

OPEN ACCESS

\*Corresponding author

Askandar Khalid Abdullah  
[askander.abduallah@su.edu.krd](mailto:askander.abduallah@su.edu.krd)

# Syllabification in English through AI : A Comparative Study of ChatGPT and Gemini

**Askandar Khalid. Abdullah/** Road Construction Department, College of Technology, University of Polytechnic - Erbil, Kurdistan Region, Iraq  
**Pakhshan Ismail Hamad/** English Department, College of Education, University of Salahaddin, Erbil, Kurdistan Region, Iraq

RECEIVED :21 /05/2025  
ACCEPTED :03/07/ 2025  
PUBLISHED :15/10/ 2025

**Keywords:**

Syllabification,  
AI,  
ChatGPT,  
Gemini



**Abstract**

This paper examines the English language syllabification task, focusing on two advanced Artificial Intelligence (AI) models: ChatGPT and Gemini. Syllabification is one of the first things one learns about phonology; it is the process of breaking words apart into their smallest phonological units, syllables. It is an essential idea that we will consider for a number of purposes, including language learning, speech synthesis, or computational linguistics, that the systematic view of syllables might help you to teach engineering development in voice recognition and the reaction process. The analysis of transcribed words is based on Peter Roach (RP) in (2009) and the Oxford Advanced Learner's Dictionary (2020). It utilizes a carefully selected dataset comprising 100 multisyllabic English words, covering all lexical categories. Then the words have been carefully analyzed to evaluate the accuracy of segmentation in the syllabification process. Through a comprehensive comparative analysis, the study identified the commonalities and differences in the syllabification patterns exhibited by the two AI models. This study significantly contributes to phonemic research, demonstrating the potential applications and limitations of AI in linguistic contexts, and shows that ChatGPT is more accurate in the task of syllabification than Gemini.



**About the Journal**

Zanco Journal of Humanity Sciences (ZJHS) is an international, multi-disciplinary, peer-reviewed, double-blind and open-access journal that enhances research in all fields of basic and applied sciences through the publication of high-quality articles that describe significant and novel works; and advance knowledge in a diversity of scientific fields. <https://zancojournal.su.edu.krd/index.php/JAHS/about>

## **1. Introduction**

### **1.1 Background of the Study**

Syllabification plays a fundamental role in the fields of phonology and speech processing. It governs how words are articulated and understood, influencing spelling, pronunciation, and rhythm in spoken English. The accurate syllabification is essential for English language learners. However, syllable boundaries in English can be unreliable due to complex phonotactic rules (Roach, 2009).

Syllables are the basic rhythmic units in speech, generally consisting of a nucleus (typically a vowel) and optional onsets and codas (Roach, 2009). Another definition for the concept of syllabification is by Ladefoged & Johnson (2014): "Syllabification is the process by which a word is broken down into syllables, each syllable typically containing a vowel or syllabic consonant as its nucleus, surrounded by consonants forming the onset and/or coda"

Accurate syllabification facilitates speech synthesis and recognition and aids in language acquisition. The placement of syllable boundaries is not always straightforward and can vary depending on pronunciation patterns.

English spelling does not always correspond neatly to its phonological structure, leading to challenges in syllable division. For example, silent letters or digraphs complicate the process of syllabification. Moreover, different dialects of English may syllabify the same word differently, depending on pronunciation variations (Ladefoged & Johnson, 2014).

### **1.2 Rise of Artificial Intelligence in Linguistic Tasks**

The arrival of AI, specifically the development of large language models (LLMs), has revolutionized numerous sectors and aspects of our daily lives. These sophisticated algorithms, capable of processing and generating human-like text, have enhanced communication, automated routine tasks, and significantly easier access to information (Davis, 2024).

ChatGPT by OpenAI and Gemini by Google DeepMind are among the leading LLMs with the capacity to perform linguistic tasks, based on their vast training quantities and sophisticated construction (OpenAI, 2023; Google DeepMind, 2024). These models use information and deep contextual analysis to replicate or even enhance traditional rule-based processing. Due to the increasing use of AI, current research aims to demonstrate its capabilities while ensuring alignment with societal values and promoting phonological tasks (Fakih, 2020).

### **1.3 Problem of the Study**

Despite advancements in AI, the consistency of syllabification across models remains underexplored. One of the problems is double-checking words; the process of syllabification sometimes varies between AI models, especially Gemini, which does not always indicate syllable boundaries clearly. Another significant challenge is that AI models can be significantly influenced by the prompts they are given. The input they receive can dramatically influence their responses, leading to variations in the quality and relevance of the output.

### **1.4 Objectives of the Study**

This study is a comparison of how well both ChatGPT and Gemini accurately identify syllables. It studies the most frequent syllabification errors performed by each model. The study goes on to analyse the consistency with which each of the models follows standard phonological rules. It aims to reveal the particularities of model behavior by analyzing a broad spectrum of multisyllabic words. In other words, it evaluates the answer of AI tools in yielding a linguistically correct syllabification.

### **1.5 Significance of the Study**

This study provides valuable insights into the effectiveness and capabilities of AI models when addressing linguistically complex tasks, such as the intricate process of syllabification. And it has some academic value, both as a contributor to computational phonology and as a modelling hinge between old-style linguistic theory and modern AI. The pedagogical value of this study represents one of this study most important significances as it provides insights that would be beneficial to both teachers and students.

## 2. Literature Review

While extensive research exists on AI and separately on English syllabification, no studies have combined these two areas. Because most prior work either focuses on the capabilities of AI in general language tasks or examines syllabification from a linguistic perspective, without exploring their intersection.

The study by Mohammadkarimi (2024) investigated the impact of AI tools (Listnr and Murf) on English pronunciation skills among 78 learners and 19 teachers in Iran using a two-month mixed-method approach. Divided into control and experimental groups, the study found that learners using AI showed significant improvement in pronunciation accuracy, increased confidence, and greater engagement, despite some challenges in interpreting feedback and recognizing nuances.

Another study by Vančová (2023) highlights the current use of AI in language pedagogy, with a focus on pronunciation training via mobile apps, virtual assistants, and chatbots. It examined learners ranging from high school students to professionals in tourism, using tools like Duolingo, Busuu, and Google Assistant to improve pronunciation. The study concludes that while AI shows promise, further research and practical implementation are needed for broader adoption in pronunciation teaching.

An additional study conducted by Duanmu et al. (2005) investigates phonological variability in English with a focus on syllable composition and stress assignment. Adopting a theoretical approach, the study evaluates four models of phonological system diversity: no-pattern, norm-and-exceptions, loose requirements, and inviolable constraints. Drawing on comprehensive data from the English lexicon, the authors argue in favor of the inviolable constraints model, presenting evidence that stressed syllables must be heavy and that syllable structures adhere to specific, predictable patterns. Their findings offer significant insights into the organization of English phonotactics and prosodic structure, enhancing the theoretical understanding of sound patterning in the language.

Although this represents only a small selection of the available studies, all have been thoroughly examined in prior research, either within the domains of English language studies or about AI and pronunciation. However, none have specifically addressed the intersection of syllabification and AI, or any existing research directly compared the syllabification performance of different AI models, such as ChatGPT and Gemini. That's why this study addresses that gap by concentrating specifically on syllable segmentation, offering a direct comparison between the two AI models, ChatGPT and Gemini.

## 3. Syllabification in English

### 3.1 Syllable Division

Dividing English words into syllables is complex, even for natives, as it follows specific linguistic principles. Syllable division respects the word's phonemic structure; each syllable contains a vowel and starts with a peak sound. Syllables can be classified by the consonants adjacent to the vowel: open syllables end with a vowel sound (e.g., go), while closed syllables end with a consonant (e.g., cat). When dividing multi-syllabic words, consonants between vowels typically start the next syllable, as in la-dy or hap-py. However, many English words have exceptions and complex structures that complicate this process.

Syllabification can be phonemically divided into three parts: onset, nucleus, and coda, as shown in Figure 1. The onset is a sequence of one or more consonants with which the syllable begins; the nucleus is the vowel or vowel-like sound that makes up the basic part of the syllable; and the coda is any consonant or consonant cluster that follows the nucleus. For example, in the word "text," the components would be written as onset /t/, nucleus /e/, and coda /kst/.

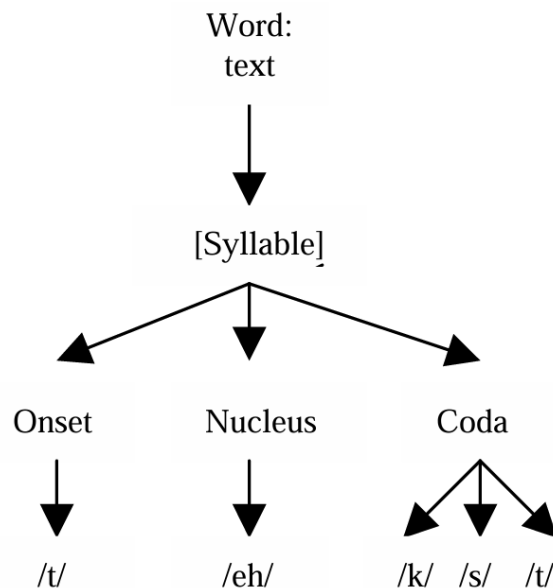


Figure 1: Diagram of the syllable structure taken from (Tian, 2004)

The number of syllables any word has can vary. Some words are monosyllabic, while others may consist of two, three, or more syllables. The division of syllables is arbitrary for certain words, particularly those containing complicated consonant clusters. In some words, there are silent letters that do not belong to any syllable. For example, "knight" is monosyllabic with four phonemes.

In the English language, there are certain principles for dividing syllables, which are as follows:

### 1. Maximum Onsets Principle (MOP)

Syllabification is based on the principle of maximizing onsets, meaning that a consonant should be placed at the beginning of the following syllable whenever possible. The significance of open syllables that end in a vowel without a coda is highlighted, as they are a universal feature across all human languages. Open syllables are considered "primitive" and are more prevalent than closed syllables, reinforcing their theoretical importance in linguistic structures.

The principle of maximum onsets means a consonant in a V.CV syllabification scheme is assigned to the second syllable, thus removing onsets in the first syllable. Onset is obligatory in a V.CV sequence; a consonant must be placed between the two vowels to ensure that the syllable boundary is placed before the consonant so that the second syllable has an onset and the first is left open. A similar bias toward onsets prevails in many languages over syllables in which the onset is optional. "A View of Onset" is discussed in Roach (2009)", ... where two syllables are to be divided, any consonant between them should be attached to the right-hand syllable, not the left, as far as possible." This rule means that open syllables are invariably prioritized, with the subsequent consonant forming the onset of its second syllable.

For instance, in the word "prefer," it must be divided as /pri.fɜ:/, not /prɪf.ɜ:/, as the second letter, 'f', is the onset for the second syllable. The MOP indicates the trend towards the syllables that have the onset, which is a consonant. Furthermore, how the functions of the different types of consonant clusters can be explained is answered by the Sonority Sequencing Principle, for the reason that it states that syllables ought to group from weak to stronger in terms of sonority (Giegerich, 1998).

### 2. Sonority Sequencing Principle (SSP)

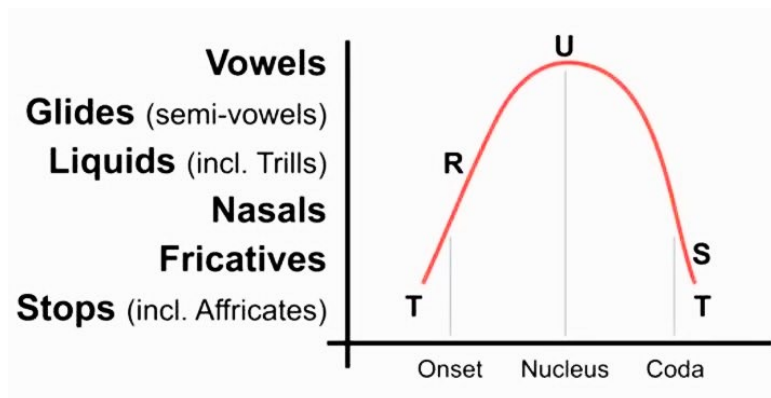
Parker, (2002) has proposed that the 'sonority scale' could be worked out in an alternative way, without having recourse to the study of phonotactic regularities, by what is generally called the sonority cycle.

Segments on the scale are labelled with a sonority feature, creating a sonority profile that rises towards the syllable peak and decreases towards the end. Segments are classified using binary feature categories: [syllabic], [voiced], [approximant], and [sonorant]. High vowels possess the

highest sonority, followed by glides, liquids, nasals, and obstruents. Additionally, the onset of a syllable has precedence over the coda of the previous syllable.

According to a sonority scale, different classes of sounds are ranked by their degree of sonority (the possible position of any of these classes of sounds in the syllable is mainly determined by the degree of sonority the sound possesses) (Parker, 2002).

There is a phonological rule for the SSP that sounds rise in sonority from the onset to the nucleus and fall from the nucleus to the coda of a syllable. The following is a sonority graph that demonstrates how SSP works:



**Figure. (2): A sonority graph for the word “trust” taken from the Wattpad Website.**

It's a sonority graph for 'trust' [trʌst]. It should rise from the onset and then fall from the nucleus to the coda, which is indeed the case in [trʌst]:

[tr]→onset, Sonority rises from [t] to [r]

[ʌ]→nucleus, Sonority rises from an [ʌ] to [ʌ]

[st]→coda, Sonority falls from [s] to [t]

The sonority hierarchy ranks speech sounds from highest to lowest amplitude. Vowels are the most sonorous, while voiceless plosives like /p/, /t/, and /k/ are the least. A simplified hierarchy orders sounds as follows: Vowels > Glides > Liquids > Nasals > Fricatives > Affricates > Plosives. Voiced sounds are more sonorous than their voiceless counterparts, such as /z/ being more sonorous than /s/.

### 3.2 Basic Rules of English Syllabification

These principles help establish consistency in English syllabification, though there are exceptions due to English's diverse etymological roots.

The division of words into smaller sound units or beats governs syllable structure in English. The following are some main rules that guide syllable formation in the English language (McMahon, 2002):

Rule 1: Syllables contain at least one vowel, usually formed by one or more vowels.

**Monophthongs** (single vowel sounds) (ɪ, e, ε, æ, ʌ, ɒ, ɔ, ʌ, ʊ, u, ə, ɜ:):

Examples: "a" /ə/, "ba-na-na" /bə'nɑ:nə/, "chil-dren" /'tʃɪl.drən/.

**Diphthongs** (two vowels as one sound) form a single syllable (ɪə, eə, ʊə, aɪ, aʊ, ɔɪ).

Examples: "cloud-y" /'klaʊ.di/, "beau-ti-ful" /'bjʊ:.tɪ.fəl/.

Rule 2: Dividing between consonants- Syllable breaks often occur with repeated consonant letters.

Examples: "let-ter" /'letə/, "rab-bit" /'ræbɪt/.

Rule 3: Single consonant between two vowel sounds"

a) A long first vowel leads to joining the next vowel.

Examples: "ro-bot" /'rəʊ.bɒt/, "wa-ter" /'wɔ:tə/.

b) Short first vowel associates with the first vowel.

Examples: "pan-el" /'pæn.əl/, "riv-er" /'rɪv.ər/.

Rule 4: Two vowels with separate sounds break between them.

Examples: "di-et" /'daɪ.ət/, "di-aer-e-sis" /,daɪ'ɛr.i.sɪs/.

Rule 5: Prefixes and suffixes to guide syllable division.

Examples: "end-less" /'ɛnd.ləs/, "pre-paid" /pri:'peɪd/.

Rule 6: Divide compound words at the boundary between the two words.

Examples: "sun-flow-er" /'sʌn.fləʊ.ər/, "how-ev-er" /haʊ'ev.ər/.

The rules of syllabification in English are not always straightforward due to the language's irregular spelling and pronunciation patterns. Unlike languages with more consistent orthographies, such as Spanish, English does not always have clear-cut rules for syllable division. However, general guidelines can be applied, such as dividing between consonants when a vowel is followed by another vowel (e.g., "create" → "cre-ate") or splitting a word into its morphemes (e.g., "unhappiness" → "un-hap-pi-ness"). These rules, though helpful, have numerous exceptions due to the complex nature of English phonology (Treiman & Zukowski, 1990).

#### 4. Methodology and Data Collection

A comparative quantitative approach is employed because the study utilizes numerical evidence to compare and evaluate the performance of ChatGPT and Gemini in syllabification. Additionally, the results are in percentages and presented quantitatively. The sample consisted of 100 English words sourced from Peter Roach's *English Phonetics and Phonology*, 4th ed., and *Oxford Advanced Learner's Dictionary*, 10<sup>th</sup> edition.

A selection of words will focus on words with more than one syllable, because as the study aims to assess the accuracy of syllable division, monosyllables do not provide this insight; therefore, the more complex the word, the more difficult it is to break it down into syllables. Furthermore, the words are classified into various grammatical categories, with twenty words from each category: nouns, verbs, adjectives, adverbs, and compound words. Part of speech matters in syllabification research because it reflects differences in morphology and helps test AI models across a variety of structures. Different parts of speech often follow different morphological rules, that is, the way words are built, usually by suffixes, for example:

Nouns often end in "-tion", "-ment", or "-ness" (e.g., action, movement, happiness).

Verbs frequently end in "-ing" or "-ed" (e.g., running, played).

Adjectives might end in "-ous", "-ful", or "-ive" (e.g., dangerous, hopeful, active).

Adverbs often end in "-ly" (e.g., quickly).

This approach ensures a balanced mix and allows for a more comprehensive analysis of the models' performance with different types of words. This study analyzes only a 20-word sample from the 100-word dataset, as examining the full set in detail would make the study excessively lengthy. However, the overall results and comparisons are based on all 100 words, with the complete dataset included in the appendix.

The process involves putting each selected word into ChatGPT and Gemini individually. Each word was delivered separately, using a standardized prompt to bring out the syllabification, and the prompt was: "Provide the phonemic transcription of 'the word,' including its syllabification according to the IPA". By comparing the models' processes with each word separately, the study can ensure that all information about the words' processing by the models was done in isolation, thus reducing any influence due to context and providing a fair comparison. All the 100 words used by the models were documented as an appendix at the end of the study.

The main reasons why this study depends on Roach as a source are that he guided the development of the modern knowledge about English pronunciation that links the aspects of theoretical phonemics and the practical terms of teaching language. Another reason is that Roach's book "English Phonetics and Phonology" is the official textbook studied by all second-year students at the English department, College of Education at Salahaddin University.

The last one is a contributing author to the widely recognized Cambridge English Pronouncing Dictionary, an important reference work in English phonetics and pronunciation. Briefly, he has a unique talent for making very complex phonological terms accessible and simple.

## 5. Data Analysis

For each part of speech, there is a table that presents a comparative analysis consisting of 20 phonemic transcriptions of selected words, but here the study shows only 3 words in each table, the rest at the end of the study, which were generated by two AI models, ChatGPT and Gemini. The prompt that asked the two AI models to show the transcription is "Provide a transcription of the word 'the word,' indicating its syllables according to the IPA."

The following tables provide a comparative analysis of the syllabification results generated by ChatGPT and Gemini. Each entry illustrates how the two AI systems parse the syllables of a given word, highlighting both similarities and differences in their phonological segmentation.

**Table 1: Instances of Verbs with Syllabification by ChatGPT and Gemini**

<i>Words</i>	<i>ChatGPT</i>	<i>Gemini</i>
<b>Commandeer</b>	/kɒm.ən'diə/	/kə'mɑ:n.diə(r)/
<b>Imagine</b>	/ɪ'mædʒ.ɪn/	/ɪ'mæ.dʒɪn/
<b>Participate</b>	/pɑ:'tɪs.i.peɪt/	/pɑ:'tɪ.sɪ.peɪt/

The first word in Table 1, **commandeer**, is transcribed by ChatGPT as /kɒm.ən.diə/, dividing it into three syllables. Similarly, PR's transcription also reads /kɒm.ən.diə/, indicating that ChatGPT closely aligns with his authoritative phonemic standard. In contrast, Gemini provides a different transcription: /kə.mɑ:n.diə(r)/. Here, Gemini reduces the first vowel to /kə/ and alters the vowel quality in the second syllable to /mɑ:n/, but it still maintains a three-syllable division. This comparison shows that while all three sources agree on the number of syllables, ChatGPT aligns more closely with Roach in the specific segmentation and phonemic realization of each syllable.

Another word in Table 2 is "**Imagine**," reveals a subtle yet significant variation in syllabification between ChatGPT and Gemini. ChatGPT transcribes the word as /ɪ'mædʒ.ɪn/, while Gemini provides /ɪ'mæ.dʒɪn/. The main difference lies in syllabification: ChatGPT differs from the Roach because ChatGPT groups the "dʒ" sound with the first syllable (mædʒ), followed by the unstressed "ɪn", which violates Onset Maximization. In contrast, Gemini separates the syllables as "mæ" and "dʒɪn", which may suggest a slightly different internal division, possibly influenced by a more phonemic or simplified segmentation approach. Since this study uses PR's transcription as the reference, Gemini's version shows greater alignment in both syllable division.

The word "**participate**" reveals subtle yet important differences in syllabification between ChatGPT and Gemini. ChatGPT transcribes it as /pɑ:'tɪs.i.peɪt/, treating the second syllable as /tɪs/ with a separate reduced vowel /i/ in the third syllable, potentially suggesting a looser syllabic boundary. In contrast, Gemini offers the transcription /pɑ:'tɪ.sɪ.peɪt/, clearly segmenting each syllable with the /sɪ/ as a full syllable rather than a continuation of the previous one. This aligns precisely with PR's transcription /pɑ:'tɪ.sɪ.peɪt/, where the word is divided into four syllables.

**Table 2: Instances of Adjectives with Syllabification by ChatGPT and Gemini**

<i>Words</i>	<i>ChatGPT</i>	<i>Gemini</i>
<b>bottled</b>	/'bɒt.əld/	/'bɒt.əld/
<b>Dangerous</b>	/'deɪn.dʒərəs/	/'deɪn.dʒərəs/
<b>Decimal</b>	/'des.i.məl/	/'de.sɪ.məl/

The first word, "bottled", highlights a key difference between the AI models and the standard reference by PR. Both ChatGPT and Gemini syllabify the word as /'bɒt.əlɪd/, indicating a two-syllable structure where the second syllable includes a schwa /ə/ followed by /l/ and /d/. This suggests the presence of a vowel-led syllable in the final position. However, PR's transcription /'bɒt.lɪd/ presents a more phonemically economical structure by using a syllabic consonant /l/ as the nucleus of the second syllable. This results that sometimes like this example AI models do not recognize syllabic consonants and will not transcribe the word in a way to consider a syllabic consonant to become the peak of a syllable.

The word "dangerous" reveals a subtle difference between the outputs of ChatGPT and Gemini. ChatGPT syllabifies it as /'deɪ.n.dʒərə.s/, where the medial syllable blends the /r/ with a reduced schwa, potentially resulting in a slightly compressed or less distinct three-syllable structure. This creates ambiguity in the syllable boundary between the second and third syllables.

In contrast, Gemini provides the syllabification /'deɪ.n.dʒə.rəs/, which does not align with PR's standard transcription. Here, the syllable boundaries are marked as maintaining a transparent three-syllable division.

The last word in the table, "decimal", ChatGPT transcribes as /'des.i.məl/, dividing the word into three syllables. Gemini, similarly, provides /'de.si.məl/ with three syllables, but differs slightly in its segmentation by placing the /s/ in the second syllable, resulting in /de.si.məl/ rather than /des.i.məl/. In contrast, PR transcribes the word as /'de.si.məl/, where the final syllable /məl/ contains a syllabic consonant /l/, indicating that the final syllable is not formed with a vowel but with a syllabic /l/ sound.

**Table 3: Instances of Nouns with Syllabification by ChatGPT and Gemini**

<i>Words</i>	<i>ChatGPT</i>	<i>Gemini</i>
<i>Capital</i>	/ 'kæp.i.təl/	[ 'kæp.i.tl]
<i>Celebration</i>	/sel.i'breɪ.fən/	/sel.ə'breɪ.fən/
<i>Disaster (N)</i>	/dɪ'zɑ:.stə/	/dɪ'zɑ:.s.tə(r)/

The first word in Table 3, which is "capital", consists of three syllables and is transcribed slightly differently by ChatGPT and Gemini. ChatGPT transcribes it as /'kæp.i.təl/, representing the final syllable with a full vowel /ə/ followed by the consonant /l/. In contrast, Gemini provides the transcription /'kæp.i.tl/, in the final syllable features a syllabic consonant /l/, where the /l/ sound serves as the syllable nucleus, replacing the vowel. This version aligns with the transcription found in PR's book, where "capital" is also given as /'kæp.i.tl/. The use of the syllabic /l/ in Roach's transcription and Gemini's output represents a more accurate reflection of how native speakers tend to articulate the word, often omitting the schwa sound before /l/.

All sources ChatGPT, Gemini, and PR, agree that the word "celebration" consists of four syllables: cel-e-bra-tion. The syllable boundaries are consistently marked as /sel.ə'breɪ.fən/ or /sel.i'breɪ.fən/, with the primary stress placed on the third syllable /'breɪ/. This uniformity in syllable count reflects a shared structural understanding of the word's phonological composition.

However, the key difference lies in the vowel quality of the second syllable, which directly affects how that syllable is interpreted and pronounced. ChatGPT uses /ɪ/, resulting in /sel.i'breɪ.fən/, while both Gemini and PR use the schwa /ə/, yielding /sel.ə'breɪ.fən/. From a syllabification perspective, the choice of /ə/ is more appropriate for an unstressed syllable in British English, as it reflects a reduced vowel that naturally occurs in fluent speech. By contrast, ChatGPT's use of /ɪ/ suggests a slightly more pronounced syllable, which may not align with the reduced quality expected in standard phonological analysis. Therefore, Gemini's syllabification more closely aligns with Roach's model, reinforcing the accuracy of the schwa in unstressed positions.

The word "disaster" ChatGPT transcribes as /dɪ'zɑ:.stə/, dividing it into three syllables, which aligns precisely with PR's transcription of /dɪ'zɑ:.stə/. This consistency suggests that ChatGPT accurately follows standard British English pronunciation norms. In contrast, Gemini transcribes the word as /dɪ'zɑ:s.tə(r)/, where the syllable boundary places the /s/ at the end of the stressed syllable "zɑ:s" and includes an optional /r/ at the end. The syllabification in Gemini differs slightly, merging the /s/ with the stressed syllable. Using PR as a reference point, it is clear that ChatGPT's transcription is more aligned with traditional British phonemic standards.

**Table 4: Instances of Adverbs with Syllabification by ChatGPT and Gemini**

<i>Words</i>	<i>ChatGPT</i>	<i>Gemini</i>
<b><i>Fortunately</i></b>	/ 'fɔ: . tʃən.ət.li/	/ 'fɔ: . tʃən.ət.li/
<b><i>Immediately</i></b>	/ɪ'mi: . di.ət.li/	/ɪ'mi:di. ət.li/
<b><i>Suddenly</i></b>	/ 'sʌd. ən.li/	[ 'sʌd. ŋ.li]

In comparing the syllabification of "fortunately" between ChatGPT, Gemini, and PR's source, both AI models divide the word into four syllables. However, the key distinction lies in the treatment of the second syllable. ChatGPT transcribes it as / 'fɔ: . tʃən.ət.li/, treating the unstressed vowel /ə/ as part of a single unit with the /n/, whereas Gemini segments the schwa into its syllable, transcribing it as / 'fɔ: . tʃən.ət.li/. This is consistent with PR's transcription, which also treats the second syllable as /tʃən.ət/. Despite this slight variation, all three sources align in the overall phonemic structure of the word.

In comparing the syllabification of "immediately", both ChatGPT and PR present the transcription as /ɪ'mi: .di.ət.li/, maintaining a clear four-syllable division. This structure distinctly separates the "di" /dɪ/ and "ət" /ət/ syllables, reflecting careful, articulated speech with fully realized vowel sounds. In contrast, Gemini offers a slightly compressed transcription: /ɪ'mi:di.ət.li/, where the boundary between the "di" and "ət" components is less distinct. This suggests a more fluid or reduced articulation, possibly merging the /d/ and /i/ into a single syllable. While all versions maintain the primary stress on the second syllable / 'mi:/, Roach and ChatGPT demonstrate a more traditional and precise syllabification, whereas Gemini reflects a more phonemically streamlined variant common in rapid or casual speech.

The word "suddenly" ChatGPT provides the transcription as / 'sʌd. ən.li/, where the second syllable contains a schwa /ə/, indicating a fully pronounced vowel. In contrast, Gemini uses / 'sʌd. ŋ.li/, where the second syllable /ŋ/ represents a syllabic nasal, meaning the /n/ sound forms an entire syllable without a separate vowel. PR's transcription / 'sʌd. ŋ.li/ aligns with Gemini's approach, suggesting that the syllabic nasal /ŋ/ is a common feature in relaxed speech, where unstressed vowels are often omitted.

**Table 5: Instances of Compound Words with Syllabification by ChatGPT and Gemini**

<i>Words</i>	<i>ChatGPT</i>	<i>Gemini</i>
<b><i>bad-tempered</i></b>	/bæd 'tɛm. pəd/ or /bæd 'tɛm. pɜ:d/	/bæd'tɛm.pəd/
<b><i>heavy-handed</i></b>	/hev.i'hæn.dɪd/	/hɛ.vi'hæn.dɪd/
<b><i>typewriter</i></b>	/ 'taɪp.raɪ.tə/	/ 'taɪp. raɪ.tər/

The first compound word in Table 5, "bad-tempered", is a compound adjective consisting of two parts: bad and tempered. In terms of syllabification and stress placement, both ChatGPT and Gemini offer slightly different interpretations. ChatGPT transcribes it as either /bæd 'tɛm.pəd/ or /bæd 'tɛm.pɜ:d/, placing primary stress on the second syllable. This transcription also reflects some flexibility in vowel quality for the second syllable of tempered (either schwa /ə/ or long /ɜ:/). Gemini, on the other hand, provides /bæd'tɛm.pəd/, merging the compound into a more compact form while still placing the primary stress on the (tɛm) syllable, aligning more closely with PR's transcription: /bæd'tɛm.pəd/. Roach's version also reflects this compactness and stresses the first syllable of tempered without separating bad as an individual stress-bearing unit. In compassion, Gemini and Roach both treat bad-tempered as a single phonological unit, emphasizing the compound's second element (tempered) as central to its stress pattern.

For the compound word "heavy-handed", according to ChatGPT transcribes the word as /hev.i'hæn.dɪd/, presenting it as a compound with three syllables. Gemini transcribes it as /hɛ.vi 'hæn.dɪd/, which aligns more closely with PR, who also uses /hɛ.vi 'hæn.dɪd/ in his transcription. In terms of syllabification, all three sources recognize the word as consisting of four syllables, typically broken as /hɛ.vi 'hæn.dɪd/. Despite a slight vowel variation in ChatGPT's transcription, its syllabification aligns with Gemini and Roach, with Gemini matching Roach more closely, highlighting its reliability in this case.

In the syllabification of "typewriter", both ChatGPT and Gemini place the primary stress on the first syllable, /'taɪp/, which aligns with PR's guideline that compounds with a nominal first element typically carry primary stress on that initial component. ChatGPT transcribes the word as /'taɪp.raɪ.tə/, clearly dividing it into three syllables with a schwa /ə/ in the final unstressed syllable, consistent with standard British English pronunciation. Gemini's version, /'taɪp.raɪ.tər/, while similar in syllable division and stress placement, includes the final /r/ sound, reflecting a more rhotic (typically American-influenced) pronunciation. Roach's transcription, also /'taɪp.raɪ.tə/ mirrors ChatGPT's rendering, reinforcing its alignment with traditional British phonological norms and syllabification patterns.

**Table 6: Some Examples of Syllabification that both AI Models Transcribe Identically.**

<i>Word</i>	<i>ChatGPT</i>	<i>Gemini</i>
<b><i>Accident</i></b>	/ ' æk.sɪ.dənt/	/ ' æk.sɪ.dənt/
<b><i>Completely</i></b>	/kəm'pli:t.li/	/kəm'pli:t.li/
<b><i>Deliver</i></b>	/dɪ'lɪv.ə/	/dɪ'lɪv.ə/
<b><i>Interrupt</i></b>	/ɪn.tə'rʌpt/	/ɪn.tə'rʌpt/
<b><i>Open-ended</i></b>	/əʊ.pən'en.dɪd/	/əʊ.pən'en.dɪd/

In Table 6, the five samples show that the words given by both ChatGPT and Gemini fully match the syllabification rules mentioned in authoritative sources and the guidelines of RP pronunciation. This harmony in syllable division shows how smoothly they highlight the dependable and steady application of linguistic standards by these transcription tools.

## 6. Results and Discussion

Based on the analysis presented in the six tables above, a thorough comparison was conducted between the two AI models, selecting 20 words from a sample of 100 words. The findings, summarized in Table 7, highlight key differences and similarities, showing their respective performances and capabilities.

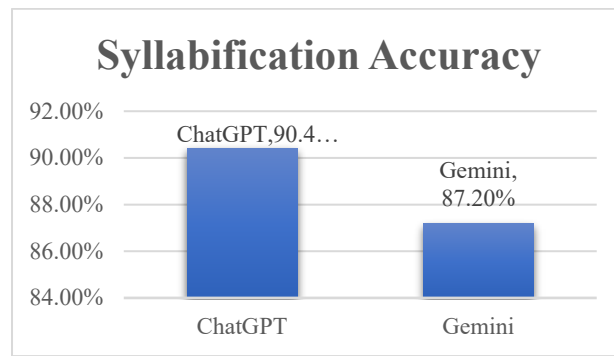
**Table 7: The Accuracy of Syllabification Results for Each AI Model**

<i>Tables</i>	<b>ChatGPT Accuracy</b>	<b>Gemini Accuracy</b>
<i>Table 1: Verbs</i>	90%	80%
<i>Table 2: Adjectives</i>	92%	94%
<i>Table 3: Nouns</i>	94%	94%
<i>Table 4: Adverbs</i>	90%	88%
<i>Table 5: Compound word</i>	86%	80%
<i>Total</i>	90.4%	87.2%

Although the methodology of this study states that only 20 words were selected as a sample, the five Tables shown in Table 7 present results based on the full set of 100 words, with 20 words from each category. Each Table provides a detailed breakdown of syllabification accuracy across different lexical word categories, revealing nuanced performance differences between ChatGPT and Gemini. ChatGPT and Gemini achieved a 94% accuracy rate for nouns. This suggests that both models handle nouns relatively well, likely due to their frequent occurrence in training datasets and consistent phonological patterns. However, the performance gap indicates that ChatGPT may better capture the syllabic boundaries of nouns, possibly due to its token-based architecture, which aligns more closely with written forms.

In contrast, verbs and adverbs showed more pronounced discrepancies. ChatGPT maintained a 90% accuracy rate for verbs, while Gemini dropped to 80%. This difference highlights Gemini's challenges with verb syllabification, which may involve more complex morphological variations (e.g., tense markers). For adverbs, both models performed comparably, with ChatGPT at 90% and Gemini at 88%, suggesting that adverbs' simpler structures pose fewer challenges. Interestingly, Gemini surpassed ChatGPT in adjective syllabification (94% vs. 92%), possibly due to adjectives' more predictable stress patterns and syllable boundaries.

The most notable divergence occurred with compound words, where Gemini (86%) slightly outperformed ChatGPT (82%). This reversal suggests that Gemini may better handle the concatenated nature of compound words, while ChatGPT's segmentation approach might struggle with their blended phonological features. Overall, the data underscores that while both models are highly reliable, their performance varies by word type, reflecting differences in training data prioritization and architectural biases. These insights are valuable for educators and developers aiming to leverage AI for linguistic tasks.



**Chart 1: Overall Accuracy Rates Between ChatGPT and Gemini**

Analysis of the data proves that ChatGPT and Gemini are good at determining where syllables start and stop in most English words. ChatGPT tends to break words into parts according to how they look in writing, rather than depending only on how the words are pronounced.

They show that LLMs can contribute to phonological learning and computational linguistics. These models do not fit all the phonotactic differences, specifically when related to elision, morphophonemic changes, or regional pronunciation.

Based on 100-word datasets using the Maximal Onset Principle, ChatGPT produced more reliable and accurate syllabified strings. It also adhered better to the standard references created by PR as well as the Oxford dictionary. On the other hand, Gemini sometimes failed to identify syllable boundaries clearly in verbs and compound words, slightly lowering its accuracy in such cases.

For instance, the primary reason for the discrepancy for the word “participate” is that English prefers open syllables (those ending in vowels) and onsets (syllables beginning with consonants), which supports Gemini’s transcription. However, ChatGPT’s division is also valid, as spoken English often allows final consonant clusters like /st/ or /s/ at the end of syllables.

The statistical significance of the difference can be best explained by the fact that syllabification is controlled by phonological rules such as Maximal Onset and Sonority Sequencing Principles. As these rules contain no variability, any deviation, including the wrongful interpolations of syllable boundaries and mis-identification of syllabic consonants, could not be claimed to be acceptable variation. So, the lack of fit between the models will not be only stylistic; it will be a phonological error.

So, these tools can make it easier for teachers to teach by giving students syllabic breakdowns immediately. Improving the AI models in text-to-speech systems can help developers create speech that flows more smoothly and has better prosody.

By using these observations, LLM developers can better design their tokenization and training approaches for improved performance in phonology. Using transcriptions of sounds for each word can also help the model improve.

The findings presented in the study demonstrate a consistent division of word specificity of word categories. e.g., ChatGPT has 90.4 % overall accuracy, but in Gemini, overall accuracy is 87.2 %.

This difference may not sound significant numerically, but in computation-based linguistics, small differences in percentage may prove statistically relevant, particularly when:

The sample can be controlled and diverse (such as the 100-word multisyllabic and morphologically differentiated sample employed in this case).

The errors are not just doubts but have definite patterns to make wrong what is considered standard phonologically.

Over all the words analyzed, ChatGPT and Gemini found 100% agreement on the total number of syllables for all words analyzed, indicating consistency in syllable count recognition. However, some syllabic consonants were skipped or misidentified in both AI models. Both of them worked best with nouns, adverbs, and adjectives, demonstrating high accuracy in these areas of speech.

## 7. Conclusions

The comparative analysis of syllabification between ChatGPT and Gemini revealed small, notable differences in the quality and consistency of outputs. Both models correctly identified the number

of syllables in all 100 multisyllabic English words, indicating a shared basic competence in syllable segmentation. However, the accuracy and linguistic validity of how those syllables were divided varied between the two. ChatGPT consistently demonstrated a stronger adherence to phonological principles, particularly in applying the Maximal Onset Principle, which helped produce syllabifications more in line with authoritative sources such as Peter Roach and the Oxford Advanced Learner's Dictionary.

One of the issues is double-checking words; the process of syllabification sometimes differs between AI models, especially Gemini, which does not always clearly indicate syllable boundaries. Another major challenge is that AI models can be greatly affected by the prompts they receive. The input they are given can significantly influence their responses, causing variations in the quality and relevance of the output.

In conclusion, while both AI systems show the possibility of handling basic syllabification, based on the results of 100-word syllabification, ChatGPT demonstrated more reliable alignment with established phonological standards and performed more consistently across various word types. Gemini, though capable, displayed specific weaknesses, particularly in dealing with compound words and maintaining consistency in more complex lexical items. This indicates that ChatGPT is currently the more dependable tool for phonological tasks like syllabification, especially for instructional or academic applications.

### Recommendations

The recommendation for educators and researchers should use these models as supplementary tools rather than standalone solutions. Pairing AI-generated outputs with human linguistic judgment can enhance both accuracy and reliability in phonetic analysis and instruction.

### Suggestions for Further Study:

- Future studies should consider testing AI performance on spoken language inputs and different English accents.
- Compare other AI models to see which are best for figuring out syllables.
- Research should extend beyond just syllabification to include other phonological topics like intonation and stress placements.

### References

- Davis, A. (2024). AI rising in higher education: opportunities, risks and limitations. *Asian Education and Development Studies*. <http://dx.doi.org/10.1108/AEDS-01-2024-0017>
- Duanmu, S., Kim, H.-Y., & Stiennon, N. (2005). Stress and Syllable Structure in English: Approaches to Phonological Variations. *Taiwan Journal of Linguistics*. [http://dx.doi.org/10.6519/TJL.2005.3\(2\).2](http://dx.doi.org/10.6519/TJL.2005.3(2).2)
- Fakih, H. (2020). The relationship between artificial intelligence and reducing the cognitive load of learners, *Zanco Journal of Human Sciences*, Vol. 24, No. 5 <https://doi.org/10.21271/zjhs.24.s5.34>.
- Giegerich, H. J. (1998). *English Phonology*. Cambridge University Press.
- Ladefoged, P., Johnson, K. (2014). *A Course in Phonetics*, 6th ed. Cengage Learning.
- McMahon, A. (2002). *An Introduction to English Phonology*. Edinburgh University Press.
- Mohammadkarimi, E. (2024). Exploring the Use of Artificial Intelligence in Promoting English Language Pronunciation Skills. *Journal on Language and Language Teaching* 27, 98–115. <https://doi.org/10.24071/llt.v27i1.8151>
- Parker, S.G. (2002). Quantifying the sonority hierarchy. PhD dissertation. University of Massachusetts Amherst.
- Roach, P. (2009). *English Phonetics and Phonology*. Cambridge University Press.
- Tian J. (2004). *Data-driven Approaches for Automatic Detection of Syllable Boundaries*. 8th International Conference on Spoken Language Processing, ICSLP pp. 61-64 <http://dx.doi.org/10.21437/Interspeech.2004-60>
- Treiman, R. and Zukowski, A. (1990). Toward an understanding of English syllabification. *Journal of Memory and Language*, 29(1), pp.66–85. [http://dx.doi.org/10.1016/0749-596X\(90\)90010-W](http://dx.doi.org/10.1016/0749-596X(90)90010-W)
- Vančová, H. (2023). AI and AI-powered tools for pronunciation training. *Journal of Language and Cultural Education* 11, 12–24. <https://doi.org/10.2478/jolace-2023-0022>
- Wattpad (n.d.) *Sonority Hierarchy*. [online] Available at: <https://www.wattpad.com/814152454-conlang-001-kjamir-sonority-hierarchy>

## برگه‌سازی له زمانی ئینگلیزی له ریگه‌ی زیره‌کی ده‌ستکرد: لیکولینه‌وه‌یه‌کی به‌راوردکاریه له Gemini و ChatGPT

به‌شی ریگاو بان، کولێژی ته‌کنه‌لوژی، زانکۆی پۆلیته‌کنیک،  
 به‌شی ئینگلیزی، کولێژی په‌روه‌ده، زانکۆی سه‌لاحه‌ددین-  
 هه‌ولێر، هه‌رێمی کوردستان، عێراق  
 askander.abduallah@su.edu.krd

اسکندر خالد عبدالله  
 به‌شی ریگاو بان، کولێژی ته‌کنه‌لوژی، زانکۆی پۆلیته‌کنیک،  
 به‌شی ئینگلیزی، کولێژی په‌روه‌ده، زانکۆی سه‌لاحه‌ددین-  
 هه‌ولێر، هه‌رێمی کوردستان، عێراق  
 Pakhshan.hamad@su.edu.krd

## پوخته

ئهم توێژینه‌وه‌یه ئه‌ركی برگه‌سازی زمانی ئینگلیزی ده‌خاته به‌ر لیکولینه‌وه، که تبايدا سه‌رنج ده‌خاته سه‌ر دوو مۆدیلی پیشکه‌وتووی زیره‌کی ده‌ستکرد Gemini و ChatGPT. برگه‌سازی بریتیه له پرۆسه‌ی دابه‌شکردنی ووشه‌کان به‌سه‌ر بزۆینه پیکه‌ینه‌ره‌کانیادا و که یه‌کیکه له چه‌مه‌که سه‌ره‌کیه‌کانی فۆنۆلوژی. ئهم چه‌مه‌که‌ش پۆلیکی گرنگی له بواری فیربوونی زماندا هه‌یه، دروستکردنی سیسته‌می ئاخاوتن، زمانه‌وانی کۆمپیوته‌ر، که تیگه‌یشتنیکی پوون له پیکه‌اته‌ی بزۆینه‌کان ده‌توانیت فیربوونی زمان و گه‌شه‌سەندنی ته‌کنه‌لوژیا له بواری زمانه‌وانیدا به‌رز بکاته‌وه. شیکاریه‌کانی ووشه‌کان له‌سه‌ر بنه‌مای کتیبه‌کانی پیتەر رۆچ، فونیتیک و فونۆلوژی و فره‌هنگی فیرخاوانی پیشکه‌وتووی ئۆکسفۆرد، چاپی ۱۰ ئەنجام دراو. له توێژینه‌وه‌که‌دا کومه‌له‌ داتایه‌ک به‌ وردی هه‌لێژێردراوه و به‌کارهاتووه که 100 وشه‌ی ئینگلیزی فره بزۆین له هه‌موو پۆله ووشه‌یه‌کان له‌خۆده‌گریت. پاشان ووشه‌کان به‌ وردی شیکراونه‌ته‌وه بۆ هه‌لسه‌نگاندنی وردی به‌شکردن و نیشاندانی جیاوازیه‌کانی هیماسازی شیکراونه‌ته‌وه. ئه‌مه‌ش له ریگه‌ی شیکاریه‌کی به‌راوردکاری گشتگیره‌وه ئه‌نجامدراوه، توێژینه‌وه‌که هه‌وایه‌شی و جیاوازیه‌کانی له شیوازه‌کانی هیماسازی ده‌ستنیشان کردووه، که له‌لایه‌ن دوو مۆدیلی AI هه‌ نیشان دراو. ئهم توێژینه‌وه‌یه به‌شداریه‌کی گرنگی له بواری زمانه‌وانی و فونۆلوژی ده‌کات، هه‌روه‌ها تیشک ده‌خاته سه‌ر توانا و سنووره‌کانی مۆدیه‌کانی AI له بواری شیکاری زمان. ئه‌نجامه‌کان پیشنیاری ئه‌وه ده‌کهن، که ChatGPT ده‌توانیت به‌ شیوه‌یه‌کی کاریگه‌رتر له Gemini له ئه‌رکه‌کانی هیماسازی به‌کاربه‌یتریت، که له بواری ته‌کنه‌لوژیادا زمانه‌ره‌خه‌ری پیشکه‌وتنی زیاتر به‌خۆیه‌وه ده‌بینیت.

وشه سه‌ره‌کیه‌کان: برگه‌سازی، زیره‌کی ده‌ستکرد، Gemini، ChatGPT

## التقسيم إلى مقاطع في اللغة الإنجليزية من خلال الذكاء الاصطناعي: دراسة مقارنة بين Gemini و ChatGPT

اسکندر خالد عبدالله  
 قسم انشاء الطرق، كلية تكنولوجيا، جامعة  
 بوليتكنيك، اربيل، العراق  
 askander.abduallah@su.edu.krd

بخشان اسماعيل حمد  
 قسم اللغة الانكليزية، كلية التربية، جامعة صلاح  
 الدين-اربيل، إقليم كردستان، العراق  
 Pakhshan.hamad@su.edu.krd

## الملخص

تسعى هذه الدراسة إلى استكشاف فعالية الذكاء الاصطناعي في مهمة تقسيم الكلمات الإنجليزية إلى مقاطع لفظية، وذلك من خلال مقارنة أداء نموذجين متقدمين وهما Gemini و ChatGPT. يُعد التقسيم إلى مقاطع لفظية أحد المفاهيم الأساسية في التحليل الفونولوجي، وله أهمية بالغة في مجالات متعددة مثل تعليم اللغة الإنجليزية، وتوليف الكلام، واللغويات الحاسوبية. يهدف هذا البحث إلى تقييم مدى دقة النموذجين في تنفيذ هذه المهمة اللغوية الدقيقة، مع التركيز على الجوانب التي تعزز أو تعيق الأداء اللغوي المدعوم بالذكاء الاصطناعي. يعتمد التحليل على إطار مرجعي علمي يتمثل في كتاب "علم الأصوات وعلم الفونولوجيا" لبينتر روتش، بالإضافة إلى "قاموس أكسفورد للمتعلمين المتقدمين" (الطبعة العاشرة) كمصدر معياري. تم اختيار عينة بحثية مكونة من 100 كلمة إنجليزية متعددة المقاطع، موزعة بشكل متوازن على الفئات المعجمية الأساسية (أسماء، أفعال، صفات، ظروف... إلخ). وقد خضعت هذه العينة لتحليل دقيق باستخدام النموذجين، بهدف قياس دقة كل نموذج في تحديد حدود المقاطع الصوتية. أظهرت النتائج من خلال تحليل مقارن مفصل أن كلا النموذجين قدما أداءً جيدًا في تحديد عدد المقاطع، إلا أن ChatGPT تفوق من حيث الدقة والاتساق، خاصة في تطبيقه لمبدأ "بداية المقطع القصوى" ومراعاته للمعايير المعجمية والفونولوجية المعتمدة. تسلط الدراسة الضوء على الإمكانيات الكبيرة التي تنتجها نماذج الذكاء الاصطناعي في مجال التحليل الصوتي، كما تكشف في الوقت نفسه عن بعض القيود التي ينبغي أخذها في الاعتبار عند استخدام هذه التقنيات في التطبيقات اللغوية المتقدمة.

الكلمات المفتاحية: التقسيم المقطعي، الذكاء الاصطناعي، Gemini، ChatGPT

**Appendix of the 100 Words in The Study:**

<b>No.</b>	<b>Word syllables</b>	<b>No.</b>	<b>Word syllables</b>
1.	Abandon	51.	Holocaust
2.	Abnormal	52.	Imagine
3.	Abstract (Adj.)	53.	Immediately
4.	Accident	54.	Important
5.	Accidentally	55.	Injure
6.	Accurate	56.	Injurious
7.	Advantage	57.	Insolent
8.	Advantageous	58.	Interrupt
9.	Amaze	59.	Japan
10.	Amazing	60.	Japanese
11.	Bad-tempered	61.	Journal
12.	Beautifully	62.	Journalese
13.	Bottle	63.	Listened
14.	Bottled	64.	Long-suffering
15.	Burden	65.	Loudspeaker
16.	Capital	66.	Negative
17.	Car ferry	67.	Observer
18.	Carnation	68.	Obviously
19.	Celebration	69.	Open-ended
20.	Certainly	70.	Paradise
21.	Cigar	71.	Participate
22.	Cigarette	72.	Particularly
23.	Climate	73.	Perceive
24.	Climatic	74.	Photo
25.	Comfort	75.	Photography
26.	Comfortable	76.	Picture
27.	Completely	77.	Picturesque
28.	Commandeer	78.	Poison
29.	Dangerous	79.	Poisonous

30.	Decimal	80.	Red-blooded
31.	Deliver	81.	Reflex
32.	Derelict	82.	Reflexive
33.	Determine	83.	Refuge
34.	Disaster	84.	Refugee
35.	Disciple	85.	Refusal
36.	Downstream	86.	Refuse (V)
37.	Elicit	87.	Resurrect
38.	Enormous	88.	Standardize
39.	Enter	89.	Second-class
40.	Entertain	90.	Suddenly
41.	Eventually	91.	Suggest
42.	Extra	92.	Suitcase
43.	Fantastic	93.	Three-wheeler
44.	Frightening	94.	Tomato
45.	Five-finger	95.	Transistor
46.	Fortunately	96.	Typewriter
47.	Fruitcake	97.	Unbelievable
48.	Furniture	98.	Unlikely
49.	Headquarters	99.	Wonder
50.	Heavy-handed	100.	Wonderful