for the first processor, or equivalently 1.93 ~ 3.80 characters, values that are well within the ranges given by neuroscience studies (Section 2); 3.22 ~ 11.63 words in the second processor, a range that includes Miller's range given by his 7 ± 2 "law"; 1.46 ~ 7.21 word intervals in the third processor.

Table 4: Fundamental Statistics of the Deep-Language Parameters

Deep-Language Parameter	μ	σ	Mode M	Median $M_{0.5}$	Mean m	Standard Deviation s
S_P	–0.358	0.373	1.61	1.70	1.75	0.29
C_S	0.479	0.184	2.56	2.62	2.64	0.31
C_P	1.297	0.199	4.62	4.66	4.73	0.75
I_P	1.581	0.261	5.79	5.86	6.03	1.33
P_F	2.716	0.286	16.04	16.12	16.76	4.60
M_F	0.526	0.434	2.50	2.69	2.86	0.85

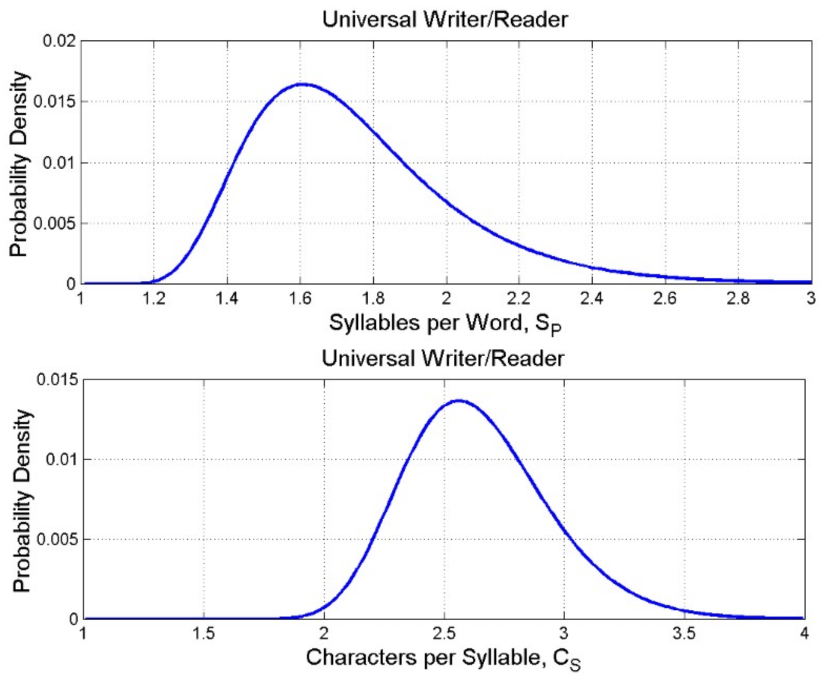


Figure 8: (a) Log-Normal Probability Density Function of S_P; (b) Log-Normal Probability Density Function of C_S

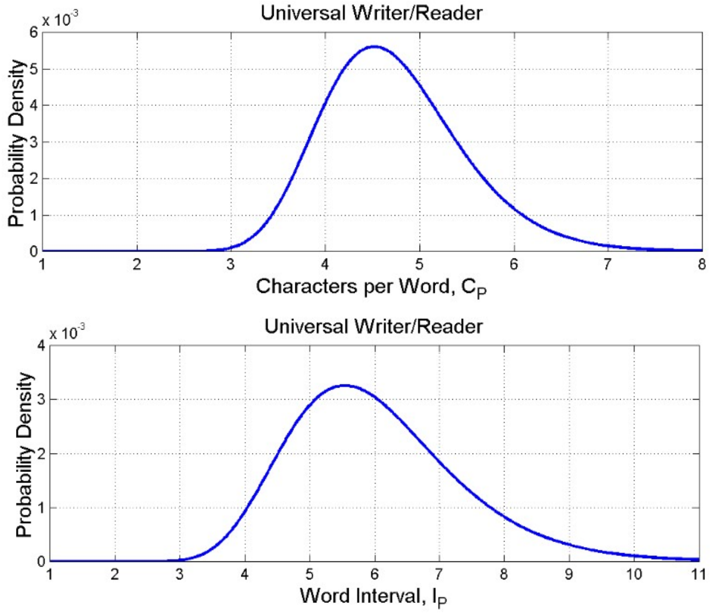


Figure 9: (a) Log-Normal Probability Density Function of C_P; (b) Log-Normal Probability Density Function of I_P

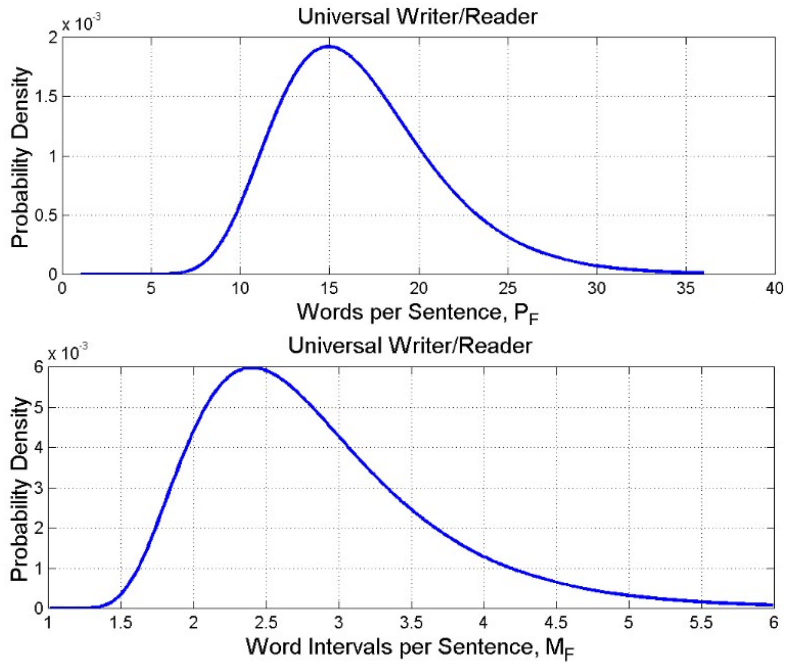


Figure 10: (a) Log-Normal Probability Density Function of P_P; (b) Log-Normal Probability Density Function of M_F

Table 5: Probability Ranges Corresponding to ±1, ±2, ±3 Standard Deviations of theLog-Normal Model (Table 4), for the Indicated Deep-Language Parameters

Probability Range (%)	S_p	C_S	C_P	I_p	P_F	M_F
68.27	1.48~2.02	2.34~2.94	4.00~5.46	4.74~7.31	12.36~21.13	2.10~3.61
95.45	1.33~2.47	2.12~3.33	3.46~6.45	3.88~9.19	9.53~27.79	1.71~5.03
99.73	1.23~3.14	1.93~3.80	3.01~7.65	3.22~11.63	7.41~36.66	1.46~7.21

It is also interesting to notice that the length of sentence, P_F, is in the range 7.41 ~ 36.66 words. In other terms, this is the number of words that humans mostly use to express a full concept. The number of concepts can be very large. For example, by considering permutations with no repetition (words are not repeated), we get the range 7! = 5040 to 36! = 3.7199 x 10^{41}, fantastic numbers. Of course, they include sequences with no meaning for the languages listed in Table 1, but they might express meaning in a language still to invent.

Finally, is also interesting to notice the range of characters that make a word, C_P, namely 3.01 ~ 7.65. Therefore, to name everything humans use a very limited number of alphabetical symbols, mostly from 3 to about 8, but the number of words that can be invented can be very large. With 26 letters, they range from 26^3 = 17576 to 26^8 = 2.0883 x 10^{11}, numbers that include, of course, sequences of letters (words) never found in the actual languages. For example, in English there are approximately 1.6 x 10^6 different words (this number refers to words that are different at least for one character), therefore there is still a large room for new words, ~10^5 times more. Similar numbers are found also for the other languages.

Now, another interesting issue can be studied. These humans, when they read a text—assuming they have the culture to understand what they read—what reading difficulty do they find? Next section addresses this issue by first recalling what readability formulae are used for and then applying a universal readability formula.

8. Human Difficulty in Reading Texts and Readability Formulae

First developed in the United States (Flesch, 1974; Kincaid *et al.*, 1975; DuBay, 2004; Bailin and Graftstein, 2001; DuBay, 2006; Zamanian and Heydari, 2012; Benjamin, 2012; Collins-Thompson, 2014; Kandel and Moles, 1958; François, 2014), a readability formula allows to design the best possible match between readers and texts. A readability formula is very attractive because it gives a quantitative judgement on the difficulty or easiness of reading a text. Every readability formula, however, gives only a partial measurement of reading difficulty because its score is mainly linked to characters, words and sentences length. It gives no clues as to the correct use of words, to the variety and richness of the literary expression, to its beauty or efficacy, to quality and clearness of ideas or give information on the correct use of grammar. The comprehension of a text—not to be confused with its readability, defined by mathematical formulae—is the result of many other factors, the most important being reader's culture. In spite of these limits, readability formulae are very useful because—besides their obvious use—they give an insight on the STM memory of readers/writers, as we suggested in References (Matricciani, 2024a and 2024b; Matricciani, 2023).

We first recall the universal readability formula (Matricciani, 2023), and secondly we find some universal human features linked to the deep-language parameters discussed in Section 3.

8.1. The Universal Readability Formula Contains the Three STM Equivalent Processors

Many readability formulae have been studied for English (Zamanian and Heydari, 2012), practically none for other languages; therefore, in Matricciani (2023) we proposed a universal readability formula—referred as the index G_U-based on a calque of the readability formula used in Italian (Lucisano and Piemontese, 1988)—applicable to any alphabetical language, given by:

$$G_U = 89 - 10kC_P + 300/P_F - 6(I_P - 6) \quad ...(7)$$

$$k = < C_{P,ITA} > / < C_{P,Lan} > \quad ...(8)$$

In Equation (8), $< C_{P,ITA} >$ is the mean value of found in Italian texts, $< C_{P,Lan} >$ is the corresponding mean value found in the language for which G_U is calculated. In Appendix A of Ref. (Matricciani, 2024) we showed that if is calculated by introducing in Equation (8) mean values, i.e., $< C_P >$, $< P_F >$ and $< I_P >$, then its value is smaller than $< G_U >$ calculated form the samples.

Very important, and contrarily to all other readability formulae, G_U is strictly connected with the three STM equivalent processors. It contains the second processor directly, trough I_P; it contains the first processor indirectly through C_P—strictly linked to S_P (Sections 4, 5); it contains the third processor trough P_F, which is very well correlated to M_F, as it was shown for Italian Literature (correlation coefficient 0.937) and English Literature (correlation coefficient 0.914) see Figure 3 of Ref. (Matricciani, 2024b).

In the present article, $< C_{P,ITA} > = 4.48$ (see Table 1); for $< C_{P,Lan} >$ we assume the unconditional mean value of a "universal reader" discussed in Section 7, hence $< C_{P,Lan} > = 4.73$, therefore in Equation (9) $k = 0.947$.

By using Equations (8) and (9), the mean value $< kC_P >$ is forced to be equal to that found in Italian, 4.48. The rationale for this choice is the following: $< C_P >$ is a parameter typical of a language (see Table 1) which, if not scaled, would bias without really quantifying the reading difficulty of readers who in their own language are used to read, on average, shorter or longer words than Italian readers do. The scaling, therefore, avoids changing G_U only because a language has, on the average, words shorter or longer than Italian.

From the probability distributions of the deep-language parameters discussed in Section 7, we can theoretically calculate the probability distribution of G_U by assuming the statistical independence of C_P, I_P and M_F. Instead of doing complex analytical calculations to arrive at a mathematical formula, we have estimated it by doing a Monte Carlo simulation of independent log-normal samples of C_P, I_P, and P_F, according to Table 4.

Figure 11 shows the probability density function (blue line) obtained with many simulations (300,000). Since $G_U > 0$ and specifically $G_U >> 0$, we can model the central part of its experimental density (blue line) as Gaussian, with the "experimental" mean value $< G_U > = 63.24$ and $S_{G_U} = 11.92$ (black line; from this model, the probability of $G_U < 0$ is negligible).

8.2. Readability Index and Universal Schooling of Humans

An interesting exercise is to link G_U to the number of school-years—i.e., the only quantitative information that can give clues on reader's reading skill—necessary to assess that a text/author is more or less difficult to read than another text/author. We do so according to the Italian school system, assumed as a common reference, Figure 12 (redrawn from Ref. (Lucisano and Piemontese, 1988)).

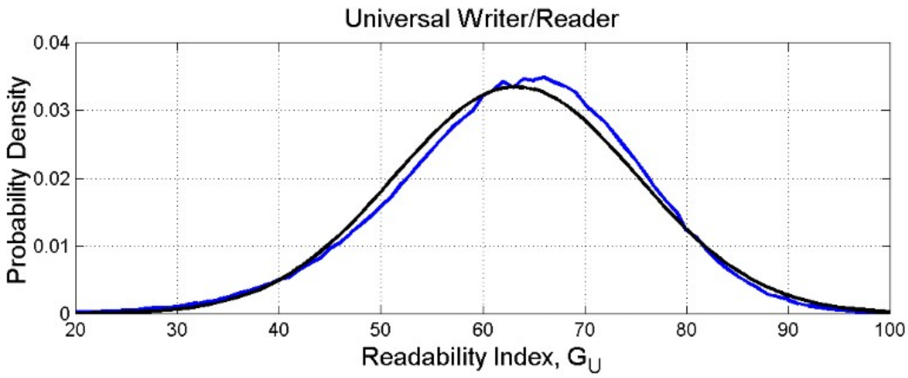


Figure 11: Experimental (Blue Line, Monte Carlo Simulations) Probability Density Function of G_U and its Gaussian Model (Black Line)

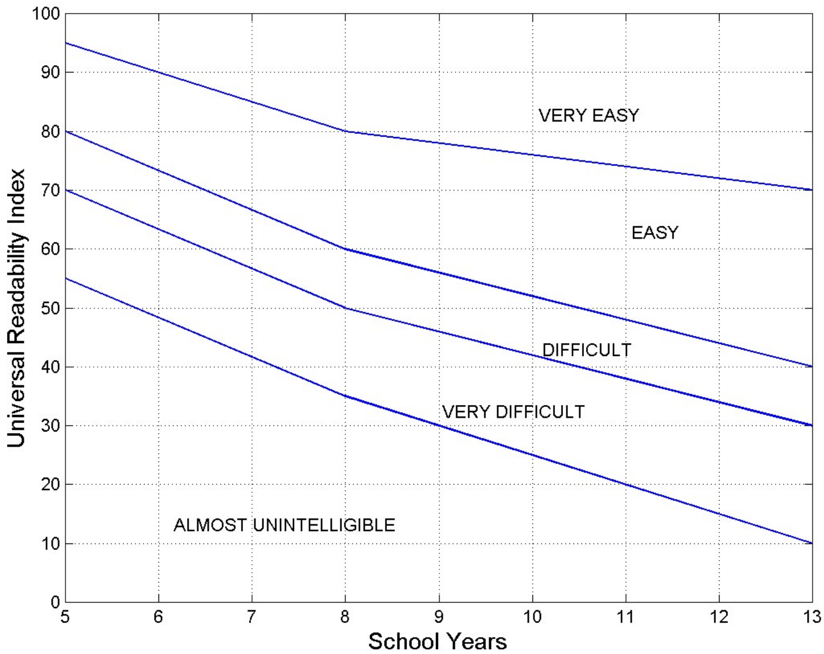


Figure 12: Chart to Estimate the Number of School-Years for a Given G_U, According to Text Reading Difficulty (Redrawn from Ref. (Lucisano and Piemontese, 1988))

This assumption does not mean, of course, that readers of any other language must attend school for the same number of years of the Italian readers, but it is only a way to do relative comparisons, otherwise difficult or impossible to do from the mere values of G_U. In other words, the number of school-years should be intended as the "equivalent" schooling and be used only as a common reference.

According to this chart, the combination of readability index and school-years that fall, for example, above the line termed "easy" corresponds to texts that the reader is likely to find "easy" or "very easy" to read. The same interpretation holds for the other cases.

According to Figure 12, texts with $G_U \approx 63$ (universal mean value) would be "easy" to read by readers of about 8 school-years. The texts at ± 1 standard deviation, namely $G_U = 63.24\text{-}11.92 = 51.32$ and $G_U = 63.24\text{+}11.92 = 75.16$, would require, respectively, about 10 and 5.5 school-years.

Now, from the probability distribution shown in Figure 11 and the "decoding/comparison" chart shown in Figure 12, we can do the following exercise: let us estimate the probability that a universal reader finds a text—extracted from the body of the universal literature written in any alphabetical language—"very easy", "easy", "difficult" or "very difficult".

Figure 13 shows the result of these exercise by reporting in ordinate the probability (%) that the readability index is greater than a certain value, $P(G_U)$—hence, this is the probability of finding texts easier to read than those of the given G_U—and in abscissa the number of school-years.

As the number of school-years increase, readers can read more texts with the same label. For example, with only 5 years of school (elementary/primary school), the reader finds 8% of the texts "easy" to read and the 92% texts left, with a continuous transition, "difficult", "very difficult" or "almost unintelligible"; with 8 school-years the percentage becomes 61% (39%) and with 13 school-years 98% (2%). In other words, with 13 school years practically all texts should be easy to read.

In conclusion, the chart shown in Figure 13 can be used for two diverse aims. The first aim is to guide a writer to establish approximately reader's school-years necessary to read the text with a "built-in" G_U; in this case, the horizontal (red) line $y = G_U$ can cross diverse areas of texts declared "very difficult", "difficult", "easy", therefore indicating a spread of readers with diverse schooling who will experience diverse difficulty. The second aim is to guide a writer to write a text with a given G_U suitable to the schooling of the intended readers (magenta line). Finally, notice that whatever the combination, it always implies considering the STM equivalent processors. In other words, every reader/writer must fall somewhere in the chart of Figure 13.

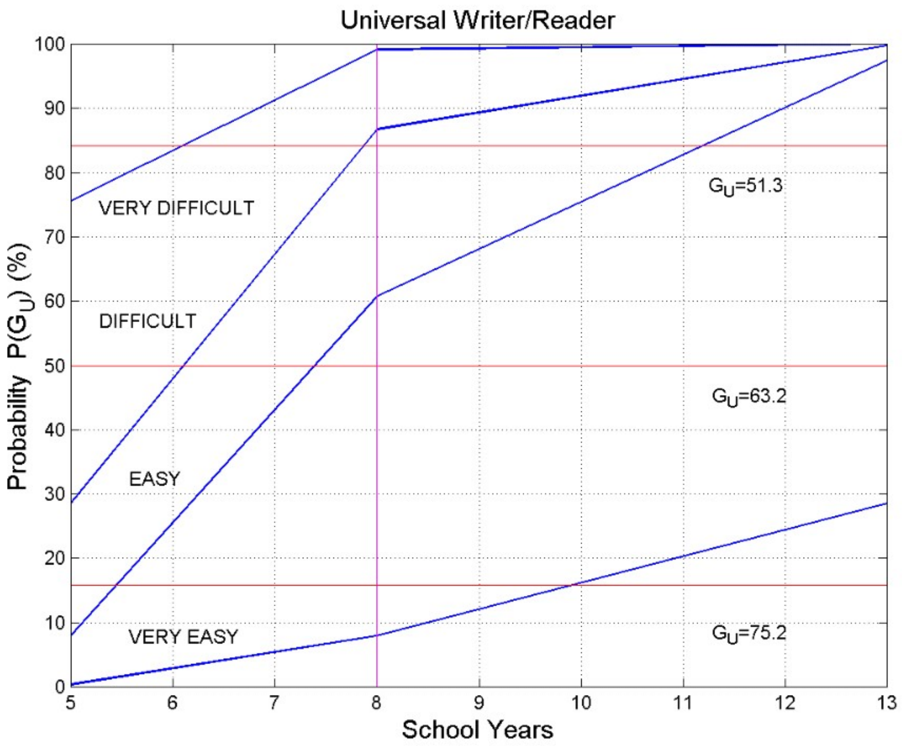


Figure 13: Probability (%) that the Readability Index is Greater Than a Certain Value, $P(G_U)$ versus Number of School-Years. Notice that in the Ordinate Scale G_U Decreases (Reading Difficulty Increases) as $P(G_U)$ Increases. Examples of Application: The Horizontal Red Line $y = G_U$—Set, in this Example, at 50% Probability (i.e., it Corresponds to $< G_U > = 63.2$)—Can Cross Texts Declared "Very Difficult", "Difficult", "Easy" as the Number of School Years Increase, Therefore Indicating that 50% of the Texts are Differently Readable by Readers of Diverse Schooling. The Other Red Lines Indicate ± 1 Standard Deviation of G_U. The Magenta Line Set, in this Example, at 8 School-Years, Shows that G_U Can Cross Percentages of Texts Declared "Very Easy", "Easy", "Difficult", "Very Difficult", "Almost Unintelligible", Therefore Indicating Percentages of Texts with Diverse Reading Difficulty

9. Conclusion

In the present article, we have further developed the model on the equivalent input-output characteristics of the STM by including a processor that memorizes syllables to produce a word. The final model is made of three equivalent processors in series, which independently process: (1) syllables to make a word; (2) words to make a word interval; (3) word intervals to make a sentence, schematically represented in Figures 1 and 4.

This is a very simple but useful approach because the multiple processing of the brain regarding speech/texts is not yet fully understood but syllables, characters, words and interpunctions—these latter used to distinguish word intervals and sentences—can be easily observed and studied in any alphabetical language. These are digital codes created by the human mind that can be fully analyzed by studying the literary texts written in any language and belonging to any historical epoch, as we have presently done with the translations of the New Testament.

We have considered these texts because they always address general audiences, with no particular or specialized linguistic culture, therefore the terms used in any language are common words because of the large degree of domestication in any translation (Matricciani, 2025). Therefore, the findings and conclusion we can draw refer to "indistinct" human readers, not to persons used to reading specialized literature, like essays or academic articles.

The deep–language parameters—linked to syllables, characters, words and interpunctions—of indistinct human readers/writers can be specifically defined with probability distribution functions that provide useful ranges on the codes that humans can invent. Their application to a universal readability formula can provide also the distribution of readers and texts that these readers can read with a given "built-in" difficulty—measured by the universal readability index—as a function of their schooling.

The main limitation of the present work is that, as is, the theory is not applicable to non-alphabetical languages. Therefore, future work should be done, very likely along the same research lines, on specialized literature (such as essays and academic texts) and, especially, on non-alphabetical languages like, for example, Chinese, Korean and Japanese. Finally, the theory can inspire new research lines in cognitive science.

Acknowledgment

The author thanks Lucia Matricciani for drawing Figures 1 and 4.

References

Atkinson, R.C. and Shiffrin, R.M. (1971). The Control of Short-Term Memory. *Scientific American*, 225(2), 82-91.

Bachelder, B.L. (2001). The Magical Number: Span Theory on Capacity Limitations. *Behavioral and Brain Sciences*, 24, 116-117. https://doi.org/10.1017/S0140525X01243921

Baddeley, A.D., Thomson, N. and Buchanan, M. (1975). Word Length and the Structure of Short-Term Memory. *Journal of Verbal Learning and Verbal Behavior*, 14, 575-589.

Bailin, A. and Graftstein, A. (2001). The Linguistic Assumptions Underlying Readability Formulae: A Critique. *Language & Communication*, 21, 285-301.

Barrouillest, P. and Camos, V. (2012). As Time Goes by: Temporal Constraints in Working Memory. *Current Directions in Psychological Science*, 413-419.

Basilakos, A., Smith, K.G., Fillmore, P., Fridriksson, J. and Fedorenko, E. (2017). Functional Characterization of the Human Speech Articulation Network. *Cereb. Cortex,* 28, 1816-1830.

Bellegarda, J.R. (2000). Exploiting Latent Semantic Information in Statistical Language Modeling. *Proceedings of the IEEE*, 88(8), 1279-1296.

Benjamin, R.G. (2012). Reconstructing Readability: Recent Developments and Recommendations in the Analysis of Text Difficulty. *Educ Psycological Review*, 63-88.

Bohland, J.W. and Guenther, F.H. (2006). An fMRI Investigation of Syllable Sequence Production. *NeuroImage,* 32, 821-841.

Burgess, N. and Hitch, G.J. (2006). A Revised Model of Short-Term Memory and Long-Term Learning of Verbal Sequences. *Journal of Memory and Language*, 55, 627-652.

Bury, K.V. (1975). *Statistical Models in Applied Science*, John Wiley, New York, USA.

Carreiras, M., Álvarez, C.J. and Devega, M. (1993). Syllable Frequency and Visual Word Recognition in Spanish. *Journal of Memory and Language*, 32(6), 766-780.

Case, R., Midian Kurland, D. and Goldberg, J. (1982). Operational Efficiency and the Growth of Short-Term Memory Span. *Journal of Experimental Child Psychology*, 33, 386-404.

Chekaf, M., Cowan, N. and Mathy, F. (2016). Chunk Formation in Immediate Memory and How it Relates to Data Compression. *Cognition*, 155, 96-107.

Collins-Thompson, K. (2014). Computational Assessment of Text Readability: A Survey of Past, in Present and Future Research, Recent Advances in Automatic Readability Assessment and Text Simplification, *ITL, International Journal of Applied Linguistics*, 97-135.

Conway, A.R.A., Cowan, N., Michael F., Bunting, M.F., Therriaulta, D.J. and Minkoff, S.R.B. (2002). A Latent Variable Analysis of Working Memory Capacity, Short-Term Memory Capacity: Processing Speed, and General Fluid Intelligence. *Intelligence*, 163-183.

Cowan, N. (2000). The Magical Number 4 in Short-Term Memory: A Reconsideration of Mental Storage Capacity. *Behavioral and Brain Sciences*, 87-114.

Crowder, R.G. (1993). Short-Term Memory: Where Do We Stand?. *Memory & Cognition*, 21, 142-145.

D'Alfonso, S. (2011). On Quantifying Semantic Information. *Information*, 2, 61-101. doi:10.3390/info2010061.

Deniz, F., Nunez–Elizalde, A.O., Huth, A.G. and Gallant Jack, L. (2019). The Representation of Semantic Information Across Human Cerebral Cortex During Listening Versus Reading is Invariant to Stimulus Modality. *J. Neuroscience*, 39, 7722-7736.

Ding, N. and Simon, J.Z. (2014). Cortical Entrainment to Continuous Speech: Functional Roles and Interpretations. *Front. Hum. Neurosci*., 8, 311.

Doelling, K.B., Arnal, L.H., Ghitza, O. and Poeppel, D. (2014). Acoustic Landmarks Drive Delta-Theta Oscillations to Enable Speech Comprehension by Facilitating Perceptual Parsing. *Neuroimage,* 85, 761-768.

DuBay, W.H. (2004). *The Principles of Readability*, Impact Information, Costa Mesa, California.

DuBay (Editor) W.H. (2006). *The Classic Readability Studies*, Impact Information, Costa Mesa, California.

Fedorenko, E., Scott, T.L., Brunner, P. and Kanwisher, N. (2016). Neural Correlate of the Construction of Sentence Meaning. *Proc. Natl Acad. Sci.,* 113, E6256-E6262, USA.

Ferrand, L., Segui, J. and Grainger, J. (1996). Masked Priming of Word and Picture Naming: The Role of Syllabic Units. *Journal of Memory and Language*, 35(5), 708-723.

Ferrand, L., Segui, J. and Humphreys, G.W. (1997). The Syllable's Role in Word Naming. *Memory & Cognition*, 25(4), 458-470.

Flesch, R. (1948). A New Readability Yardstick. *Journal of Applied Psychology*, 222-233.

Flesch, R. (1974). *The Art of Readable Writings*. Harper and Row, New York. Revised and Enlarged 25th Anniversary Edition.

François, T. (2014). An Analysis of a French as Foreign Language Corpus for Readability Assessment. *Proceedings of the 3rd Workshop on NLP for CALL*, NEALT Proceedings Series 22, Linköping, Electronic Conference Proceedings, 107, 13-32.

Ghitza, O. (2011). Linking Speech Perception and Neurophysiology: Speech Decoding Guided by Cascaded Oscillators Locked to the Input Rhythm. *Front. Psychol*., 2, 130.

Gignac, G.E. (2015). The Magical Numbers 7 and 4 are Resistant to the Flynn Effect: No Evidence for Increases in Forward or Backward Recall across 85 Years of Data. *Intelligence*, 48, 85-95. https://doi.org/10.1016/j.intell.2014.11.001

Giraud, D. (2012). Poeppel, Cortical Oscillations and Speech Processing: Emerging Computational Principles and Operations. *Nat. Neurosci*., 15, 511-517.

Giroud, J., Lerousseau, J.P., Pellegrino, F. and Morillon, B. (2023). The Channel Capacity of Multilevel Linguistic Features Constrains Speech Comprehension. *Cognition,* 232, 105345.

Glanz, O., Hader, M., Schulze-Bonhage, A., Auer, P. and Ball, T. (2021). A Study of Word Complexity Under Conditions of Non-Experimental, Natural Overt Speech Production Using ECoG Front. *Hum. Neurosci*., 15. https://doi.org/10.3389/fnhum.2021.711886

Grondin, S. (2000). A Temporal Account of the Limited Processing Capacity. *Behavioral and Brain Sciences*, 24, 122-123.

Houdt, G.V., Mosquera, C. and Napoles, G., (2020). A Review on the Long Short-Term Memory Model. *Artificial Intelligence Review*, 53, 5929-5955.

Islam, M., Sarkar, A., Hossain, M., Ahmed, M. and Ferdous, A. (2023). Prediction of Attention and Short-Term Memory Loss by EEG Workload Estimation. *Journal of Biosciences and Medicines*, 304-318. doi: 10.4236/jbm.2023.114022.

Jones, G. and Macken, B. (2015). Questioning Short-Term Memory and its Measurements: Why Digit Span Measures Long-Term Associative Learning. *Cognition*, 1-13.

Jonides, J., Lewis, R.L., Nee, D.E., Lustig, C.A., Berman, M.G. and Moore, K.S. (2008). The Mind and Brain of Short-Term Memory. *Annual Review of Psychology*, 69, 193-224.

Kaminski, J. (2017). Intermediate-Term Memory as a Bridge between Working and Long-Term Memory. *The Journal of Neuroscience*, 37(20), 5045-5047.

Kandel, L. and Moles, A. (1958). *Application de l'indice de Flesch à la langue française*. *Cahiers Etudes de Radio–Télévision,* 253-274.

Kazanina, N., Bowers, J.S. and Idsardi, W. (2018). Phonemes: Lexical Access and Beyond. *Psychon. Bull. Rev*., 25, 560-585.

Keitel, A., Gross, J. and Kayser, C. (2018). Perceptually Relevant Speech Tracking in Auditory and Motor Cortex Reflects Distinct Linguistic Features. *PLOS Biol.,* 16, e2004473.

Khanna, A.R., Muñoz, W., Joon Kim, Y.J., Kfir, Y., Paulk, A.C., Jamali, M., Cai, J., Mustroph, M.L., Caprara, I., Hardstone, R., Mejdell, M., Meszéna, D., Zuckerman, A., Schweitzer, J., Cash, S. and Williams, Z.M. (2024). Single-Neuronal Elements of Speech Production in Humans. *Nature*, 626, 603-610. https://doi.org/10.1038/s41586–023–06982–w

Kincaid, J.P., Fishburne, R.P., Rogers, R.L. and Chissom, B.S. (1975). *Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel*, Research Branch Report, 8-75, Chief of Naval Technical Training, Naval Air Station, Memphis, TN, USA.

Lee, D.K., Fedorenko, E., Simon, M.V., Curry, W.T., Nahed, B.V., Cahill, D.P. and Williams, Z.M. (2018). Neural Encoding and Production of Functional Morphemes in the Posterior Temporal Lobe. *Nat. Commun*., 9, 1877.

Levelt, W.J.M., Roelofs, A. and Meyer, A.S. (1999). *A Theory of Lexical Access in Speech Production*, Cambridge Univ. Press.

Lindgren, B.W. (1968). *Statistical Theory*, 2nd Edition, MacMillan Company, New York, NY, USA.

Lisman, J.E. and Idiart, M.A.P. (1995). Storage of 7 ± 2 Short-Term Memories in Oscillatory Subcycles. *Science*, 267(5203), 1512-1515.

Lubinus, C., Keitel, A., Obleser, J., Poeppel, D. and Rimmele, J.M. (2023). Explaining Flexible Continuous Speech Comprehension from Individual Motor Rhythms. *Proc. Biol. Sci*., 290, 20222410.

Lucisano, P. and Piemontese, M.E. (1988). *GULPEASE: Una formula per la predizione della difficoltà dei testi in lingua italiana*. *Scuola e città*, 110-124.

Lundmark, M.S. and Erickson, D. (2022). Segmental and Syllabic Articulations: A Descriptive Approach. *Journal of Speech, Language, and Hearing Research*, 67, 3974-4001.

Luo, H. and Poeppel, D. (2007). Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. *Neuron,* 54, 1001-1010.

Luo, X., Chen, H.H. and Guo, Q. (2022). Semantic Communications: Overview, Open Issues, and Future Research Directions. *IEEE Wireless Communications*, 29(1), 210-219.

Mathy, F. and Feldman, J. (2012). What's Magic about Magic Numbers?. Chunking and Data Compression in Short-Term Memory. *Cognition*, 346-362.

Matricciani, E. (2019). Deep Language Statistics of Italian throughout Seven Centuries of Literature and Empirical Connections with Miller's 7-2 Law and Short-Term Memory. *Open Journal of Statistics*, 9, 373-406. https://doi.org/10.4236/ojs.2019.93026

Matricciani, E. (2020). Statistical Theory of Language Translation Based on Communication Theory. *Open J. Stat.*, 10, 936-997. https://doi.org/10.4236/ojs.2020.106055

Matricciani, E. (2022). Linguistic Mathematical Relationships Saved or Lost in Translating Texts: Extension of the Statistical Theory of Translation and its Application to the New Testament. *Information*, 13, 20. https://doi.org/10.3390/info13010020

Matricciani, E. (2023). Capacity of Linguistic Communication Channels in Literary Texts: Application to Charles Dickens' Novels. *Information*, 14, 68. https://doi.org/10.3390/info14020068

Matricciani, E. (2023). Linguistic Communication Channels Reveal Connections between Texts: The New Testament and Greek Literature. *Information*, 14, 405. https://doi.org/10.3390/info14070405

Matricciani, E. (2023). Readability across Time and Languages: The Case of Matthew's Gospel Translations. *AppliedMath*, 3, 497-509. https://doi.org/10.3390/appliedmath3020026

Matricciani, E. (2023). Readability Indices Do Not Say it All on a Text Readability. *Analytics*, 2, 296-314. https://doi.org/10.3390/analytics2020016

Matricciani, E. (2023). Short-Term Memory Capacity across Time and Language Estimated from Ancient and Modern Literary Texts. Study-Case: New Testament Translations. *Open Journal of Statistics*, 13, 379-403. doi: 10.4236/ojs.2023.133019.

Matricciani, E. (2024). A Mathematical Structure Underlying Sentences and Its Connection with Short-Term Memory. *AppliedMath*, 4, 120-142. https://doi.org/10.3390/appliedmath4010007

Matricciani, E. (2024). Is Short-Term Memory Made of Two Processing Units? Clues from Italian and English Literatures down Several Centuries. *Information*, 15, 6. https://doi.org/10.3390/info15010006

Matricciani, E. (2024). Multi-Dimensional Data Analysis of Deep Language in J.R.R. Tolkien and C.S. Lewis Reveals Tight Mathematical Connections. *AppliedMath*, 4, 927-949. https://doi.org/10.3390/appliedmath4030050

Matricciani, E. (2025). A Mathematical Analysis of Texts of Greek Classical Literature and their Connections. *Open Journal of Statistics*, 15, 1-34. doi: 10.4236/ojs.2025.151001.

Matricciani, E. (2025). Domestication of Source Text in Literary Translation Prevails over Foreignization. *Analytics*, 4, 17. https://doi.org/10.3390/analytics4030017

Matricciani, E. (2025). Translation Can Distort the Linguistic Parameters of Source Texts Written in Inflected Language: Multidimensional Mathematical Analysis of "The Betrothed", a Translation in English of "I Promessi Sposi" by A. Manzoni. *AppliedMath*, 5, 24. https://doi.org/10.3390/appliedmath5010024

Melton, A.W. (1963). Implications of Short-Term Memory for a General Theory of Memory. *Journal of Verbal Learning and Verbal Behavior*, 2, 1-21.

Miller, G.A. (1955). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review*, 343-352.

Murdock, B.B. (1972). Short-Term Memory. *Psychology of Learning and Motivation*, 5, 67-127.

Nelson, M.J., Karoui, I.E., Giber, K., Yang, X., Cohen, L., Koopman, H., Cash, S.S., Naccache, L., Hale, J.T., Pallier, C. and Dehaene, S. (2017). Neurophysiological Dynamics of Phrase-Structure Building During Sentence Processing. *Proc. Natl Acad. Sci.,* 114, E3669-E3678, USA.

Norris, D. (2017). Short-Term Memory and Long-Term Memory are Still Different. *Psychological Bulletin*, 143(9), 992-1009.

Oganian, Y. and Chang, F.A. (2019). Speech Envelope Landmark for Syllable Encoding in Human Superior Temporal Gyrus. *Sci. Adv.,* 5.

Papoulis Papoulis, A. (1990). *Probability & Statistics*, Prentice Hall, Hoboken, NJ, USA.

Park, H.R., Ince, R.A.A., Schyns, P.G., Thut, G. and Gross, J. (2023). Frontal Top-Down Signals Increase Coupling of Auditory Low-Frequency Oscillations to Continuous Speech in Human Listeners. *Curr. Biol*., 25, 1649-1653.

Parkes, Malcolm B. (2016). *Pause and Effect: An Introduction to the History of Punctuation in the West*, Abingdon, Routledge.

Peelle, J.E., Gross, J. and Davis, M.H. (2013). Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. *Cereb. Cortex,* 23, 1378-1387.

Poeppel, D. (2003). The Analysis of Speech in Different Temporal Integration Windows: Cerebral Lateralization as 'Asymmetric Sampling in Time'. *Speech Commun*., 41, 245-255.

Poeppel, D. and Assaneo, M.F. (2020). Speech Rhythms and their Neural Foundations. *Nat. Rev. Neurosci*., 21, 322-334.

Pothos, E.M. and Joula, P., (2000). Linguistic Structure and Short-Term Memory. *Behavioral and Brain Sciences*, 138-139.

Potter, M.C. (2012). Conceptual Short-Term Memory in Perception and Thought. *Frontiers in Psychology*. doi: 10.3389/fpsyg.2012.00113.

Richardson, J.T.E. (2007). Measures of Short-Term Memory: A Historical Review. *Cortex*, 43(5), 635-650.

Rosen, S. (1992). Temporal Information in Speech: Acoustic, Auditory and Linguistic Aspects. *Philos. Trans. R. Soc. Lond. B Biol. Sci*., 336, 367-373.

Rosenzweig, M.R., Bennett, E.L., Colombo, P.J. and Lee, P.D.W. (1993). Short-Term, Intermediate-term and Long-Term Memories. *Behavioral Brain Research*, 57(2), 193-198.

Saaty, T.L. and Ozdemir, M.S. (2003). Why the Magic Number Seven Plus or Minus Two. *Mathematical and Computer Modelling*, 233-244.

Saberi, K. and Perrott, D.R. (1999). Cognitive Restoration of Reversed Speech. *Nature*, 398, 760.

Shi, G., Xiao, Y., and Li, Xie, (2021). From Semantic Communication to Semantic-Aware Networking: Model, Architecture, and Open Problems. *IEEE Communications Magazine*, 59(8), 44-50.

Strinati, E.C. and Barbarossa, S. (2021). 6G Networks: Beyond Shannon Towards Semantic and Goal-Oriented Communications. *Computer Networks*, 190(8), 1-17.

Trauzettel-Klosinski, S. and Dietz, K. (2012). Standardized Assessment of Reading Performance: The New International Reading Speed Texts IreST. *IOVS*, 5452-5461.

Walenski, M., Europa, E., Caplan, D. and Thompson, C.K. (2019). Neural Networks for Sentence Comprehension and Production: An ALE-Based Meta-Analysis of Neuroimaging Studies. *Hum. Brain Mapp.,* 40, 2275-2304.

Wanting, Y., Hongyang, D., Liew, Z.Q., Lim, W.Y.B., Xiong, Z., Niyato, D., Chi, X., Shen, X. and Miao, C. (2023). Semantic Communications for Future Internet: Fundamentals, Applications, and Challenges. *IEEE Communications Surveys & Tutorials*, 25(1), 213-250.

Xie, H., Qin, Z., Li, G.Y. and Juang, B.H. (2021). Deep Learning Enabled Semantic Communication Systems. *IEEE Trans. Signal Processing*, 69, 2663-2675.

Zamanian, M. and Heydari, P. (2012). Readability of Texts: State of the Art. *Theory and Practice in Language Studies*, 43-53.

Zhong, Y. (2017). A Theory of Semantic Information. *China Communications*, 1-17.

Appendix

Appendix A: List of Mathematical Symbols and Definition

Symbol	Definition
C_P	Characters per word
I_p	Word interval
m	Linear mean value
M	Mode
$M_{0.5}$	Median
M_F	Word intervals per sentence
P_F	Words per sentence
s	Linear standard deviation
n_C	Number of characters per chapter
n_W	Number of words per chapter
n_S	Number of sentences per chapter
n_I	Number of interpunctions per chapter
μ	Natural log mean value
σ	Natural log standard deviation

Appendix B: Scatterplots Involving Syllables

Figures A1-A5 shows the scatterplot between words and syllables and between characters and syllables, for the indicated languages. Figures A5, A6 show the scatterplots between S_p and words, and between C_s and characters for different languages.

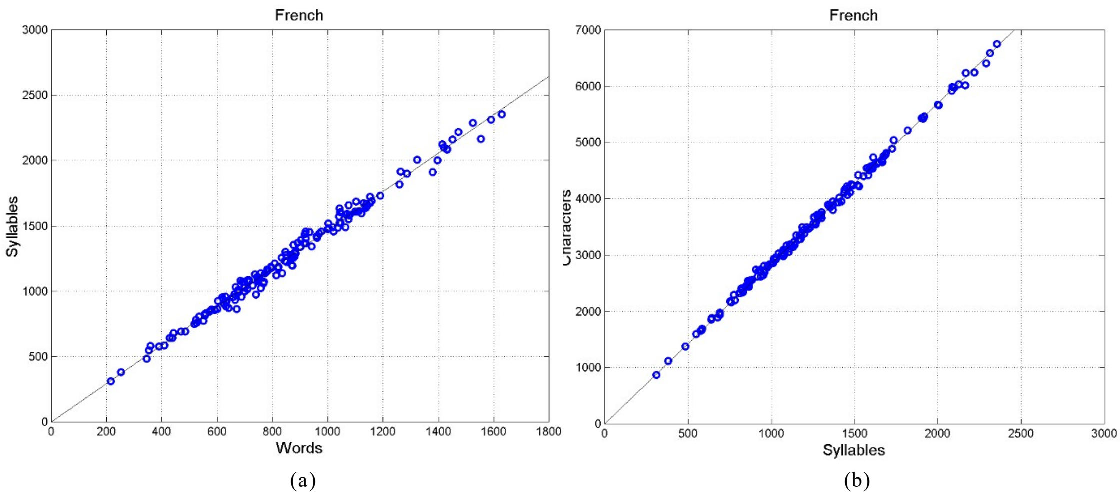
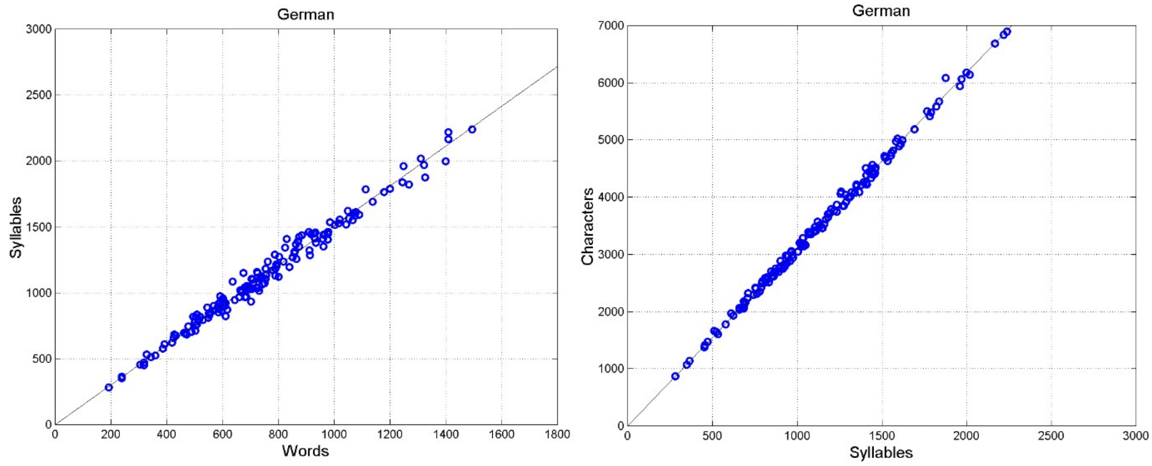


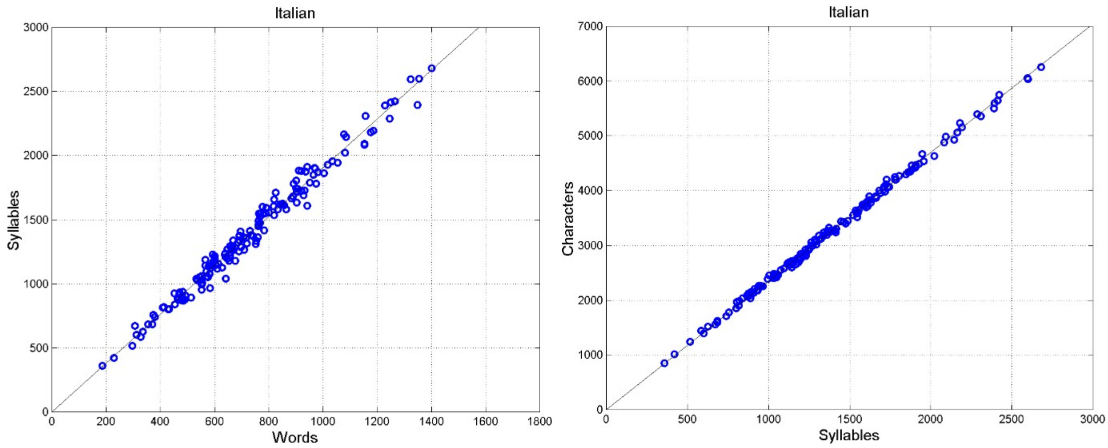
Figure A1: French: (a) Scatterplot between Words and Syllables; (b) Scatterplot between Characters and Syllables (155 Samples)

Appendix (Cont.)



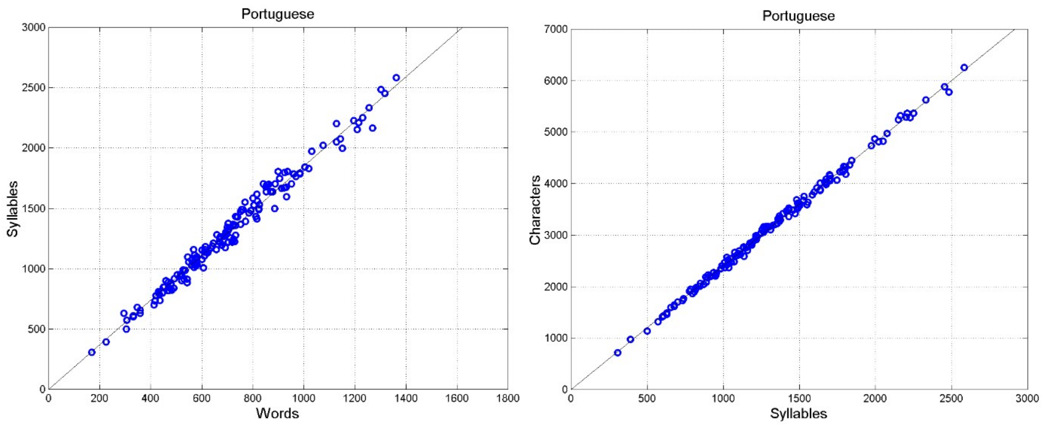
(a) (b)

Figure A2: German (a) Scatterplot between Words and Syllables; (b) Scatterplot between Characters and Syllables (155 Samples)



(a) (b)

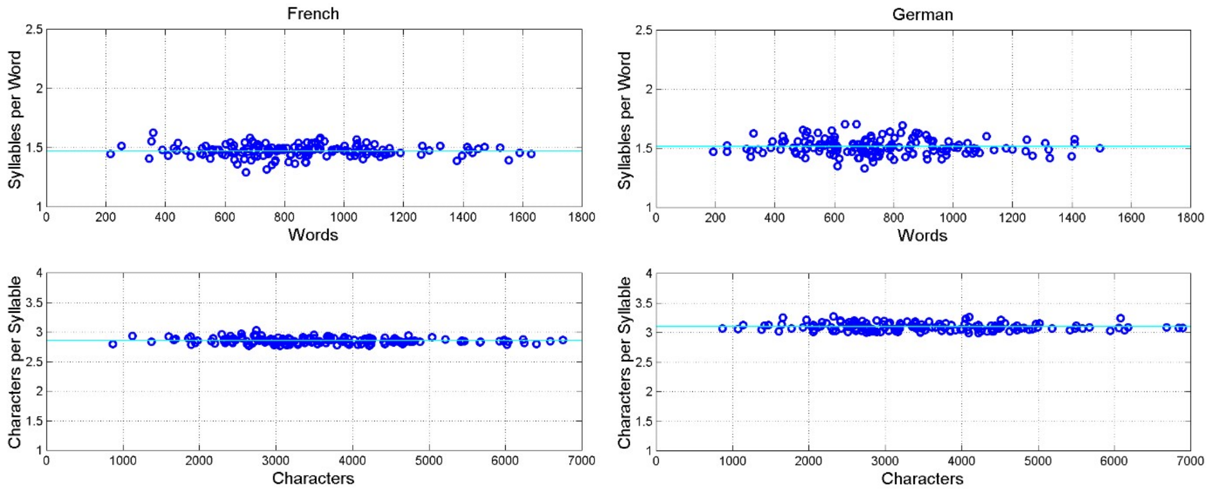
Figure A3: Italian: (a) Scatterplot between Words and Syllables; (b) Scatterplot between Characters and Syllables (155 Samples)



(a) (b)

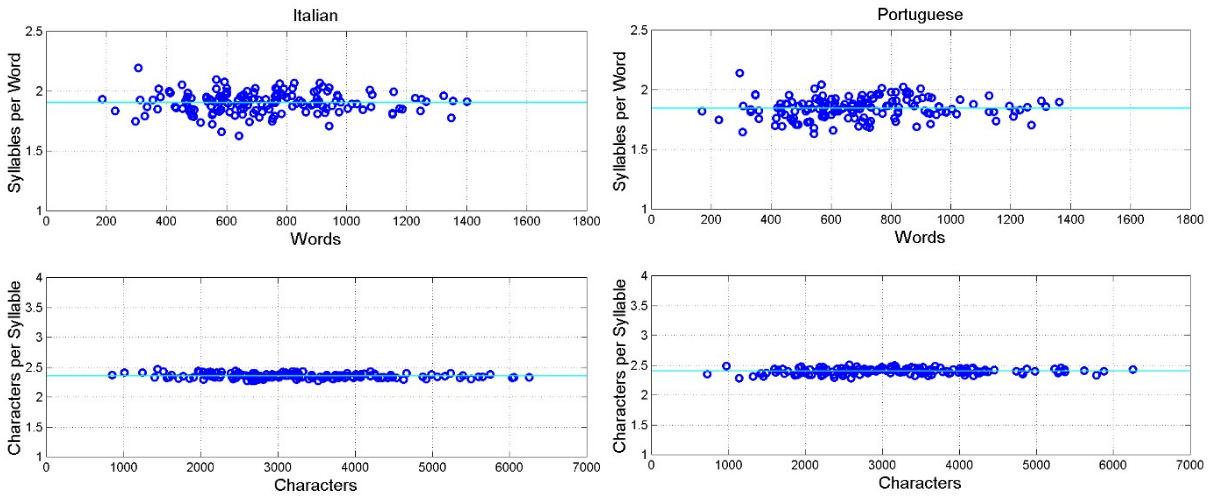
Figure A4: Portuguese: (a) Scatterplot between Words and Syllables; (b) Scatterplot between Characters and Syllables (155 Samples)

Appendix (Cont.)



(a) (b)

Figure A5: (a) French: Upper Panel: Scatterplot between S_P and Words; Lower Panel: Scatterplot between C_S and Characters (155 Samples); (b) German: Upper Panel: Scatterplot between S_P and Words; Lower Panel: Scatterplot between C_S and Characters (155 Samples); The Mean Value is Drawn with the Cyan Line



(a) (b)

Figure A6: (a) Italian: Upper Panel: Scatterplot between S_P and Words; Lower Panel: Scatterplot between C_S and Characters (155 Samples); (b) Portuguese: Upper Panel: Scatterplot between S_P and Words; Lower Panel: Scatterplot between C_S and Characters (155 Samples); The Mean Value is Drawn with the Cyan Line

Appendix C: Unconditional Mean and Variance

Let m_k and S_k be the mean value and standard deviation of samples belonging to set $k-th$, out of N sets of the ensemble. From statistical theory, the unconditional mean (ensemble mean) m is given by the mean of means:

$$m = \frac{1}{N}\sum_{k=1}^{N} m_k \quad \text{...(A1)}$$

The unconditional variance (ensemble variance) S^2 (S is the unconditional standard deviation) is given:

$$s^2 = \mathrm{var}(m_k) + \frac{1}{N}\sum_{k=1}^{N} s_k^2 \quad \text{...(A2)}$$

With

$$\mathrm{var}(m_k) = \frac{1}{N}\sum_{k=1}^{N} m_k^2 - m^2 \quad \text{...(A3)}$$

Appendix (Cont.)

Appendix D: Three-Parameter Log-Normal Probability Density Modelling

Let us consider a stochastic variable $x \geq 1$ (in linear units) with mean value $m_x = <x>$ and standard deviation S_x (as those reported in Table 1).

Let us consider the variable linear transformation $y = x - 1$, with mean value $m_y = m_x - 1$ and standard deviation $S_y = S_x$, then the log-normal probability density function of y is given by Papoulis Papoulis (1990), Lindgren (1968) and Bury (1975):

$$f(y) = \frac{1}{\sqrt{2\pi} s_y y} \exp\left\{-\frac{1}{2}\left[\frac{log(y)-\mu}{\sigma}\right]^2\right\} \quad ...(A4)$$

where the mean μ (Np) and the standard deviation σ (Np) are given by:

$$\sigma^2 = \ln\left[\left(\frac{s_y}{m_y}\right)^2 + 1\right] \quad ...(A5)$$

$$\mu = \ln\left[m_y - \frac{\sigma^2}{2}\right] \quad ...(A6)$$

Returning to the variable x, its mode M (most probable value) is given by:

$$M = \exp(\mu - \sigma^2) + 1 \quad ...(A7)$$

The median $M_{0.5}$(value exceeded with 0.5 probability) is given by:

$$M_{0.5} = \exp(\mu) + 1 \quad ...(A8)$$

Now, by assuming histogram bins centered at x_k, $k = 1, 2, ...$, with bin width Δx, the number of samples $n(x_k)$ per bin, out of Z total samples, is given by:

$$n(x_k) \approx Z \times f(x_k) \times \Delta x \quad ...(A9)$$

In our case, $Z = 2 \times 155 = 310$, therefore we get the estimated histogram reported in Figure 4.

Appendix E: Correlation Coefficients between Deep-Language Parameters

Table A1 reports the correlation coefficient between the logarithms of the indicated deep-language parameters, for each language. From this table we can conclude that C_P, I_P and M_F are practically uncorrelated. Because of the bivariate Gaussian modelling, they can be assumed independent.

Table A1: Correlation Coefficient between the Logarithms of the Indicated Deep-Language Parameters, for Each Language

Language	C_P versus I_P	C_P versus M_F	M_F versus I_P	S_P versus I_P	S_P versus M_F
Greek	0.3176	–0.0988	0.0686	—	—
Latin	0.3568	0.0944	0.0157	—	—
Esperanto	0.0747	0.0680	0.2735	—	—
French*	0.1439	–0.0643	–0.3932	0.1403	–0.0689
Italian*	–0.0117	0.2950	–0.2399	–0.0785	0.2999
Portuguese*	–0.2227	–0.0814	–0.0798	–0.1029	–0.0094
Romanian°	–0.0366	0.4040	–0.1230	—	—
Spanish	0.1052	0.1074	–0.0320	—	—
Danish°	0.4588	–0.2698	–0.2699	—	—
English*	0.1568	0.3103	–0.2650	0.0858	0.3320
Finnish°	0.2906	0.2024	–0.1003	—	—
German*	0.2895	0.1232	–0.3047	0.3531	0.0285
Icelandic	0.1487	–0.0834	–0.2382	—	—
Norwegian°	0.1708	–0.1665	–0.6105	—	—
Swedish°	0.2352	–0.0189	–0.6260	—	—
Bulgarian	0.2739	–0.1167	–0.2792	—	—

Appendix (Cont.)

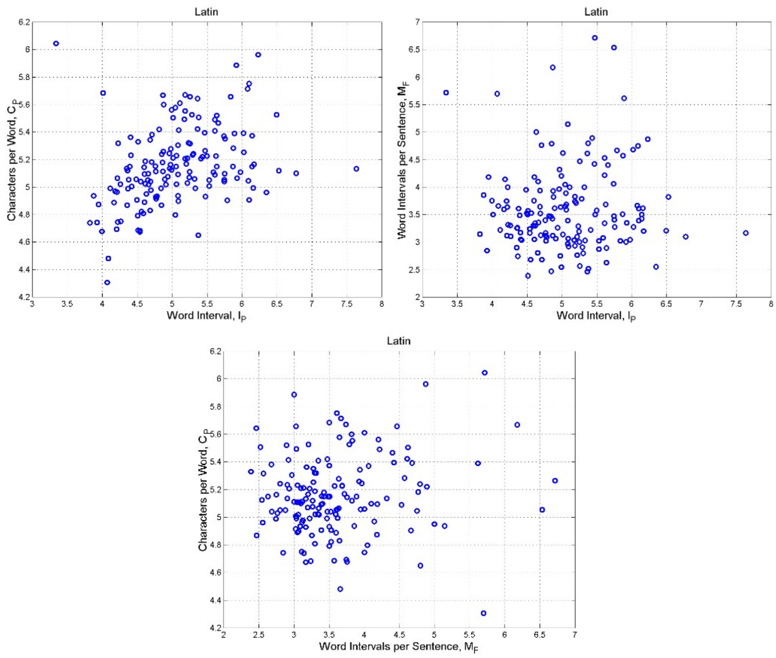
Table A1 (Cont).

Czech°	0.1935	0.0708	–0.1467	—	—
Croatian°	0.1446	0.1111	0.0007	—	—
Polish°	0.2423	–0.0994	–0.2518	—	—
Russian	0.0686	0.1369	0.0508	—	—
Serbian	0.0424	0.2215	–0.2461	—	—
Slovak	0.2195	–0.0266	–0.2101	—	—
Ukrainian	0.3878	–0.3976	–0.5274	—	—
Estonian°	0.3783	0.1358	–0.1090	—	—
Hungarian°	0.2464	–0.2439	–0.0263	—	—
Albanian	–0.0244	0.1433	0.0388	—	—
Armenian	0.2587	0.2234	0.1047	—	—
Welsh	–0.0800	0.0988	–0.0313	—	—
Basque	0.2283	–0.0678	–0.0450	—	—
Hebrew	0.2078	0.2880	–0.1593	—	—
Cebuano	0.1130	–0.2173	–0.5899	—	—
Tagalog	0.2838	–0.3915	–0.3069	—	—
Chichewa	–0.0003	–0.0935	–0.4157	—	—
Luganda	–0.0569	0.1538	–0.1771	—	—
Somali	–0.0243	–0.1853	0.1491	—	—
Haitian	0.2753	0.0173	–0.3800	—	—
Nahuatl	–0.2043	–0.0967	–0.5228	—	—

Appendix F: Scatterplots between S_P (or C_P) I_P and M_F

Figures A7-A11 show the scatterplots between S_P (or C_P) I_P and M_F for the indicated languages.

(a) (c)



(b)

Figure A7: Latin: (a) Scatterplot between C_P and I_P; (b) Scatterplot between M_F and I_P. (c) Scatterplot between C_P and M_F

Appendix (Cont.)

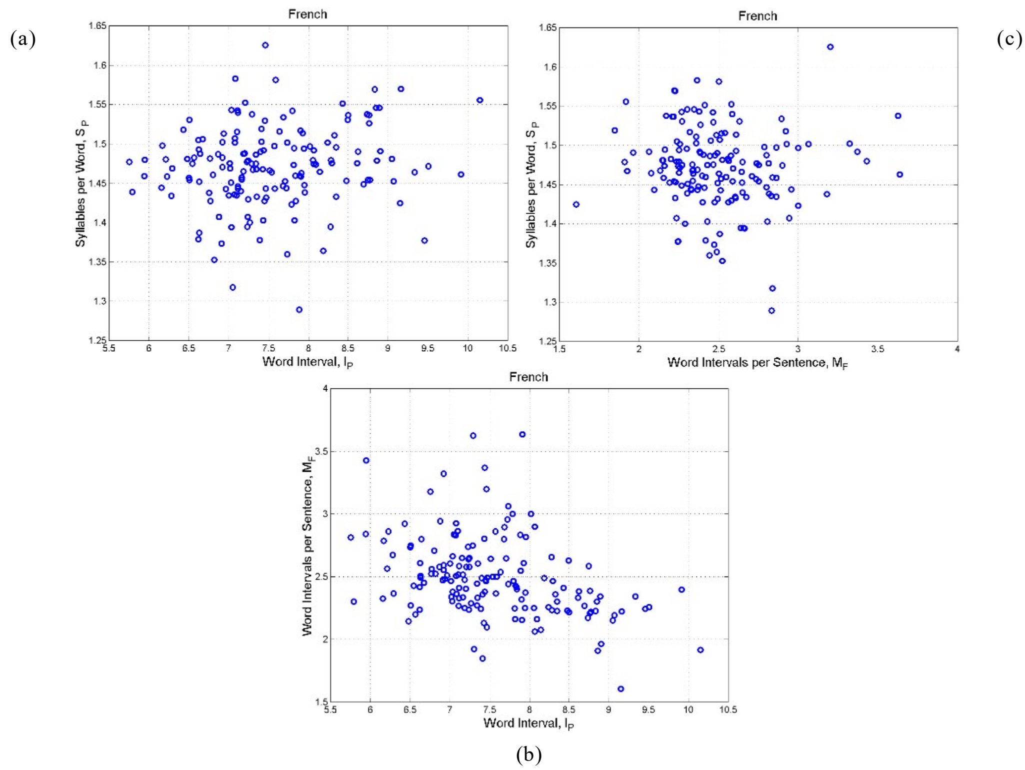


Figure A8: French: (a) Scatterplot between S_P and I_P; (b) Scatterplot between M_F and I_P; (c) Scatterplot between S_P and M_F

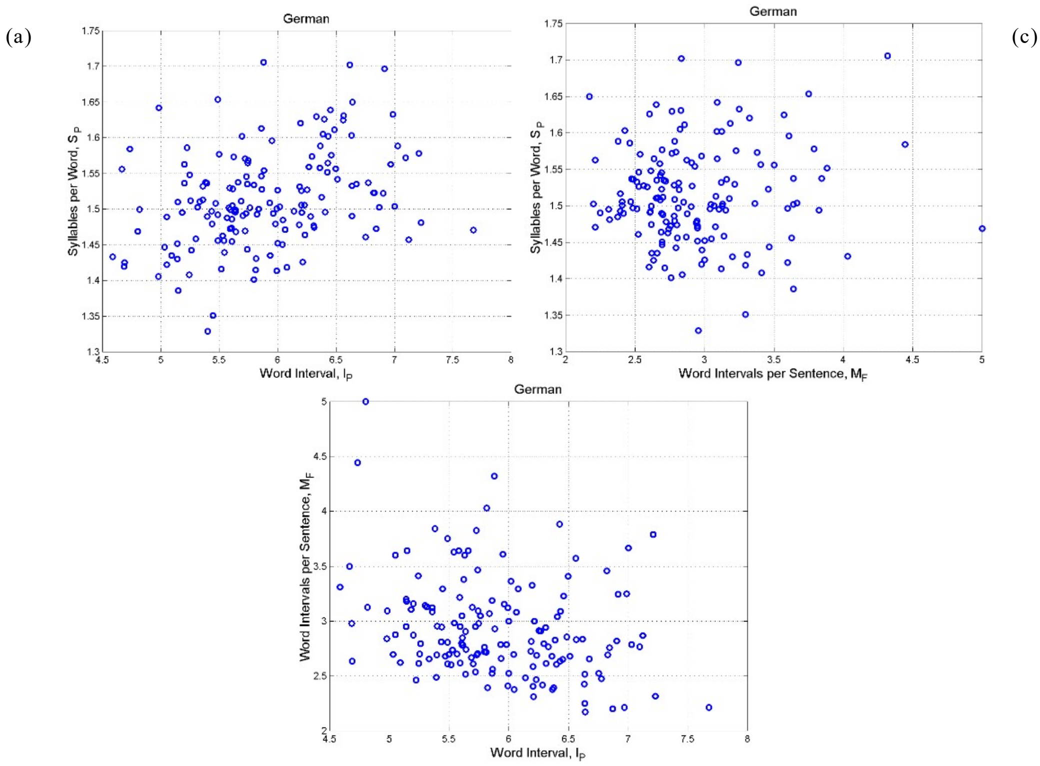
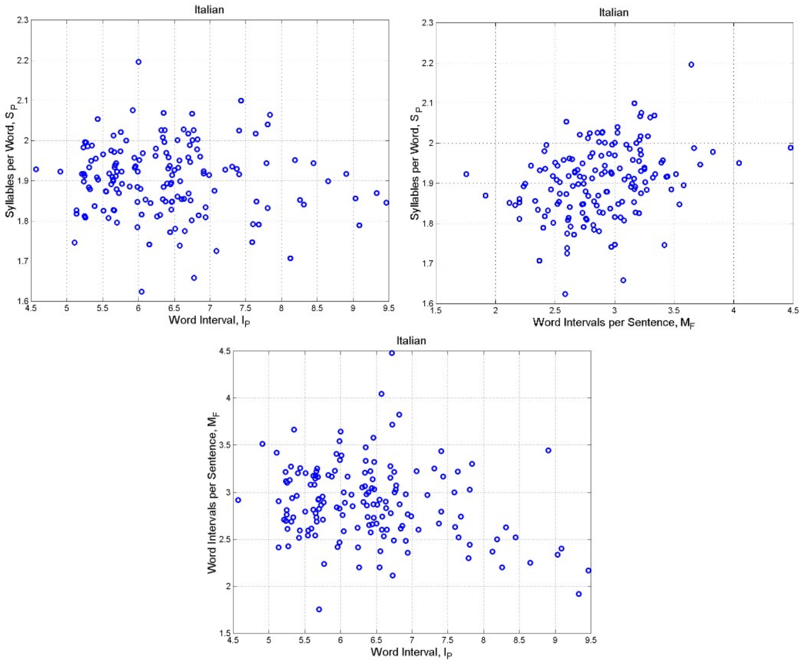


Figure A9: German: (a) Scatterplot between S_P and I_P; (b) Scatterplot between M_F and I_P; (c) Scatterplot between S_P and M_F

Appendix (Cont.)

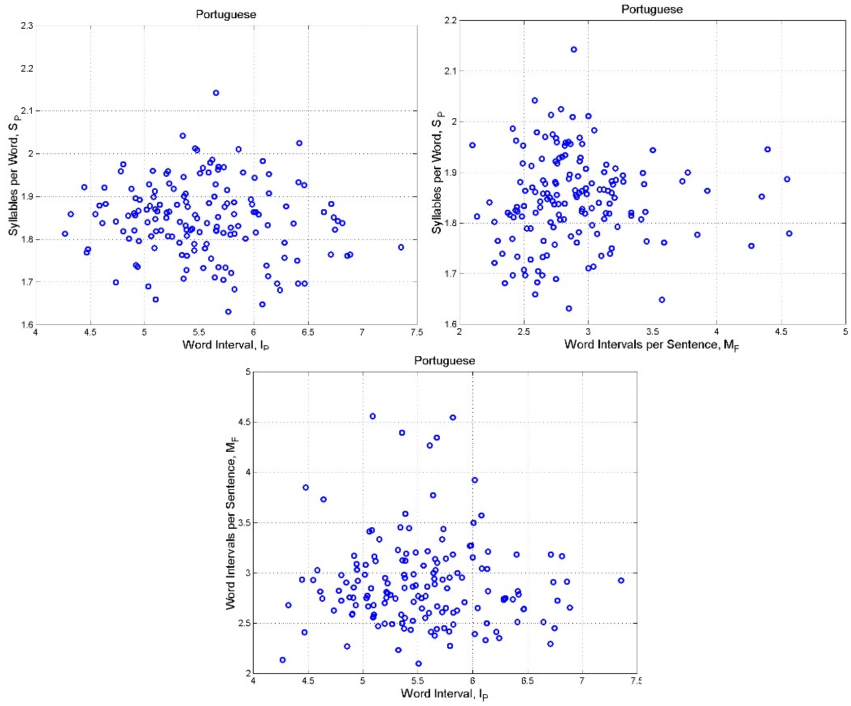
(a) (c)



(b)

Figure A10: Italian: (a) Scatterplot between S_P and I_P; (b) Scatterplot between M_F and I_P; (c) Scatterplot between S_P and M_F

(a) (c)



(b)

Figure A11: Portuguese: (a) Scatterplot between S_P and I_P; (b) Scatterplot between M_F and I_P; (c) Scatterplot between S_P and M_F

Cite this article as: Emilio Matricciani (2025). Equivalent Processors Modelling the Short-Term Memory. *International Journal of Pure and Applied Mathematics Research*, 5(2), 1-28. doi: 10.51483/IJPAMR.5.2.2025.1-28.