The Theory of Sentencoids

Created by Jacek Hawrot

May 31, 2013

May 16, 2014

Abstract

In this paper, I will present the theory of sentencoids that I have created. The paper contains four sections, and over 40 subsections. Some subsections consist of smaller units too.

The first section presents the structure of the theory of sentencoids,

the second section describes how to morph nouns, verbs, attributive adjectives, and wh–adverbs,

the third section contains a text that will be used to illustrate how the theory of sentencoids works. Some sentences from the text are analyzed again; this time according to the grammar of the English language,

the fourth section explains the similarities, and differences between the grammar of the English language, and the theory. The section also points the topics for the further research.

Changes that have been made

The following list presents the most important changes that have been made:

text has been corrected,

more examples have been added in order to make the theory easier to understand,

the glossary of terms has been added,

the list of assumptions has been added,

the status of a coordinator has been described,

the status of a subordinator has been described,

the subsection describing phrasal verbs has been removed. The verbs require more detailed research,

the language morphing of attributive adjectives has been described,

the language morphing of the selected wh–adverbs has been described,

the second text from the third section was removed. I have decided to focus more on the first text,

the analysis of a created sentencoid has been corrected. The analysis contained an error,

the subsection that confronts the theory with the grammar of the English language has been added,

the index of terms has been added.

Introduction

The proper use of articles can lead to confusion during the process of learning English as a second language. The same situation happened to me; however, I have asked myself the following question—is it possible to transform the problem into a solution? If yes, it may support persons that learn English as a second language. This idea has motivated me to create a new approach that makes possible to create grammatically correct texts.

The theory describes a new way of classifying sentences—the criterion for classification is based on the quantity of words that a sentence contains. The traditional grammar terms: a phrase, and clause have been replaced in favor of the term “sentencoid.” The term is used to name the unit that stands between words, and sentences. The theory describes two types of borders that are introduced by punctuation marks. One type of border allows to divide a sentence, and the second one signals where the previous sentences ends, and where the next one starts.

The theory, however, does not totally replace the rules that govern the grammar of the English language. The task of the theory is to provide a simple way that allows to write, and analyze English texts.

Glossary of terms

The following glossary briefly explains the terms that appear in the paper. The terms appear in the alphabetical order.

Advanced language addition: in this type of language addition, the repeated words must be removed before the created sentencoids are added in the same order they were created. Advanced language addition consists of three stages.

Classification Paradox: a situation in which a sentence that contains a proper noun has more words that it should contain according to its type.

External boundary: a boundary that signalizes the end of the previous sentence, and the beginning of the next original sentence. External boundaries are introduced by exclamation points, interrobangs, periods, and question marks.

First level sentence (FLS): a sentence that contains maximum four words.

Fourth level sentence (FhLS): a sentence that contains at least 15 words.

Internal boundary: a boundary that is present inside an original sentence. The internal boundary makes possible to divide the original sentence into parts called sentencoids. Internal boundaries are introduced by commas, en–dashes, parentheses, quotation marks, and semicolons.

Language addition: a type of addition in which not numbers are added to other numbers, but words (sentencoids) are added to other words (other sentencoids). They are two types of language addition: simple, and advanced language addition.

Language morphing: the change of a word into another word within the same word class. Morphing within each word class has its own limitations.

Morphing into nothingness: a type of language morphing in which word is removed from the sentence containing it.

Negative result (the negative result of search): it means that the searched sentencoid was found less than three times in the available language material.

Original sentence: a sentence that is going to be divided into sentencoids, in order to check its correctness.

Parenthetical material: any text inside a pair of parentheses.

Positive result (the positive result of search): it means that the searched sentencoid was found at least three times in the available language material.

Repeated Words (RW): the last word(s) from the previous sentencoid that is (are) repeated at the beginning of the next sentencoid. RW can take positive values, (at least one word was repeated) or the neutral value (no word was repeated).

Second level sentence (SLS): a sentence that contains five to maximum seven words.

Sentencoid: a sequence of words that usually consists of maximum four words.

Simple language addition: in this type of language addition, sentencoids are added in the same order they were created to form the original sentence. Simple language addition consists of two stages.

The available language material: a set of grammatically correct sentences.

Third level sentence (TLS): a sentence containing eight to maximum 14 words.

The process of breaking and re–creating original sentences: a process consisting of six phases, during which an original sentence is divided into sentencoids, and re–created again.

The short passive: a passive construction in which the performer of an action is absent.

Zero result: the most extreme variant of the negative result. This result signalizes that the searched sentencoid is not present in the available language material.

List of assumptions

The following list contains assumptions on which the theory of sentencoids is based:

[1] A text can be divided into parts called sentencoids. A sentencoid usually consists of maximum four words.

[2] The analysis of a sentencoid consisting of one word requires one line.

[3] No matter how many words an idiom, a proverb, and a saying may contain, each of them can be considered as:

1) one word

2) a sentencoid consisting of only one word.

However, any change will transform its status.

[4] A combination of various letters, marks, symbols, words, etc. can be considered as:

1) one word,

2) a sentencoid consisting of only one word.

[5] A list of words is always classified as a sentencoid that consists of only one word.

[6] A proper noun that consists of more than just one word cannot be divided into parts; words belonging to the proper noun cannot be present in adjacent sentencoids. All words that belong to a proper noun are present in the same sentencoid.

[7] No matter how many words a proper noun contains, it can be considered as:

1) one word,

2) a sentencoid consisting of only one word.

[8] A sentencoid that consists of only one word is not searched in the available language material.

[9] A coordinator always stands as a separate sentencoid.

[10] No matter how many words a subordinator contains, it always creates a sentencoid that stands alone.

[11] No matter how a number is written (3; 6/2; 3,0; 3.0000000000; etc.) it can be considered as:

1) one word,

2) a sentencoid consisting of only one word.

[12] According to the theory of sentencoids, they are two types of boundaries: external, and internal boundaries. Both types of boundaries are introduced by punctuation marks; however, each punctuation mark introduces only one type of border.

[13] External boundaries are introduced by exclamation points, interrobangs, periods, and question marks.

[14] Internal boundaries are introduced by commas, en–dashes, parentheses, quotation marks, and semicolons.

[15] A period located next to an abbreviation does not introduce the external border.

[16] The first pair of quotation marks begins an independent content within a sentence, and the second pair of quotation marks ends the content.

[17] The content inside a pair of quotation marks does not create sentencoids with the content outside both pairs of quotation marks.

[18] A compound word is always considered as one word.

[19] The content inside a pair of parentheses is analyzed independently from the content outside parentheses—i.e. there is no sentencoid that contains words from both the content inside, and outside parentheses.

[20] Each type of sentence is divided into sentencoids. The only exception takes place when a sentence consists of maximum four words, and does not contain any introducer of the internal border.

[21] The value of RW determinates the number of lines that language addition will contain. The neutral value of RW signals simple language addition; the positive value of RW signals advanced language addition.

[22] Words that interrupt the flow of the sentence are always analyzed separately.

[23] A word can be morphed into another word when both of them are members of the same word class.

1 The structure of the theory of sentencoids

The section describes terms, and assumptions on which the theory of sentencoids is based. The section also makes references to the third section.

1.1 The new classification

The reason for the new classification of sentences was to replace the present criterion that is based on the presence of clauses together with their coordinators, and subordinators. I have assumed that the more words sentence contains, the more it is problematic for a person that learns English as a second language. That is why the criterion for the new classification of sentences is based on the quantity of words that a sentence contains.

Sentences have been categorized into four types. The status of a sentence is based on the quantity of words that the sentence contains; the letter "W" refers to the quantity. The classification follows:

1) the acronym "FLS" stands for a first level sentence ($W \leq 4$)

2) the acronym "SLS" stands for a second level sentence ($5 \leq W \leq 7$)

3) the acronym "TLS" stands for a third level sentence ($8 \leq W \leq 14$)

4) the acronym "FhLS" stands for a fourth level sentence ($W \geq 15$)

1.2 The same content, but different types of sentences

The same content can be often written in different ways, and each of them will be grammatically correct. The following examples illustrate this question:

(1a) Do you want to fight me?
(1b) Do you want to fight against me?

(2a) One of my friends won this prize.
(2b) A friend of mine won this prize.

(3a) We will visit Paris.
(3b) We will visit the capital of France.

(4a) I am 26.
(4b) I am 26 years old.

(5a) How will this effect the process of learning English?
(5b) How will this effect learning English?

(6a) You are the winner!
(6b) You are victorious!

(7a) I live in Seattle.
(7b) I live in the city of Seattle.

(8a) Her dream is to visit Europe.
(8b) Her dream is to visit the continent of Europe.

As we can see, the same content can be written in more than one way.
Each sentence will be now classified:

(1a) Do you want to fight me?
Type: SLS ($5 \leq W \leq 7$)

(1b) Do you want to fight against me?
Type: SLS ($5 \leq W \leq 7$)

(2a) One of my friends won this prize.
Type: SLS ($5 \leq W \leq 7$)

(2b) A friend of mine won this prize.
Type: SLS ($5 \leq W \leq 7$)

(3a) We will visit Paris.
Type: FLS ($W \leq 4$)

(3b) We will visit the capital of France.
Type: SLS ($5 \leq W \leq 7$)

(4a) I am 26.
Type: FLS ($W \leq 4$)

(4b) I am 26 years old.
Type: SLS ($5 \leq W \leq 7$)

(5a) How will this effect the process of learning English?
Type: TLS ($8 \leq W \leq 14$)

(5b) How will this effect learning English?
Type: SLS ($5 \leq W \leq 7$)

(6a) You are the winner!
Type: FLS ($W \leq 4$)

(6b) You are victorious!
Type: FLS ($W \leq 4$)

(7a) I live in Seattle.
Type: FLS ($W \leq 4$)
(7b) I live in the city of Seattle.
Type: SLS ($5 \leq W \leq 7$)

(8a) Her dream is to visit Europe.
Type: SLS ($5 \leq W \leq 7$)

(8b) Her dream is to visit the continent of Europe.
Type: TLS ($8 \leq W \leq 14$)

The same content may be classified as different types of sentences, as illustrated in the sentences 7a and 7b. It may also happen that the same content will be classified as two sentences of the same type; however, these sentences may contain different quantities of words.

1.3 Sentencoids

A sentencoid is the basic unit that is present in every text. The following assumption describes the status of a sentencoid:

[1] A text can be divided into parts called sentencoids. A sentencoid usually consists of maximum four words.

The following examples explain how to read sentencoids that were created according to the pattern:

s 1 The first sentencoid,
s 2 The second sentencoid,
s 5 The fifth sentencoid.

1.4 Sentencoids consisting of one word

The following assumption describes the status of a sentencoid consisting of one word:

[2] The analysis of a sentencoid consisting of one word requires one line.

Such sentencoids do not create sentencoids with other words. The short analyses of the following sentences illustrate this:

(9) Tom likes horses, and cats.

9.3 The list of the created sentencoids

9.3.1 sentencoid 1: Tom likes horses
9.3.2 sentencoid 2: and
9.3.3 sentencoid 3: cats

(10) Our research team consists of chemists, and physicists

10.3 The list of the created sentencoids

10.3.1 sentencoid 1: Our research team consists
10.3.2 sentencoid 2: consists of chemists
10.3.3 sentencoid 3: and

10.3.4 sentencoid 4: physicists

(11) His answer was “Yes, or no.”

11.3 The list of the created sentencoids

11.3.1 sentencoid 1: His answer was
11.3.2 sentencoid 2: Yes
11.3.3 sentencoid 3: or
11.3.4 sentencoid 4: no

1.5 Idioms, proverbs, and sayings

The presence of idioms, proverbs, and sayings is a source of the classification paradox. A first level sentence consists of maximum four words; however, an idiom, a proverb, and a saying may consist of many words—consider the following sentences:

Do not cry over spilt milk;

This time no news is good news;

Remember, a bird in the hand is worth two in the bush.

The following assumption has been made to solve the problem related with the presence of an idiom, a proverb, and a saying in any type of sentence:

[3] No matter how many words an idiom, a proverb, and a saying may contain, each of them can be considered as:

1) one word

2) a sentencoid consisting of only one word.

However, any change will transform its status.

The analyses of the two following dialogs will illustrate this problem:

The first dialog contains the original idiom “a piece of cake,” and its altered version: ”a piece of pizza”:

- Uff! Making that big pizza was pretty easy and fast.
- I agree with you. It was a piece of cake!
- No, it was a piece of pizza! Ha ha ha
- Ha ha ha ha

The analysis of the original idiom:
It was a piece of cake! → Type: FLS ($W \leq 4$)

the first word: “It”

the second word: “was”

the third word: “a piece of cake” (an unchanged proverb that is equal to only one word)

$1 + 1 + 1 = 3$

$3 \leq 4 \rightarrow$ meets the condition to be classified as a first level sentence

The analysis of the altered version of the idiom:

No, it was a piece of pizza! → Type: SLS ($5 \leq W \leq 7$)

the first word: “No”

the second word: “it”

the third word: “was”

the fourth word: “a”

The fifth word: “piece”

the sixth word: “of”

the seventh word: “pizza”

$1 + 1 + 1 + 1 + 1 + 1 + 1 = 7$ (a changed proverb loses its unique status)

$7 \rightarrow (5 \leq W \leq 7) \rightarrow$ meets the condition to be classified as a second level sentence.

The second dialog contains the original idiom “Van Gogh’s ear for music,” and its altered version: “Tom’s ear for music”:

- Tom that music you played on the guitar was horrible! You have Van Gogh’s ear for music!
- Do not tell me that I am so horrible…
- No, you are not. You just have Tom’s ear for music!
- Very funny

The analysis of the original idiom:
You have Van Gogh’s ear for music! → Type: FLS ($W \leq 4$)

the first word: “You”

the second word: “have”

the third word: “Van Gogh’s ear for music” (an unchanged proverb that is equal to only one word)

1 + 1 + 1 = 3

$3 \leq 4 \rightarrow$ meets the condition to be classified as a first level sentence

The analysis of the altered version of the idiom:

You just have Tom's ear for music! → Type: SLS ($5 \leq W \leq 7$)

the first word: "You"

the second word: "just"

the third word: "have"

the fourth word: "Tom's"

the fifth word: "ear"

the sixth word: "for"
the seventh word: "music"

1 + 1 + 1 + 1 + 1 + 1 + 1 = 7 (a changed proverb loses its unique status)

$7 \rightarrow (5 \leq W \leq 7) \rightarrow$ meets the condition to be classified as a second level sentence.

Conclusion: any alternation of an idiom changes the status of the idiom. The idiom will no longer be classified as one word, or as a sentencoid that consists of only one word. The same rule applies to proverbs, and sayings.

1.6 The case of Latin proverbs and sayings

The case of Latin proverbs, and sayings is similar to the case of English proverbs, and sayings. However, the difference is that Latin proverbs cannot be altered. The same rule applies to all non–English proverbs, and sayings.

1.7 Combinations

Combinations consisting of various letters, numbers, and typographical marks are commonly used in English. They are often used to "encapsulate" certain content because of limited space. Instead of writing that a shop is always open, the combination "24/7" is used. This subsection will describe various kinds of combinations.

The following assumption describes the status of a combination in the sentence:

[4] A combination of various letters, marks, symbols, words, etc. can be considered as:

1) one word,

2) a sentencoid consisting of only one word.

1.7.1 Combinations of letters and numbers

Type of combination: a combination of letters, and numbers

Examples of sentences that contain the combination:

The section 2N is where the materials room is located.

The meteorite has crashed in the sector 3J.

Our machine is working almost 24h!

1.7.2 A combination of a number and the symbol “+”

Type of combination: a combination of a number, and the symbol “+”

Examples of sentences that contain the combination:

I have gathered 76+ points in just 17 minutes!

My collection consists of 100+ porcelain dolls.

It is at least 70+ years old.

1.7.3 A combination of a number and the symbol “%”

Type of combination: a combination of a number, and the symbol “%”

Examples of sentences that contain the combination:

Our incomes have increased by 23%!

Only today—17% discount!

1.7.4 A combination of the letter “x” and numbers

Type of combination: a combination of the letter “x,” and numbers

Examples of sentences that contain the combination:

They have selected the only available resolution: 800 x 600.

Our program requires a resolution higher than 640 x 400

My aunt invented a logic game that requires a 5 x 5 grid.

1.7.5 A combination of the currency symbol and a number

Type of combination: a combination of the currency symbol, and a number

Examples of sentences that contain the combination:

I have just won £500!

Will you borrow me $83?

I have lost $20 two hours ago.

1.8 Time expressing combinations

The term “time expressing combinations” includes combinations of words, and numbers that are used to write the time of an event. The following minisections will present the combinations.

1.8.1 Types of date expressing combinations

1) The first type consists of a word, and a number. The following dates were created according to the combination:

January 1, December 31.

2) The second type consists of three numbers with two forward slashes between them. The following dates were created according to the combination:

04/30/86, 09/29/93.

3) The third type consists of a word, two numbers, and a comma between the numbers. The following dates were created according to the combination:

November 1, 1880; June 8, 2005.

4) The fourth type consists of two words. The following dates were created according to the combination:

February first, March fourteenth.

5) The fifth type consists of the word “in,” and a number. The following dates were created according to the combination:

in 1952, in 1997, in 2004.

1.8.2 The combination for naming a decade of a century

The combination that allows to name the certain decade of the century consists of six words. The following examples were created according to the combination:

the twenties of the twentieth century,

the fifties of the twentieth century.

1.8.3 The combination for localizing an event in a decade of a century

The combination that allows to localize an event in a decade of a century consists of seven words. The following examples were created according to the combination:

in the twenties of the twentieth century,

in the fifties of the twentieth century.

1.8.4 The combination for localizing an event in a century

The combination that allows to localize an event in the certain century consists of four words. The following examples were created according to the combination:

in the fifteenth century,

in the twentieth century.

1.9 Is a sentencoid equivalent to a first level sentence?

Did the fact that a sentencoid may contain exactly the same quantity of words like a first level sentence means we can say that they are one, and the same thing?

To answer this question, the two following sentences will be analyzed:

(12) She was very angry, and very tired.

(13) This equation consists of more than three symbols.

The first sentence:

(12) She was very angry, and very tired.

The first four words can stand as a first level sentence. All what is needed is to put a period after the word “angry”:

She was very angry.

However, these four words may be, and in fact are, a part of the larger whole:

“She was very, angry and very tired.”

The sentence shows that in some cases a sentencoid may be equal to a first level sentence. Only a period is needed.

The second sentence:

(13) This equation consists of more than three symbols.

The first four words cannot stand as a first level sentence.

Conclusion: in some cases a sentencoid may be equal to a first level sentence.

1.10 Is a first level sentence always equal to a sentencoid?

Both a first level sentence, and a sentencoid consist of maximum four words. Does this fact mean that a first level sentence is always equal to a sentencoid? To answer the question, the following sentences will be analyzed:

I am ready. = a first level sentence that consists of one sentencoid. This first level sentence is equal to the sentencoid.

He is very sick. = a first level sentence that consists of one sentencoid. This first level sentence is equal to the sentencoid.

Yes, I can swim. = a first level sentence that consists of two sentencoids

Change your decision, please = a first level sentence that consists of to two sentencoids

Conclusion: because a first level sentence may consist of more than one sentencoids, the sentence is not always equal to one sentencoid.

1.11 Is a sentencoid equal to a clause?

Both a sentencoid, and a clause are basic elements of a sentence. Does this fact mean that a sentencoid is equal to a clause? The answer the question, various types of sentences will be analyzed.

Simple sentences

1. They really like oranges!

The first sentence consists solely of an independent clause. The clause contains four words.

2. Tom is driving his car now.

The second sentence consists solely of an independent clause. The clause contains six words.

Conclusion: a sentencoid may be equal to an independent clause. This depends on the quantity of words that the independent clause contains.

Compound sentences

1. Tom is a clerk, and Sarah works in a shop.

The first sentence consists of two independent clauses, and one coordinating conjunction between them.

2. Robert is a very professional physician, but Tom is the best teacher in our school.

The second sentence consists of two independent clauses, and one coordinating conjunction between them.

Conclusion: a sentencoid may be equal to an independent clause that belongs to the compound sentence. This depends on the quantity of words that a clause contains.

Complex sentences

1. When Tom is very hungry, he eats almost everything.

The first sentence consists of an independent clause, a subordinate clause, and the subordinator “when” that begins the sentence.

2. My parents tell me that I should spend more time in our library.

The second sentence consists of an independent clause, a subordinate clause, and the subordinator “that” between the clauses.

Conclusion: a sentencoid may be equal to a subordinate clause. The same applies to an independent clause. All depends on the quantity of words that a clause contains.

1.12 The status of a list of words

The following assumption describes the status of a list of words:

[5] A list of words is always classified as a sentencoid that consists of only one word.

A list often ends with the abbreviation “etc.” However, the abbreviation itself does not belong to the list. The role of the abbreviation is to signal that no more words will be listed.

Because a list of words is always considered as a sentencoid containing only one word, only one line is required to analyze such sentencoid. The line contains an equation. The left side of the equation contains a pair of parentheses containing the letter “L” inside them. The letter signals that the analyzed sentencoid contains only a list of words. The right side of the equation contains the word “OK.”

1.13 Proper nouns containing more than one word

The fact that sentences can be classified on the basis of how many words they contain may look simple; however, we must take into account that many proper nouns consist of more than just one word. These proper nouns include: official country names, names, and second names of persons, names of cities, etc.

The fact that a proper noun may consist of more than one word is problematic because it means that some proper nouns will consist of four, or even more words. Such proper noun will contain more words than a sentencoid normally contains. In order to solve this problem, the following assumptions have been made:

[6] A proper noun that consists of more than just one word cannot be divided into parts; words belonging to the proper noun cannot be present in adjacent sentencoids. All words that belong to a proper noun are present in the same sentencoid.

[7] No matter how many words a proper noun contains, it can be considered as:

1) one word,

2) a sentencoid consisting of only one word.

In the first case, a proper noun can be added to the rest of a sentencoid. This leads to the classification paradox. The second case makes possible to omit limitations of the available language material.

The presence of proper nouns may determinate character of text. The presence of historical leaders will point to a historical paper, while the presence of many terms related with human anatomy will point to a medical paper. However, it may happen that, for example, names of historical persons will appear in a text about astronomy, when the discoverer of a celestial body is mentioned, or interesting, but less known fact was presented in order to attract more readers.

The problem appears when there is no connection between the content of the available language material, and the text that we want to create. It will be very hard to create a text about the middle ages in reference to the text about information technology. To solve the problem related with content of the available language material, an additional assumption has been made:

[8] A sentencoid that consists of only one word is not searched in the available language material.

The analysis of a sentencoid that contains one word (see the analysis of the sentence 1 from the text #1) requires only one line of text. The line contains an equation. The left part of the equation contains a pair of parentheses containing the number “1” inside them. The number signals that the analyzed sentencoid consists of only one word. The right part of the equation contains the word “OK.”

The problem between the classification of sentences, and the existence of proper nouns containing more than one word is the source of the classification paradox. The term describes a situation in which a sentence containing a proper noun that consists of more than one word will be normally classified as a second level sentence, a third level sentence, or a fourth level sentence, but in fact, it is another type of sentence. The following sentences illustrate the paradox:

(14) He photographed the Sydney Opera House.

The quantity of words:	six words
The classification based on the theory of sentencoids:	four words → FLS

(15) He is a citizen of the United States of America.

The quantity of words:	ten words
The classification based on the theory of sentencoids:	seven words → SLS

(16) We saw the Eiffel Tower.

The quantity of words:	five words
The classification based on the theory of sentencoids:	four words → SLS

1.14 The status of a coordinator

Coordinating conjunctions (coordinators) join two, or more independent clauses in the compound sentence.

In order to describe the status of a coordinator, the following assumption has been made:

[9] A coordinator always stands as a separate sentencoid.

The following three sentences illustrate the assumption:

(17) Tom is very smart, but he is lazy.

17.3 The list of the created sentencoids

17.3.1 sentencoid 1:	Tom is very smart
17.3.2 sentencoid 2:	but
17.3.3 sentencoid 3:	he is lazy

(18) Sarah is a lawyer, and Robert is a doctor

18.3.1 sentencoid 1:	Sarah is a lawyer
18.3.2 sentencoid 2:	and
18.3.3 sentencoid 3:	Robert is a doctor

(19) I will select tea, or orange juice

19.3.1 sentencoid 1:	I will select tea
19.3.2 sentencoid 2:	or
19.3.3 sentencoid 3:	orange juice

Conclusion: a coordinator never creates sentencoids with other words.

1.15 The status of a subordinator

Subordinating conjunctions (subordinators) are present in complex sentences.

In order to describe the status of a subordinator, the following assumption has been made:

[10] No matter how many words a subordinator contains, it always creates a sentencoid that stands alone.

The three following sentences illustrate the assumption:

(20) She will do the analysis just as it should be done.

(21) The book was interesting as I expected.

(22) I am sure that he will win the race

(20) She will do the analysis just as it should be done.

20.3 The list of the created sentencoids

20.3.1 sentencoid 1: She will do the
20.3.2 sentencoid 2: do the analysis
20.3.3 sentencoid 3: just as
20.3.4 sentencoid 4: it should be done

(21) The book was interesting as I expected.

21.3 The list of the created sentencoids
21.3.1 sentencoid 1: The book was interesting
21.3.2 sentencoid 2: as
21.3.3 sentencoid 3: I expected

(22) I am sure that he will win the race.

22.3 The list of the created sentencoids

22.3.1 sentencoid 1: I am sure
22.3.2 sentencoid 2: that
22.3.3 sentencoid 3: he will win the
22.3.4 sentencoid 4: win the race

Conclusion: no matter how many words a subordinator contains, it always stands alone.

Conclusion: no matter how many words a subordinator contains, it always stands alone.

1.16 Words belonging to more than one word class

The fact that the same word may appear twice in the same sentence, and each time it belongs to a different word class is problematic. The analysis of the sentence 23 illustrates the problem:

(23) This mirror will mirror your image.

23.1 Type: SLS ($5 \leq W \leq 7$)

23.2 SLS → divide into sentencoids

23.3 The list of the created sentencoids

23.3.1 sentencoid 1: This mirror will mirror
23.3.2 sentencoid 2: mirror your image

RW: mirror

23.4 Analyses of the created sentencoids

23.4.1 sentencoid 1

23.4.1.1 “This mirror will mirror” = OK ($R \geq 3$)

23.4.2 sentencoid 2

23.4.2.1 “mirror your image” = OK ($R \geq 3$)

23.5 RW > 0 → Advanced language addition

23.6 The addition of the created sentencoids:

This mirror will mirror + mirror your image.

This mirror will mirror + ~~mirror~~ your image.

This mirror will mirror your image.

The sentence 23 is grammatically correct; however, the first sentencoid seems to be unique. The same word appears twice in the same sentencoid—first as a noun, and again as a verb.

The same problem will be illustrated by the analysis of the sentence 24:

(24) I will train in the second train.

24.1 Type: SLS ($5 \leq W \leq 7$)

24.2 SLS → divide into sentencoids

24.3 The list of the created sentencoids

24.3.1 sentencoid 1: I will train in
24.3.2 sentencoid 2: in the second train

RW: in

24.4 Analyses of the created sentencoids

24.4.1 sentencoid 1

24.4.1.1 “I will train in” = OK ($R \geq 3$)

24.4.2 sentencoid 2

24.4.2.1 “in the second train” = OK ($R \geq 3$)

24.5 $RW > 0 \rightarrow$ Advanced language addition

24.6 The addition of the created sentencoids:

I will train in + in the second train

I will train in + ~~in~~ the second train

I will train in the second train.

Again, the analyzed sentence is grammatically correct; however, the same word appears twice in the same sentence.

A possible solution to the problem is to use synonyms instead of writing the same word more than one time. In the sentences above, the word “train” can be replaced by the word “exercise,” and the word “mirror” can be replaced by the word “reflects.”

1.17 The status of a number

The presence of a number in a sentence can be described by the following assumption:

[11] No matter how a number is written (3; 6/2; 3,0; 3.0000000000; etc.) it can be considered as:

1) one word,

2) a sentencoid consisting of only one word.

The second case makes possible to omit limitations of the available language material.

The same assumption applies to a group of digits connected by hyphen(s).

1.18 The question of boundaries

The theory of sentencoids assumes the existence of two types of borders. The following assumptions describe their function:

[12] According to the theory of sentencoids, they are two types of boundaries: external, and internal boundaries. Both types of boundaries are introduced by punctuation marks; however, each punctuation mark introduces only one type of border.

[13] External boundaries are introduced by exclamation points, interrobangs, periods, and question marks.

[14] Internal boundaries are introduced by commas, en–dashes, parentheses, quotation marks, and semicolons.

1.18.1 Abbreviation vs. external borders

To solve the conflict between the presence of an abbreviation in a sentence, and the external border, the following assumption has been made:

[15] A period located next to an abbreviation does not introduce the external border.

The following sentences illustrate the exception:

Mr. Robert will be visiting this office at the evening.

This book was co–authored by Prof. Tom Green.

Well, Mrs. Sarah was late again.

1.19 Comma(s) in a sentence

The presence of a single comma makes possible to continue the sentence without ending it, and starting the next sentence. A comma can be used to separate a question from the rest of the sentence—the following sentences illustrate this:

He can write in English, can he?

You are ready, aren't you?

The presence of two commas in a sentence is necessary to:

1) set off an additional description,

2) set off words that interrupt the flow of a sentence,

3) set off an opinion, or a commentary,

4) set off the conjunctive adverb "however."

The part of a sentence that is set off by two commas can be syntactically integrated, or syntactically unintegrated with the rest of the sentence. In many cases, the two parts of a sentence that are separated by the part that is set off will create a sentencoid containing at least one word from each of these two parts—see the analysis of the sentence 11 from the text #1.

1.20 Semicolon(s) in a sentence

A semicolon is a punctuation mark that can be used between two independent clauses in order to avoid the comma splice error. According to the theory of sentencoids, a semicolon has one more function—it introduces an internal border.

Usually only one semicolon is used in a sentence; however, an exception exists. The exception takes place when we make a list of things that require to have semicolons between them. Consider the two following sentences:

I have learned about European Mediterranean countries, and also about their capitals: Spain, Madrid; France, Paris; Italy, Rome; and Greece, Athens.

My relatives live in many American cities: Boston, Massachusetts; Miami, Florida; Providence, Rhode Island; and San Diego, California.

1.21 Interrobang—the “forgotten” punctuation mark

An interrobang is a punctuation mark. Similarly to other members of its family, this punctuation mark also introduces an external border.

1.22 Quotation marks in a sentence

The following assumptions explain the role of quotation marks:

[16] The first pair of quotation marks begins an independent content within a sentence, and the second pair of quotation marks ends the content.

[17] The content inside a pair of quotation marks does not create sentencoids with the content outside both pairs of quotation marks.

The following question arises—why is the content inside both pairs of quotation marks analyzed separately?

The content is analyzed separately because it will not be easy to find a sentencoid containing words from both before, and after a single pair of quotation marks. The content between quotation marks may be a direct speech, or another content that cannot be altered.

1.23 En–dash(es) in a sentence

A single en en–dash allows to divide a sentence into sentencoids. A pair of en–dashes is used to surround a non–essential clause. In many cases, the two parts of a sentence that are separated by a non–essential clause will create a sentencoid containing at least one word from each of these two parts.

1.24 Hyphen(s) in a compound word

The following assumption describes the status of a compound word:

[18] A compound word is always considered as one word.

The presence of the compound word(s) in a sentence does not influence the process of creating sentencoids. The compound words like re–create, re–installation, merry–go–round, Tic–tac–toe, etc. will not be divided into parts, because they contain one, or more hyphens.

1.25 The presence of parentheses

The following assumption describes the status of the content that is inside a pair of parentheses.

[19] The content inside a pair of parentheses is analyzed independently from the content outside parentheses—i.e. there is no sentencoid that contains words from both the content inside, and outside parentheses.

The assumption is based on the fact that the presence of a pair of parentheses is not compulsory to understand the meaning of a sentence. Parentheses only provide additional information. They can be removed, and the sentence will be still grammatically correct, and understandable. The five following sentences illustrate this:

(25a) Tom is drinking a glass of tea (with three ice cubes) at the moment.

(26a) The process of learning physics (and chemistry) requires a lot of patience.

(27a) In my opinion, he was too lazy (and too nervous) to accomplish his personal goals.

(28a) All cars in this city (except the car of my old friend) are very expensive.

(29a) All students (except me, and two good friends of mine) have agreed with the lecturer.

The same sentences without parenthetical material:

(25b) Tom is drinking a glass of tea at the moment.

(26b) The process of learning physics requires a lot of patience.

(27b) In my opinion, he was too lazy to accomplish his personal goals.

(28b) All cars in this city are very expensive.

(29b) All students have agreed with the lecturer.

It should also be taken into account that the content inside a pair of parentheses is not always syntactically integrated with the content outside parentheses; the pair may, for example, contain an interjection.

1.26 The question of capitalization

This subsection uses the following four sentences (30-33) to explain the question of the capitalization from the theory's point of view:

(30) He is teaching English in the primary school.
(31) He is teaching english in the primary school.
(32) Well, I am writing my essay now.
(33) Well, i am writing my essay now.

According to the grammar of the English language the sentence 31, and the sentence 33 are incorrect. Both sentences lack of the proper capitalization. From the theory's point of view, however, all four sentences are correct.

The question of capitalization illustrates how much depends on the available language material. To avoid any grammatical errors, the available language material must be written in accordance with the rules that govern the grammar of the English language.

1.26.1 Writing all in capital letters

Sometimes a word is written all in capital letters. This is done in order to underline the significance of the word—see the following sentences:

Can you tell us WHEN did your friend left the meeting?

WHERE did she go?

I will NEVER trust you again.

The presence of such words does not influence the process of breaking and re-creating an original sentence.

1.26.2 Capitalization vs. sentencoids

We do not capitalize the first letter of the first word of a sentencoid. However, in some situations the first letter of the word that starts a sentencoid must be capitalized. These situations occur when:

1) the word that starts a sentencoid is a proper noun,

2) the word that starts a sentencoid is also the first word of the original sentence,

3) the word that begins a sentencoid is a fragment of direct speech.

1.27 The process of breaking and re–creating an original sentence

The process of breaking and re–creating an original sentence consists of six phases during which we divide sentences into sentencoids. We analyze the created sentencoids, and later add them in the same order they were created to form the original sentence.

The first four phases stand as the “process of creating sentencoids.” The last two phases stand as language addition. The following minisection will present the first four phases.

1.27.1 The process of creating sentencoids

The following assumption has been made:

[20] Each type of sentence is divided into sentencoids. The only exception takes place when a sentence consists of maximum four words, and does not contain any introducer of the internal border.

The classification of a sentence (the first phase) begins the process. The classification is based on the quantity of words that the analyzed sentence contains. The classification requires only one line of text. The left side of the line contains the word “Type,” and a colon. The right side of the line contains an acronym that signals which type of sentence is analyzed.

The second phase informs what will happen with the classified sentence:

The third phase of the process contains two vertical columns:

The column on the left contains names of sentencoids.

The column on the right contains content of each sentencoid.

At the end of the third phase, repeated words (RW) are listed. The term “repeated words” means words that appear more than one time; at the end of the previous sentencoid, and again at the beginning of the next sentencoid as a result of the process.

RW can take positive values, or the neutral value. The first possibility means that at least one word appeared more than one time as a result of the process of creating sentencoids. The sentence number 18 from the text #1 illustrates the situation where only word is repeated. In most cases, more than one word is repeated. It may happen that all repeated words will create a set of words that will look almost like a sentencoid—see the analysis of the sentence number 15 from the text #1.

The neutral value of RW, which is symbolized by the number zero, means that no word was repeated as a result of the process of creating sentencoids. This situation is illustrated by the sentence number 2 from the text #1.

The value of RW determinates what type of language addition will be used in the last two phases of the process of breaking and re–creating an original sentence.

In the fourth phase, most of the analyzed sentencoids will be searched in the available language material. The search can bring the positive result, or the negative result. In the second case, it may happen that the search will bring the most extreme variant of the negative result—the zero result.

The positive result of the search (or the “positive result” in short) informs that the searched sentencoid appeared at least three times in the available language material. The result is written in the form of the following equation:

“abc” = OK (R ≥ 3)

The left side of the equation contains the three following letters “abc.” The letters symbolize the content of the searched sentencoid.

The right side of the equation contains the word “OK” together with a pair of parentheses containing the information that the searched sentencoid must appear at least three times in order to gain the positive result.

The letter “R” that informs how many times the searched sentencoid was found.

The negative result of the search (or the “negative result” in short) informs that the searched sentencoid was found less than three times in the language material. The result is written in the form of the following equation:

“abc” = INCORRECT (R < 3)

The left side of the equation contains the three following letters “abc.” The letters symbolize the content of the searched sentencoid.
The right side of the equation contains the word “INCORRECT” that is written in capital letters. A pair of parentheses contains the condition that must be fulfilled by the searched sentencoid.

The letter “R” that informs how many times the searched sentencoid appeared.

It may happen that the value of “R” will be equal to the number “zero.” Such result is called the “zero result,” and has the form of the following equation:

“abc” = ZERO (R = 0)

The left side of the equation contains the three following letters “abc.” The letters symbolize the content of the searched sentencoid.

The right side of the equation contains the word “ZERO.” A pair of parentheses contains an equation; the left side contains the letter “R,” while the right side contains the number “0.”

The analysis of a sentencoid that consists of only one word requires one line. The line contains an equation. The left side of the equation contains a pair of parentheses, and the number “1” inside them. The right side of the equation contains the word “OK.”

1.28 Repeated words

An original sentence is divided into sentencoids in order to check its correctness. Each created sentencoid consists of maximum four words. The first, and the last word are crucial. The last word (or the two last words) of the previous sentencoid is (are) repeated at the beginning of the next sentencoid. Repeated word (words) begins (begin) the new sentencoid.

The following question arises:

Why it so important to repeat the last word(s) of the previous sentencoid at the beginning of the next sentencoid?

This repetition is caused by the fact that it will be hard, or even impossible to find a sentence that is classified as a third level sentence, or a fourth level sentence. It is much easier to find a part of the original sentence instead of the whole sentence. By progressively checking sentencoids that have been created, we are checking the original sentence. The purpose of repeating the last word(s) of the previous sentencoid is to create continuity between neighboring sentencoids.

If a sentencoid ends with an article, or the word "of," the next sentencoid begins with the last two words of the previous sentencoid—see the analysis of the sentence number 10 from the text #1. However, there are situations that do not require to repeat the last word(s) of the previous sentencoid at the beginning of the next sentencoid. These situations take place when:

1) the word that starts an original sentence, starts the first sentencoid too,

2) the new sentencoid starts after a single en–dash,

3) the new sentencoid starts after a single comma,

4) the new sentencoid starts after a semicolon.

1.29 What is responsible for the negative result?

The negative result does not always mean that the analyzed sentencoid is incorrect. The following reasons explain why the negative result may be obtained:

1) the presence of an archaism in a sentencoid,

2) the presence of a (too) specialized terminology in a sentencoid,

3) the presence of a less-known proper noun in a sentencoid.

A good example that illustrates the last reason is the case of an ancient mythological hero, Hercules. He originates from Greek mythology; however, the name "Hercules" is a Roman name, not a Greek name. The Greek name is "Heracles." The Greek name may appear in a sentencoid, while the Roman name may appear in the available language material. These different names may cause the negative result, despite the fact that both names are correct, and refer to the same mythological character.

1.30 Language addition

The last two phases of the process of breaking and re–creating an original sentence stand as language addition. In most cases, each sentencoid begins with the same word that ends the previous sentencoid. The problem appears when we want to add all sentencoids in

order to form the original sentence. To avoid chaos, a new type of addition, which I have named “language addition,” will be presented.

1.30.1 How does language addition work?

The word “language” signals two features. The first feature points to the nature of sentencoids. In mathematics numbers can be added to each other; regardless of their order of appearance. The case of sentencoids is different. They cannot be added to each other without respecting their order of creation. The solution is to add all sentencoids in the same order they were created.

The presence of the second feature relies on the selected type of language addition. Before the second feature will be explained, two types of language addition will be presented. Language addition may be simple, or advanced. The diversity is caused by the fact that RW may take two values:

1) the positive value of RW (RW > 0) means that the advanced language addition will be used. This value signals that at least one word was repeated as a result of the process of creating sentencoids.

2) the neutral value of RW (RW = 0) means that simple language addition will be used. This value signals that no word was repeated during the process of creating sentencoids. In this sense, this type of language addition is simple—there is no need to remove repeated words, because they do not exist.

The second feature of language addition can be explained now:

The second feature of language addition points to the structure of sentencoids. The positive value of RW means that repeated words are present. Their presence means that the original sentence will be unreadable, when all sentencoids will be added. The solution is to remove the second (or the third appearance) of the same word.

1.30.2 Phases of language addition

The following assumption describes the role of RW:

[21] The value of RW determinates the number of lines that language addition will contain. The neutral value of RW signals simple language addition; the positive value of RW signals advanced language addition.

Simple language addition consists of two lines of text:

1) the first line in which sentencoids appear in the same order they were created, and checked during the process of creating sentencoids. However, this time sentencoids are listed horizontally, not vertically. The symbols of addition are located not only between sentencoids, but also before, and after punctuation marks.

2) the second line that contains the original sentence.

Advanced language addition consists of the three following lines of text:

1) the first line in which sentencoids appear in the same order they were created, and checked during the process of creating sentencoids. However, this time sentencoids are listed horizontally, not vertically. The symbols of addition are located not only between sentencoids, but also before, and after punctuation marks.

2) the second line in which the second (or the third) appearance of the same word is marked by a horizontal line. The marked words together with the symbol(s) of addition will be removed,

3) the third line that contains the original sentence.

1.31 Words that interrupt the flow of a sentence

The following assumption has been made to describe the status of words that interrupt the flow of a sentence:

[22] Words that interrupt the flow of the sentence are always analyzed separately.

Firstly, the part that does not interrupt the flow is divided into sentencoids, and RW for the part are defined. Next, the part that is interrupting the flow is divided into sentencoids. If the part is equal to one sentencoid, there is no sense in defining its RW.

There is, however, a difference between a text that is between two commas/en–dashes, and a text inside a pair of parentheses. The difference reflects the fact that the text inside a pair of parentheses can be always removed from the sentence. That is why the words inside a pair of parentheses are not taken into account when classifying the sentence that contains them. In contrast, words that interrupt the flow of a sentence, and are surrounded by two commas/en–dashes are taken in account when classifying the sentence that contains them.

Language addition of a sentence containing words that interrupt the flow has additional lines. Their function is exactly the same as the function of the “standard” lines. The only difference is that they describe language addition for the part that is interrupting the flow of a sentence.

Each sentencoid created from the part that interrupts the flow of a sentence is numbered. The number itself consists of two numbers, and the symbol of addition between them. The first number is, however, identical to the number of the sentencoid created from non-interrupting content that precedes the first sentencoid created from the interrupting part. The second number signalizes which place a sentencoid occupies in the order of sentencoids created from the part that interrupts the flow of a sentence. The symbol of addition joins these two numbers together.

The case of an en–dash is similar to the case of other punctuation marks. The difference lies in the fact that an en–dash is placed inside a pair of curly braces. The pair itself is surrounded by symbols of addition.

1.32 Language morphing

The word “morphing” means:

to change something,

to change shapes of something,

to change into something else.

The term “language morphing” is used to describe the change of one word into another one. The following assumption describes how this type of morphing works:

[23] A word can be morphed into another word when both of them are members of the same word class.

Additional conditions for nouns, verbs, attributive adjectives, and WH–adverbs are described in the second section.

Language morphing is used when the negative result has been obtained—see the analysis of the sentence 16 from the text #1. The negative result means that we must add four more lines of text. The first added line contains an equation. Each side of the equation contains only one word. The left side contains brackets containing the word that will be morphed into another word. The right side of the equation contains the word that will replace the word inside the brackets.

The second added line contains an equation. The left side of the equation contains the analyzed sentencoid. However, the “new” word has replaced the “old” word. The right side of the equation contains the result of the search. The negative or the positive result has exactly the same form as it was presented earlier.

If the positive result was obtained:

the third added line contains an equation. Each side of the equation contains only one word; however, this time brackets are not present. The left side of the equation contains the “old” word, and the right side contains the “new” world. This equation signals that:

1) both words are equal in the sense that they are members of the same word class,

2) the new word can replace the old word, and the sentence will be still grammatically correct.

The last fourth line contains the original sentencoid together with brackets containing the word “OK” next to it. We can assume that the sentencoid is correct.

The difference between the word “OK” in brackets, and the same word without brackets helps to understand the role that language morphing plays. The “[OK]” signals that the positive result was obtained only because language morphing was used.

If the negative result was obtained, it is necessary to follow the grammar of the English language in order to check correctness.

The use of language morphing has its limitations. These limitations are presented in the second section of this paper.

2 The selected aspects of language morphing

The section illustrates how to morph countable nouns into uncountable nouns, and vice–versa. The section also describes:

the question of morphing verbs,

the question of morphing attributive adjectives,

the question of morphing wh–adverbs.

The section uses simplified notation of language morphing. The third section of the paper uses full notation of language morphing.

2.1 Morphing within the class of nouns

The following subsection presents the rules that must be taken into account when morphing countable nouns into uncountable nouns, and vice–versa. Each case is illustrated with one example.

2.1.1 The first case

The first case illustrates how to morph an uncountable noun into the plural form of a countable noun. The sentence 34a contains the uncountable “rice,” and the partitive noun “ton”:

(34a) I have bought a ton of rice.

The following question arises:

Can an uncountable noun simply be morphed into the plural form of a countable noun without changing anything else in the sentence?

rice → tomatoes

(34b) I have bought a ton of tomatoes.

The sentence 34b is grammatically correct.

Conclusion: it is possible to morph an uncountable noun into the plural form of a countable noun without changing anything else in the sentence.

2.1.2 The second case

The second case illustrates how to morph an uncountable noun into the plural form of a countable noun when we also want to add a cardinal number before the partitive noun.

The sentence 35a contains the uncountable noun “rice”:

(35a) I have bought a ton of rice.

The uncountable noun “rice” will be morphed into the plural form of the countable noun “tomato.” The cardinal number “three” will be added before the partitive noun:

rice → tomatoes

(35b) I have bought a three ton of tomatoes.

The sentence 35b can be understood; however, it is incorrect. In order to correct it, we must follow the following rule:

To morph an uncountable noun into the plural form of a countable noun, and to add a cardinal number, we must remove an indefinite article, and change the form of a partitive noun from the singular to the plural form. The value of the cardinal number must be at least equal to the number “2.”

The rule will be applied to the analyzed sentence:

I have bought ~~a~~ three ton(+s) of tomatoes.

(35c) I have bought three tons of tomatoes.

The sentence 35c is now grammatically correct.

The last sentence of the previous rule underlines the importance of the value of a cardinal number. If the value is equal to one, there is no need to add the cardinal number “one” before the partitive expression. What is more, we do not change the form of the partitive noun. The partitive noun stands in its singular form.

Compare the following sentences:

(36a) I have bought a ton of tomatoes. [CORRECT]

(36b) I have bought a tons of tomatoes. [INCORRECT]

(37a) I have drunk a bottle of tomato juice. [CORRECT]

(37b) I have drunk a bottles of tomato juice. [INCORRECT]

2.1.3 The third case

The third case illustrates how to morph the singular form of a countable noun into an uncountable noun.

The sentence 38a contains the singular form of the noun “computer”:

(38a) I have bought a computer.

The singular form of the countable noun “computer” will be morph into the uncountable noun “tea.” The question is:

Can we simply change the singular form of a countable noun into an uncountable noun without changing anything else?

a computer → tea

(38b) I have bought a tea.

The sentence 38b can be understood; however, it is incorrect. This is because we are talking about the quantity of an uncountable noun. The uncountable noun “tea” must be preceded with the pretitive expression.

We must apply the following rule:

to morph the singular form of a countable noun into an uncountable noun we must add a partitive expression. The expression will replace an indefinite article that is located before the singular form of the countable noun.

The partitive expression “a cup of” will be added before the uncountable noun:

I have bought + a cup of + tea.

(38c) I have bought a cup of tea.

The sentence 38c is grammatically correct.

2.1.4 The fourth case

The fourth case illustrates how to morph the plural form of a countable noun into an uncountable noun.

The sentence 39a contains the plural form of the countable noun “shoe”:

(39a) I have bought shoes.

The plural form of the countable noun “shoe” will be morph into the uncountable noun “tea.” The question is:

Can we simply change the plural form of a countable noun into the uncountable noun without changing anything else?

shoes → coffee

(39b) I have bought coffee.

Because we are talking about the quantity of an uncountable noun, the uncountable noun "coffee" must be preceded with the pretitive expression:

I have bought + a cup of + coffee.

(39c) I have bought a cup of coffee.

The sentence 39c is grammatically correct.

We must follow the following rule:

To morph the plural form of a countable noun into an uncountable noun we must add a partitive expression before the uncountable noun.

(39c) I have bought a cup of coffee.

The sentence 39c is grammatically correct.

2.2 Morphing the main verb

Each tense, except the Present Simple Tense and the Past Simple Tense, requires the presence of more than one verb. We can divide them into the main verb, and at least one auxiliary verb. The new verb that will replace the previous verb as the main verb must have exactly the same form as its predecessor. This condition cannot be violated. Three tenses have been chosen to illustrate this rule. In each case, the main verb will be morphed four times:

The present perfect tense

(40a) I have written the book.

written → seen

(40b) I have seen the book. [CORRECT]

written → see

(40c) I have see the book. [INCORRECT]

written → saw

(40d) I have saw the book. [INCORRECT]

written → seeing

(40e) I have seeing the book. [INCORRECT]

The past perfect continuous tense

(41a) These students had been swimming for two hours.

swimming → eating

(41b) These students had been eating for two hours. [CORRECT]

swimming → eat

(41c) These students had been eat for two hours. [INCORRECT]

swimming → ate

(41d) These students had been ate for two hours. [INCORRECT]

swimming → eaten

(41e) These students had been eaten for two hours. [INCORRECT]

The future continuous tense

(42a) I will be playing in four days.

playing → singing

(42b) I will be singing in four days. [CORRECT]

playing → sing

(42c) I will be sing in four days. [INCORRECT]

playing → sang

(42d) I will be sang in four days. [INCORRECT]

playing → sung

(42e) I will be sung in four days. [INCORRECT]

2.2.1 The meaning of subject–verb agreement

The relation between the subject and the main verb in a sentence helps to distinguish the singular form of the subject from its plural form. This relation is called subject–verb agreement. The two following sentences perfectly illustrate the agreement:

The following list lists the most important issues.
The following lists list the most important issues.

If the previous main verb indicates the singular form of the subject, the new main verb cannot indicate the plural form of the subject. The following sentence illustrates this relation:

The sentence 43a contains the main verb “repairs”:

(43a) Tom repairs buildings on every Thursday.

The main verb will be morphed:

repairs → paint

(43b) Tom paint buildings on every Thursday.

The sentence 43b is incorrect. The new main verb suggests the presence of the plural form of the subject. We make a correction, and obtain a grammatically correct sentence:

(43c) Tom paints buildings on every Thursday.

The sentence 43c is grammatically correct.

2.2.2 Subject–verb agreement vs. all grammatical tenses

The following sentences illustrate that subject–verb agreement is present only in the Present Simple Tense. Each tense contains the singular, and plural form of the subject.

The Present Simple Tense

(44a) My rabbit eats carrots on Saturdays.
(44b) My rabbits eat carrots on Saturdays.

The Present Continuous Tense

(44c) My rabbit is eating carrots right now.
(44d) My rabbits are eating carrots right now.

The Present Perfect Tense

(44e) My rabbit has already eaten carrots.
(44f) My rabbits have already eaten carrots.
The Present Perfect Continuous Tense

(44g) My rabbit has been eating carrots since yesterday.
(44h) My rabbits have been eating carrots since yesterday.

The Past Simple Tense

(44i) My rabbit ate carrots one hour ago.
(44j) My rabbits ate carrots one hour ago.

The Past Continuous Tense

(44k) My rabbit was eating carrots when she called me.
(44l) My rabbits were eating carrots when she called me.

The Past Perfect Tense

(44m) My rabbit had eaten carrots for three hours.
(44n) My rabbits had eaten carrots for three hours.

The Past Perfect Continuous Tense

(44o) My rabbit had been eating carrots since 2007.
(44p) My rabbits had been eating carrots since 2007.

The Future Simple Tense

(44q) My rabbit will always eat carrots.
(44r) My rabbits will always eat carrots.

The Future Continuous Tense

(44s) My rabbit will be eating carrots in two days.
(44t) My rabbits will be eating carrots in two days.

The Future Perfect Tense

(44u) My rabbit will have eaten carrots by night.
(44v) My rabbits will have eaten carrots by night.

The Future Perfect Continuous Tense

(44w) By the end of this day my rabbit will have been eating carrots for four hours.
(44x) By the end of this day my rabbits will have been eating carrots for four hours.

Conclusion: subject–verb agreement is present only in the Present Simple Tense.

2.2.3 The question of auxiliary verbs

Some auxiliary verbs indicate the presence of the singular form of the subject. However, this insight applies only to the four following tenses:

1) The Present Continuous Tense,

2) The Present Perfect Tense,

3) The Present Perfect Continuous Tense,

4) The Past Continuous Tense.

The table #1 illustrates these auxiliary verbs together with their corresponding grammar tenses.

Grammatical Tense	Auxiliary Verb
The Present Continuous Tense	is
The Present Perfect Tense	has
The Present Perfect Continuous Tense	has
The Past Continuous Tense	was

The table #1: Auxiliary verbs with their corresponding grammatically tenses.

2.2.4 The short passive as a result of language morphing

The words that can replace the previous word, and meet other additional conditions create a specific set of words. Such set contains the so–called "empty set." The empty set is, as its name suggests, empty. The existence of the empty set allows to morph a word into nothingness—i.e. a word can be deleted. This fact makes possible to change the passive voice (45a) into the short passive (45b):

(45a) Car was repaired by Tom.

Because the short passive requires absence of the performer of an activity, the performer "Tom" will be removed:

Tom → Ø

"Car was repaired by."
Because the performer has been removed, the same must happen with the preposition "by":

by → Ø

"Car was repaired"

A period will be placed after the word "repaired":

(45b) "Car was repaired."

The sentence 45b illustrates the short passive.

2.3 Morphing adjectives

The subsection presents the rules that must be taken into account when morphing adjectives, and also describes the limitations of this kind of morphing. The subsection will focus only on attributive adjectives—i.e. adjectives that appear before nouns.

The gradability of adjectives vs. the morphing of adjectives

The fact that adjectives are gradable limits the number of results that can be achieved by using language morphing within members of this word class. In order to explore this question, language morphing of various forms of adjectives will be performed.

Gradable vs non–gradable adjectives

The next issue is related with the fact that some adjectives are non–gradable. Such adjective does not have the comparative, and superlative form. It cannot be used with most of gradable adverbs.

2.3.1 Language morphing of adjectives in their positive form

Language morphing of adjectives that share the positive form will be described.
The sentence 46a contains the adjective “smart.” The adjective will be morphed into its antonym:

(46a) That was a smart decision.

smart → stupid

(46b) That was a stupid decision.

The sentence 46b is grammatically correct, and makes sense.

The sentence 47a contains the adjective “rich” that will be morphed into its antonym:

(47a) My uncle is a rich man.

rich → poor

(47b) My uncle is a poor man.

The sentence 47b is grammatically correct, and makes sense.

The sentence 48a contains the adjective “dirty” that will be changed into its antonym:

(48a) What a dirty room this is!

dirty → clean

(48b) What a clean room this is!

The sentence 48b is grammatically correct, and makes sense.

Conclusion: it is possible to morph an adjective into its antonym that shares the positive form, and to create a sentence that makes sense.

The morphing of adjectives that share the positive form, but are not antonyms will be performed now. The sentence 49a contains the adjective “green.” The adjective will be changed into other adjectives:

(49a) I have bought a green car as the present.

green → red

(49b) I have bought a red car as the present.
green → blue

(49c) I have bought a blue car as the present.

green → yellow

(49d) I have bought a yellow car as the present.

green → fast

(49e) I have bought a fast car as the present.

green → proud

(49f) I have bought a proud car as the present.

green → late

(49g) I have bought a late car as the present.

Language morphing may create sentences that sound quite bizarre (49f, 49g). To explore this question, the morphing of stative, and dynamic adjectives will be performed.

The sentence 50a contains the dynamic adjective “shy.” The dynamic adjective will be morphed into several other dynamic adjectives:

(50a) What a shy person he is!

shy → rude

(50b) What a rude person he is!

shy → friendly

(50c) What a friendly person he is!

shy → patient

(50d) What a patient person he is!

shy → calm

(50e) What a calm person he is!

shy → suspicious

(50f) What a suspicious person he is!

shy → foolish

(50g) What a foolish person he is!

shy → patient

(50h) What a patient person he is!

shy → strange

(50i) What a strange person he is!

shy → cruel

(50j) What a cruel person he is!

Conclusion: the morphing of adjectives that share the same positive form, but are not antonyms, may give sentences that are both grammatical correct, and have sense.

The sentence 51a contains the dynamic adjective “shy.” The dynamic adjective will be morphed into several other stative adjectives:

(51a) What a shy person he is!

shy → tall

(51b) What a tall person he is!

shy → small

(51c) What a small person he is!

shy → young

(51d) What a young person he is!

Conclusion: language morphing within the adjectives that share the positive form, but are attributive, and dynamic adjectives, is possible, and may give correct sentences.

2.3.2 Language morphing of adjectives that share comparative form

Language morphing of adjectives that share the comparative form will be described.

The sentence 52a contains the comparative form of the adjective “fast.” The form will be changed into its antonym:

(52a) My friend is faster than the previous winner.

faster → slower

(52b) My friend is slower than the previous winner.

The sentence 52b is grammatically correct, and makes sense.

The sentence 53a contains the comparative form of the adjective “smart.” The form will be changed into its antonym:

(53a) He is smarter than your brother.

smarter → dumber

(53b) He is dumber than your brother.

The sentence 53b is grammatically correct, and makes sense.

Conclusion: it is possible to morph an adjective into its antonym that shares the comparative form, and to create sentences that make sense.

The morphing of adjectives that share the comparative form, but are not antonyms will be performed now. The sentence 54a contains the comparative form of the adjective “smart.” The form will be morphed into other adjectives that share the same form:

(54a) He is smarter than your brother.

smarter → richer

(54b) He is richer than your brother.

smarter → faster

(54c) He is faster than your brother.

smarter → weaker

(54d) He is weaker than your brother.

Conclusion: language morphing of adjectives that share the comparative form, but are not antonyms may give sentences that are both grammatical correct, and have sense. However,

carefulness is recommended when performing language morphing of adjectives that are not antonyms.

2.3.3 The language morphing of adjectives that share the superlative form

The sentence 55a contains the superlative form of the adjective “big.” The form will be morphed into its antonym:

(55a) This is the biggest cat I have ever seen!

the biggest → the smallest

(55b) This is the smallest cat I have ever seen!

The next sentence (56a) contains the superlative form of the adjective “fast.” The form will be changed into its antonym:

(56a) This is the fastest car in the world.
the fastest → the slowest

(56b) This is the slowest car in the world.

Both the sentence 55b, and the sentence 56b are grammatically correct, and make sense.
An adjective in its superlative form can be morphed into its antonym that shares the same form. However, because the superlative form of some adjectives requires the word “most,” this kind of morphing may change the quantity of words in the newly created sentence. See the following example:

The sentence 57a contains the superlative form of the adjective “simple.” The form will be morphed into its antonym:

(57a) You have chosen the simplest way of doing this.

the simples → the most complicated

(57b) You have chosen the most complicated way of doing this.

The sentence 58a contains the superlative form of the adjective “good.” The form will be morphed into its antonym:

(58a) He has made the smartest decision.

the smartest → the most foolish

(58b) He has made the most foolish decision.

Conclusion: language morphing of adjectives that share the superlative form may change quantity of words.

2.3.4 Morphing (not only) into antonyms

The meaning of a sentence limits the number of words that can be used to replace the "original" word in the sentence. The following example illustrates this problem:

The sentence 59a contains the superlative form of the adjective "good" The superlative form will be morphed into the superlative form of the adjective "lazy":

(59a) This is the best pencil I have ever used!

the best → the laziest

(59b) This is the laziest pencil I have ever used!

The sentence 59b makes no sense.

Conclusion: to avoid situations in which the newly created sentence makes no sense, it is better to morph the superlative form of an adjective into its antonym that shares the same form:

(59a) This is the best pencil I have ever used!

the best → the worst

(59c) This is the worst pencil I have ever used!

The sentence 59c makes sense.

2.4 Morphing the selected wh–adverbs

The subsection explains how to morph the three following wh–adverbs: when, where, and why. The subsection also describes how a wh–question behaves when a wh–adverb is removed from it.

2.4.1 Morphing the three wh-adverbs: when, where, and why

The purpose is to examine how the three wh–adverbs can be morphed into each other. The sentence 60a contains the wh–adverb "where." The wh–adverb will be morph into the other two wh–adverbs:

(60a) Where you want to meet him?

Where → When

(60b) When you want to meet him?

Where → Why

(60c) Why you want to meet him?

Both the sentence 60b, and the sentence 60c are grammatically correct, and make sense.

The sentence 61a contains the wh–adverb “when.” The wh–adverb will be morph into the other two wh–adverbs:

(61a) When did this event happen?

When → Where

(61b) Where did this event happen?

When → Why

(61c) Why did this event happen?

Both, the sentence 61b, and the sentence 61c are grammatically correct, and make sense.

The sentence 62a contains the wh–adverb “why.” The wh–adverb will be morph into the other two wh–adverbs:

(62a) Why did you buy the notebook?

Why → Where

(62b) Where did you buy the notebook?

Why → When

(62c) When did you buy the notebook?

Both, the sentence 62b, and the sentence 62c are grammatically correct, and make sense.

Conclusion: they were no problems with morphing within the three selected wh–adverbs. Each newly created sentence is grammatically correct.

2.4.2 Removing wh–adverbs

Wh–adverbs are present in interrogative sentences, known as wh–questions. If a wh–adverb is removed, the wh–question becomes an “ordinary” interrogative sentence. The following sentences illustrate this.

The sentence 63 contains the wh–adverb “where.” The wh–adverb will be morphed into nothingness:

(63) Where you want to meet him?

Where → Ø

you want to meet him?

The letter “y” will be capitalized:

(64) You want to meet him?

The new interrogative sentence is grammatically correct, and makes sense.

The sentence 65 contains the wh–adverb “when.” The wh–adverb will be morphed into nothingness:

(65) When did this event happen?

When → Ø

did this event happen?

The letter “d” will be capitalized:

(66) Did this event happen?

The new interrogative sentence is grammatically correct, and makes sense.

The sentence 67 contains the wh–adverb “when.” The wh–adverb will be morphed into nothingness:

(67) Why are you digging?

Why → Ø

are you digging?

The letter “a” will be capitalized:

(68) Are you digging?

The new interrogative sentence is grammatically correct, and makes sense.

Conclusion: despite the removal of wh–adverbs, the sentences are still grammatically correct. The sentences are still classified as interrogative sentences.

3 The theory in practice

The section contains one text; the text about the structure of an atom. The text will be analyzed in accordance with the theory of sentencoids, and the grammar of the English language.

This section is based on the assumption that the available language material contained enough content to give positive results.

3.1 The text #1

The Text #1

An atom consists of three kinds of subatomic particles—electrons, neutrons, and protons. Protons, and electrons are electrically charged. Protons have positive electric charges, and electrons have negative electric charges. The nucleus is formed by protons, and neutrons. The consequence of this fact is that the nucleus is positively charged. Because the number of protons, and electrons is equal, the atom is electrically neutral.

The subatomic particles differ in mass. The mass of each electron is a fraction when we will compare it to the unified atomic mass unit. In contrast, the mass of each proton (and also the mass of each neutron) is almost equal to the unified atomic mass unit. These facts lead to the conclusion that the nucleus holds almost all of the mass of every atom.

Probably the best way to visualize the atom, its inner structure, is to compare it to the Solar System. According to the planetary model, electrons act like planets. They orbit the nucleus of the atom. Valence electrons, the most distant electrons, create chemical bonds between atoms. The nucleus plays the role of the Sun. The planetary model is easy to understand; however, it confused physicists. Why negatively charged electrons do not fall into the positively charged nucleus? The subatomic world required a correction.

New views on the atom have been born. The quantum theory states that it is impossible to give the exact location of an electron. Its location can be estimated only. Such calculation is based on probability.

The theory was established in the twentieth century. It describes, and rules the subatomic world.

3.2 The analysis of the first text

1. Sentence: **An atom consists of three kinds of subatomic particles—electrons, neutrons, and protons.**

1.1 Type: TLS ($8 \leq W \leq 14$)

1.2 TLS → divide into sentencoids

1.3 The list of the created sentencoids

1.3.1 sentencoid 1: An atom consists of
1.3.2 sentencoid 2: consists of three kinds
1.3.3 sentencoid 3: kinds of subatomic particles
1.3.4 sentencoid 4: electrons, neutrons,
1.3.5 sentencoid 5: and
1.3.6 sentencoid 6: protons

RW: consists, of, kinds

1.4 Analyses of the created sentencoids

1.4.1 sentencoid 1

1.4.1.1 “An atom consists of” = OK ($R \geq 3$)

1.4.2 sentencoid 2

1.4.2.1 “consists of three kinds” = OK ($R \geq 3$)

1.4.3 sentencoid 3

1.4.3.1 “kinds of subatomic particles” = OK ($R \geq 3$)

1.4.4 sentencoid 4

1.4.4.1 (L) = OK

1.4.5 sentencoid 5

1.4.5.1 (1) = OK

1.4.6 sentencoid 6

1.4.6.1 (1) = OK

1.5 RW > 0 → Advanced language addition

1.6 The addition of the created sentencoids:

An atom consists of + consists of three kinds + kinds of subatomic particles + {—} + electrons, neutrons + , + and + protons.

An atom consists of + ~~consists of~~ three kinds + ~~kinds~~ of subatomic particles + {—} + electrons, neutrons + , + and + protons.

An atom consists of three kinds of subatomic particles—electrons, neutrons, and protons.

2. Sentence: **Protons, and electrons are electrically charged.**

2.1 Type: SLS ($5 \leq W \leq 7$)

2.2 SLS → divide into sentencoids

2.3 The list of the created sentencoids

2.3.1 sentencoid 1: Protons
2.3.2 sentencoid 2: and
2.3.3 sentencoid 3: electrons are electrically charged

RW: 0

2.4 Analyses of the created sentencoids

2.4.1 sentencoid 1

2.4.1.1 (1) = OK

2.4.2 sentencoid 2

2.4.2.1 (1) = OK

2.4.3 sentencoid 3

2.4.3.1 “electrons are electrically charged” = OK ($R \geq 3$)

2.5 RW = 0 → Simple language addition

2.6 The addition of the created sentencoids:

Protons + , + and + electrons are electrically charged.

Protons, and electrons are electrically charged.

3. Sentence: **Protons have positive electric charges, and electrons have negative electric charges.**

3.1 Type: TLS ($8 \leq W \leq 14$)

3.2 TLS → divide into sentencoids

3.3 The list of the created sentencoids

3.3.1 sentencoid 1: Protons have positive electric
3.3.2 sentencoid 2: electric charges
3.3.3 sentencoid 3: and
3.3.4 sentencoid 4: electrons have negative electric
3.3.5 sentencoid 5: electric charges

RW: electric, electric

3.4 Analyses of the created sentencoids

3.4.1 sentencoid 1

3.4.1.1 “Protons have positive electric” = OK ($R \geq 3$)

3.4.2 sentencoid 2

3.4.2.1 “electric charges” = OK ($R \geq 3$)

3.4.3 sentencoid 3

3.4.3.1 (1) = OK

3.4.4 sentencoid 4

3.4.4.1 “electrons have negative electric” = OK (R ≥ 3)

3.4.5 sentencoid 5

3.4.5.1 “electric charges” = OK (R ≥ 3)

3.5 RW > 0 → Advanced language addition

3.6 The addition of the created sentencoids:

Protons have positive electric + electric charges + , + and + electrons have negative electric + electric charges.

Protons have positive electric + ~~electric~~ charges + , + and + electrons have negative electric + ~~electric~~ charges.

Protons have positive electric charges, and electrons have negative electric charges.

4. Sentence: **The nucleus is formed by protons, and neutrons.**

4.1 Type: TLS (8 ≤ W ≤ 14)

4.2 TLS → divide into sentencoids

4.3 The list of the created sentencoids
4.3.1 sentencoid 1: The nucleus is formed
4.3.2 sentencoid 2: formed by protons
4.3.3 sentencoid 3: and
4.3.4 sentencoid 4: neutrons

RW: formed

4.4 Analyses of the created sentencoids

4.4.1 sentencoid 1

4.4.1.1 “The nucleus is formed” = OK (R ≥ 3)

4.4.2 sentencoid 2

4.4.2.1 “formed by protons” = OK (R ≥ 3)

4.4.3 sentencoid 3

4.4.3.1 (1) = OK

4.4.4 sentencoid 4

4.4.4.1 (1) = OK

4.5 RW = 0 → Simple language addition

4.6 The addition of the created sentencoids:

The nucleus is formed + formed by protons + , + and + neutrons.

The nucleus is formed + ~~formed~~ by protons + , + and + neutrons.

The nucleus is formed by protons, and neutrons.

5. Sentence: **The consequence of this fact is that the nucleus is positively charged.**

5.1 Type: TLS ($8 \leq W \leq 14$)

5.2 TLS → divide into sentencoids

5.3 The list of the created sentencoids

5.3.1 sentencoid 1:	The consequence of this
5.3.2 sentencoid 2:	this fact is
5.3.3 sentencoid 3:	that
5.3.4 sentencoid 4:	the nucleus is positively
5.3.5 sentencoid 5:	positively charged

RW: this, positively

5.4 Analyses of the created sentencoids

5.4.1 sentencoid 1

5.4.1.1 “The consequence of this” = OK ($R \geq 3$)

5.4.2 sentencoid 2

5.4.2.1 “this fact is” = OK ($R \geq 3$)

5.4.3 sentencoid 3

5.4.3.1 (1) = OK

5.4.4 sentencoid 4

5.4.4.1 “the nucleus is positively” = OK (R ≥ 3)

5.4.5 sentencoid 5

5.4.5.1 “positively charged” = OK (R ≥ 3)

5.5 RW > 0 → Advanced language addition

5.6 The addition of the created sentencoids:

The consequence of this + this fact is + that + the nucleus is positively + positively charged.

The consequence of this + ~~this~~ fact is + that + the nucleus is positively + ~~positively~~ charged.

The consequence of this fact is that the nucleus is positively charged.

6. Sentence: **Because the number of protons, and electrons is equal, the atom is electrically neutral.**

6.1 Type: TLS (8 ≤ W ≤ 14)

6.2 TLS → divide into sentencoids

6.3 The list of the created sentencoids

6.3.1 sentencoid 1:	Because
6.3.2 sentencoid 2:	the number of protons
6.3.3 sentencoid 3:	and
6.3.4 sentencoid 4:	electrons is equal
6.3.5 sentencoid 5:	the atom is electrically
6.3.6 sentencoid 6:	electrically neutral

RW: electrically

6.4.1 sentencoid 1

6.4.1.1 (1) = OK

6.4.2 sentencoid 2

6.4.2.1 “the number of protons” = OK (R ≥ 3)

6.4.3 sentencoid 3

6.4.3.1 (1) = OK

6.4.4 sentencoid 4

6.4.4.1 “electrons is equal” = OK (R ≥ 3)

6.4.5 sentencoid 5

6.4.5.1 “the atom is electrically” = OK (R ≥ 3)

6.4.6 sentencoid 6

6.4.6.1 “electrically neutral” = OK (R ≥ 3)

6.5 RW > 0 → Advanced language addition

6.6 The addition of the created sentencoids:

Because + the number of protons + , + and + electrons is equal + , + the atom is electrically + electrically neutral.

Because + the number of protons + , + and + electrons is equal + , + the atom is electrically + ~~electrically~~ neutral.

Because the number of protons, and electrons is equal, the atom is electrically neutral.

7. Sentence: **The subatomic particles differ in mass.**

7.1 Type: SLS (5 ≤ W ≤ 7)

7.2 SLS → divide into sentencoids

7.3 The list of the created sentencoids

7.3.1 sentencoid 1: The subatomic particles differ
7.3.2 sentencoid 2: differ in mass

RW: differ

7.4 Analyses of the created sentencoids

7.4.1 sentencoid 1

7.4.1.1 “The subatomic particles differ” = OK (R ≥ 3)

7.4.2 sentencoid 2

7.4.2.1 “differ in mass” = OK (R ≥ 3)

7.5 RW > 0 → Advanced language addition

7.6 The addition of the created sentencoids:

The subatomic particles differ + differ in mass.

The subatomic particles differ + ~~differ~~ in mass.

The subatomic particles differ in mass.

8. Sentence: **The mass of each electron is a fraction when we will compare it to the unified atomic mass unit.**

8.1 Type: FhLS (W ≥ 15)

8.2 FhLS → divide into sentencoids

8.3 The list of the created sentencoids

8.3.1 sentencoid 1: The mass of each
8.3.2 sentencoid 2: each electron is a
8.3.3 sentencoid 3: is a fraction
8.3.4 sentencoid 4: when
8.3.5 sentencoid 5: we will compare it
8.3.6 sentencoid 6: it to the unified
8.3.7 sentencoid 7: unified atomic mass unit

RW: each, is, a, it, unified

8.4 Analyses of the created sentencoids

8.4.1 sentencoid 1

8.4.1.1 “The mass of each” = OK (R ≥ 3)

8.4.2 sentencoid 2

8.4.2.1 “each electron is a” = OK (R ≥ 3)

8.4.3 sentencoid 3

8.4.3.1 “is a fraction” = OK (R ≥ 3)

8.4.4 sentencoid 4

8.4.4.1 (1) = OK

8.4.5 sentencoid 5

8.4.5.1 “we will compare it” = OK (R ≥ 3)

8.4.6 sentencoid 6

8.4.6.1 “it to the unified” = OK (R ≥ 3)

8.4.7 sentencoid 7

8.4.7.1 “unified atomic mass unit” = OK (R ≥ 3)

8.5 RW > 0 → Advanced language addition

8.6 The addition of the created sentencoids:

The mass of each + each electron is a + is a fraction + when + we will compare it + it to the unified + unified atomic mass unit.

The mass of each + ~~each~~ electron is a + ~~is a~~ fraction + when + we will compare it + ~~it~~ to the unified + ~~unified~~ atomic mass unit.

The mass of each electron is a fraction when we will compare it to the unified atomic mass unit.

9. Sentence: **In contrast, the mass of each proton (and also the mass of each neutron) is almost equal to the unified atomic mass unit.**

9.1 Type: FhLS (W ≥ 15)

9.2 FhLS → divide into sentencoids

9.3 The list of the created sentencoids

9.3.1 sentencoid 1:	In contrast
9.3.2 sentencoid 2:	the mass of each
9.3.3 sentencoid 3:	each proton is almost
9.3.4 sentencoid 4:	almost equal to the
9.3.5 sentencoid 5:	to the unified atomic
9.3.6 sentencoid 6:	atomic mass unit

RW: each, almost, to, the, atomic

9.3.3 + 1:	and
9.3.3 + 2:	also the mass of
9.3.3 + 3:	mass of each neutron

RW for 9.3.3 + 1–3: mass, of

9.4 Analyses of the created sentencoids

9.4.1 sentencoid 1

9.4.1.1 “In contrast” = OK (R ≥ 3)

9.4.2 sentencoid 2

9.4.2.1 “the mass of each” = OK (R ≥ 3)

9.4.3 sentencoid 3

9.4.3.1 “each proton is almost” = OK (R ≥ 3)

9.4.4 sentencoid 4

9.4.4.1 “almost equal to the” = OK (R ≥ 3)

9.4.5 sentencoid 5

9.4.5.1 “to the unified atomic” = OK (R ≥ 3)

9.4.6 sentencoid 6

9.4.6.1 “atomic mass unit” = OK (R ≥ 3)

9.4.3 + 1

9.4.3 + 1.1 (1) = OK

9.4.3 + 2

9.4.3 + 2.1 “also the mass of” = OK (R ≥ 3)

9.4.3 + 3

9.4.3 + 3.1 “mass of each neutron” = OK (R ≥ 3)

9.5 RW > 0 → Advanced language addition

9.6 The addition of the created sentencoids:

In contrast + , + the mass of each + each proton is almost + almost equal to the + to the unified atomic + atomic mass unit.

and + also the mass of + mass of each neutron

In contrast + , + the mass of each + ~~each~~ proton is almost + ~~almost~~ equal to the + ~~to the~~ unified atomic + ~~atomic~~ mass unit.

and + also the mass of + ~~mass of~~ each neutron

In contrast, the mass of each proton (and also the mass of each neutron) is almost equal to the unified atomic mass unit.

10. Sentence: **These facts lead to the conclusion that the nucleus holds almost all of the mass of every atom.**

10.1 Type: FhLS (W ≥ 15)

10.2 FhLS → divide into sentencoids

10.3 The list of the created sentencoids

10.3.1 sentencoid 1: These facts lead to
10.3.2 sentencoid 2: to the conclusion
10.3.3 sentencoid 3: that
10.3.4 sentencoid 4: the nucleus holds almost
10.3.5 sentencoid 5: almost all of the
10.3.6 sentencoid 6: of the mass of
10.3.7 sentencoid 7: mass of every atom.

RW: to, almost, of, the, mass, of

10.4 Analyses of the created sentencoids

10.4.1 sentencoid 1

10.4.1.1 “These facts lead to” = OK (R ≥ 3)

10.4.2 sentencoid 2

10.4.2.1 “to the conclusion” = OK (R ≥ 3)

10.4.3 sentencoid 3

10.4.3.1 (1) = OK

10.4.4 sentencoid 4

10.4.4.1 “the nucleus holds almost” = OK (R ≥ 3)

10.4.5 sentencoid 5

10.4.5.1 “almost all of the” = OK (R ≥ 3)

10.4.6 sentencoid 6

10.4.6.1 “of the mass of” = OK (R ≥ 3)

10.4.7 sentencoid 7

10.4.7.1 “mass of every atom” = OK (R ≥ 3)

10.5 RW > 0 → Advanced language addition

10.6 The addition of the created sentencoids:

These facts lead to + to the conclusion + that + the nucleus holds almost + almost all of the + of the mass of + mass of every atom.

These facts lead to + ~~to~~ the conclusion + that + the nucleus holds almost + ~~almost~~ all of the + ~~of the~~ mass of + ~~mass of~~ every atom.

These facts lead to the conclusion that the nucleus holds almost all of the mass of every atom.

11. Sentence: **Probably the best way to visualize the atom, its inner structure, is to compare it to the Solar System.**

11.1 Type: FhLS (W ≥ 15)

11.2 FhLS → divide into sentencoids

11.3 The list of the created sentencoids

11.3.1 sentencoid 1: Probably the best way
11.3.2 sentencoid 2: way to visualize the
11.3.3 sentencoid 3: visualize the atom is
11.3.4 sentencoid 4: is to compare it
11.3.5 sentencoid 5: it to the Solar System

RW: way, visualize, the, is, it

11.3.3 + 1 its inner structure

11.4 Analyses of the created sentencoids

11.4.1 sentencoid 1

11.4.1.1 “Probably the best way” = OK (R ≥ 3)

11.4.2 sentencoid 2

11.4.2.1 “way to visualize the” = OK (R ≥ 3)

11.4.3 sentencoid 3

11.4.3.1 “visualize the atom is” = OK (R ≥ 3)

11.4.4 sentencoid 4

11.4.4.1 “is to compare it” = OK (R ≥ 3)

11.4.5 sentencoid 5

11.4.5.1 “it to the Solar System” = OK (R ≥ 3)

11.4.3 + 1

11.4.3 + 1.1 “its inner structure” = OK (R ≥ 3)

11.5 RW > 0 → Advanced language addition

11.6 The addition of the created sentencoids:

Probably the best way + way to visualize the + visualize the atom is + is to compare it + it to the Solar System.

its inner structure

Probably the best way + ~~way~~ to visualize the + ~~visualize the~~ atom is + ~~is~~ to compare it + ~~it~~ to the Solar System.

its inner structure

Probably the best way to visualize the atom, its inner structure, is to compare it to the Solar System.

12. Sentence: **According to the planetary model, electrons act like planets.**

12.1 Type: TLS (8 ≤ W ≤ 14)

12.2 TLS → divide into sentencoids

12.3 The list of the created sentencoids

12.3.1 sentencoid 1: According to the planetary
12.3.2 sentencoid 2: planetary model
12.3.3 sentencoid 3: electrons act like planets

RW: planetary

12.4 Analyses of the created sentencoids

12.4.1 sentencoid 1

12.4.1.1 “According to the planetary” = OK (R ≥ 3)

12.4.2 sentencoid 2

12.4.2.1 “planetary model” = OK (R ≥ 3)

12.4.3 sentencoid 3

12.4.3.1 “electrons act like planets” = OK (R ≥ 3)

12.5 RW > 0 → Advanced language addition

12.6 The addition of the created sentencoids:

According to the planetary + planetary model + , + electrons act like planets.

According to the planetary + ~~planetary~~ model +, + electrons act like planets.

According to the planetary model, electrons act like planets.

13. Sentence: **They orbit the nucleus of the atom.**

13.1 Type: SLS ($5 \leq W \leq 7$)

13.2 SLS → divide into sentencoids

13.3 The list of the created sentencoids

13.3.1 sentencoid 1: they orbit the nucleus
13.3.2 sentencoid 2: nucleus of the atom

RW: nucleus

13.4 Analyses of the created sentencoids

13.4.1 sentencoid 1

13.4.1.1 “they orbit the nucleus” = OK ($R \geq 3$)

13.4.2 sentencoid 2

13.4.2.1 “nucleus of the atom” = OK ($R \geq 3$)

13.5 RW > 0 → Advanced language addition

13.6 The addition of the created sentencoids:

They orbit the nucleus + nucleus of the atom.

They orbit the nucleus + ~~nucleus~~ of the atom.

They orbit the nucleus of the atom.

14. Sentence: **Valence electrons, the most distant electrons, create chemical bonds between atoms.**

14.1 Type: TLS ($8 \leq W \leq 14$)

14.2 TLS → divide into sentencoids

14.3 The list of the created sentencoids

14.3.1 sentencoid 1: Valence electrons create chemical
14.3.2 sentencoid 2: chemical bonds between atoms

RW: chemical

14.3.1 + 1 the most distant electrons

14.4 Analyses of the created sentencoids

14.4.1 sentencoid 1

14.4.1.1 “Valence electrons create chemical” = OK ($R \geq 3$)

14.4.2 sentencoid 2

14.4.2.1 “chemical bonds between atoms” = OK ($R \geq 3$)

14.4.1 + 1

14.4.1 + 1.1 “the most distant electrons” = OK ($R \geq 3$)

14.5 RW > 0 → Advanced language addition

14.6 The addition of the created sentencoids:

Valence electrons create chemical + chemical bonds between atoms.

the most distant electrons

Valence electrons create chemical + ~~chemical~~ bonds between atoms.

the most distant electrons

Valence electrons, the most distant electrons, create chemical bonds between atoms.

15. Sentence: **The nucleus plays the role of the Sun.**

15.1 Type: TLS ($8 \leq W \leq 14$)

15.2 TLS → divide into sentencoids

15.3 The list of the created sentencoids

15.3.1 sentencoid 1: The nucleus plays the
15.3.2 sentencoid 2: plays the role of
15.3.3 sentencoid 3: role of the Sun

RW: plays, the, role, of

15.4 Analyses of the created sentencoids

15.4.1 sentencoid 1

15.4.1.1 “The nucleus plays the” = OK (R ≥ 3)

15.4.2 sentencoid 2

15.4.2.1 “plays the role of” = OK (R ≥ 3)

15.4.3 sentencoid 3

15.4.3.1 “role of the Sun” = OK (R ≥ 3)

15.5 RW > 0 → Advanced language addition

15.6 The addition of the created sentencoids:

The nucleus plays the + plays the role of + role of the Sun.

The nucleus plays the + ~~plays the~~ role of + ~~role of~~ the Sun.

The nucleus plays the role of the Sun.

16. Sentence: **The planetary model is easy to understand; however, it confused physicists.**

16.1 Type: TLS (8 ≤ W ≤ 14)

16.2 TLS → divide into sentencoids

16.3 The list of the created sentencoids

16.3.1 sentencoid 1: The planetary model is
16.3.2 sentencoid 2: is easy to understand
16.3.3 sentencoid 3: however
16.3.4 sentencoid 4: it confused physicists
RW: is

16.4 Analyses of the created sentencoids

16.4.1 sentencoid 1

16.4.1.1 “The planetary model is” = OK (R ≥ 3)

16.4.2 sentencoid 2

16.4.2.1 “is easy to understand” = OK (R ≥ 3)

16.4.3 sentencoid 3

16.4.3.1 (1) = OK

16.4.4 sentencoid 4

16.4.4.1 “it confused physicists” = INCORRECT (R < 3)

16.4.4.2 [physicists] = chemists

16.4.4.3 “it confused chemists” = OK (R ≥ 3)

16.4.4.4 physicists = chemists

16.4.4.5 “it confused physicists” [OK]

16.5 RW > 0 → Advanced language addition

16.6 The addition of the created sentencoids:

The planetary model is + is easy to understand + ; + however + , + it confused physicists.

The planetary model is + ~~is~~ easy to understand + ; + however + , + it confused physicists.

The planetary model is easy to understand; however, it confused physicists.

17. Sentence: **Why negatively charged electrons do not fall into the positively charged nucleus?**

17.1 Type: TLS (8 ≤ W ≤ 14)

17.2 TLS → divide into sentencoids

17.3 The list of the created sentencoids

17.3.1 sentencoid 1: Why
17.3.2 sentencoid 2: negatively charged electrons do
17.3.3 sentencoid 3: do not fall into
17.3.4 sentencoid 4: into the positively charged
17.3.5 sentencoid 5: charged nucleus

RW: do, into, charged

17.4 Analyses of the created sentencoids

17.4.1 sentencoid 1

17.4.1.1 (1) = OK

17.4.2 sentencoid 2

17.4.2.1 “negatively charged electrons do” = OK (R ≥ 3)

17.4.3 sentencoid 3

17.4.3.1 “do not fall into”= OK (R ≥ 3)

17.4.4 sentencoid 4

17.4.4.1 “into the positively charged” = OK (R ≥ 3)

17.4.5 sentencoid 5

17.4.5.1 “charged nucleus” = OK (R ≥ 3)

17.5 RW > 0 → Advanced language addition

17.6 The addition of the created sentencoids:

Why + negatively charged electrons do + do not fall into + into the positively charged + charged nucleus

Why + negatively charged electrons do + ~~do~~ not fall into + ~~into~~ the positively charged + ~~charged~~ nucleus

Why negatively charged electrons do not fall into the positively charged nucleus?

18. Sentence: **The subatomic world required a correction.**

18.1 Type: SLS (5 ≤ W ≤ 7)

18.2 SLS → divide into sentencoids

18.3 The list of the created sentencoids

18.3.1 sentencoid 1: The subatomic world required
18.3.2 sentencoid 2: required a correction

RW: required
18.4 Analyses of the created sentencoids

18.4.1 sentencoid 1

18.4.1.1 “The subatomic world required” = OK (R ≥ 3)

18.4.2 sentencoid 2

18.4.2.1 “required a correction” = OK (R ≥ 3)

18.5 RW > 0 → Advanced language addition

18.6 The addition of the created sentencoids:

The subatomic world required + required a correction.

The subatomic world required + ~~required~~ a correction.

The subatomic world required a correction.

19. Sentence: **New views on the atom have been born.**

19.1 Type: TLS ($8 \leq W \leq 14$)

19.2 TLS → divide into sentencoids

19.3 The list of the created sentencoids

19.3.1 sentencoid 1: New views on the
19.3.2 sentencoid 2: on the atom have
19.3.3 sentencoid 3: have been born

RW: on, the, have

19.4 Analyses of the created sentencoids

19.4.1 sentencoid 1

19.4.1.1 “New views on the” = OK ($R \geq 3$)

19.4.2 sentencoid 2

19.4.2.1 “on the atom have” = OK ($R \geq 3$)

19.4.3 sentencoid 3

19.4.3.1 “have been born” = OK ($R \geq 3$)

19.5 RW > 0 → Advanced language addition

19.6 The addition of the created sentencoids:

New views on the + on the atom have + have been born.

New views on the + ~~on the~~ atom have + ~~have~~ been born.

New views on the atom have been born.

20. Sentence: **The quantum theory states that it is impossible to give the exact location of an electron.**

20.1 Type: FhLS ($W \geq 15$)

20.2 FhLS → divide into sentencoids

20.3 The list of the created sentencoids

20.3.1 sentencoid 1:	The quantum theory states
20.3.2 sentencoid 2:	that
20.3.3 sentencoid 3:	it is impossible to
20.3.4 sentencoid 4:	to give the exact
20.3.5 sentencoid 5:	exact location of an
20.3.6 sentencoid 6:	of an electron

RW: to, exact, of, an

20.4 Analyses of the created sentencoids

20.4.1 sentencoid 1

20.4.1.1 “The quantum theory states” = OK (R ≥ 3)

20.4.2 sentencoid 2

20.4.2.1 (1) = OK

20.4.3 sentencoid 3

20.4.3.1 “it is impossible to” = OK (R ≥ 3)

20.4.4 sentencoid 4

20.4.4.1 “to give the exact” = OK (R ≥ 3)

20.4.5 sentencoid 5

20.4.5.1 “exact location of an” = OK (R ≥ 3)

20.4.6 sentencoid 6

20.4.6.1 “of an electron” = OK (R ≥ 3)

20.5 The addition of the created sentencoids:

20.6 RW > 0 → Advanced language addition

The quantum theory states + that + it is impossible to + to give the exact + exact location of an + of an electron.

The quantum theory states + that + it is impossible to + ~~to~~ give the exact + ~~exact~~ location of an + ~~of an~~ electron.

The quantum theory states that it is impossible to give the exact location of an electron.

21. Sentence: **Its location can be estimated only.**

21.1 Type: SLS ($5 \leq W \leq 7$)

21.2 SLS → divide into sentencoids

21.3 The list of the created sentencoids

21.3.1 sentencoid 1: Its location can be
21.3.2 sentencoid 2: be estimated only

RW: be

21.4 Analyses of the created sentencoids

21.4.1 sentencoid 1

21.4.1.1 “Its location can be” = OK ($R \geq 3$)

21.4.2 sentencoid 2

21.4.2.1 “be estimated only” = OK ($R \geq 3$)

21.5 The addition of the created sentencoids:

21.6 RW > 0 → Advanced language addition

Its location can be + be estimated only

Its location can be + ~~be~~ estimated only

Its location can be estimated only.

22. Sentence: **Such calculation is based on probability.**

22.1 Type: SLS ($5 \leq W \leq 7$)

22.2 SLS → divide into sentencoids

22.3 The list of the created sentencoids

22.3.1 sentencoid 1: Such calculation is based
22.3.2 sentencoid 2: based on probability

RW: based

22.4 Analyses of the created sentencoids

22.4.1 sentencoid 1

22.4.1.1 “Such calculation is based” = OK (R ≥ 3)

22.4.2 sentencoid 2

22.4.2.1 “based on probability” = OK (R ≥ 3)

22.5 The addition of the created sentencoids:

22.6 RW > 0 → Advanced language addition

Such calculation is based + based on probability.

Such calculation is based + ~~based~~ on probability.

Such calculation is based on probability.

23. Sentence: **The theory was established in the twentieth century.**

23.1 Type: SLS (5 ≤ W ≤ 7)

23.2 SLS → divide into sentencoids

23.3 The list of the created sentencoids

23.3.1 sentencoid 1: The theory was established
23.3.2 sentencoid 2: established
23.3.3 sentencoid 3: in the twentieth century

RW: established

23.4 Analyses of the created sentencoids

23.4.1 sentencoid 1

23.4.1.1 “The theory was established” = OK (R ≥ 3)

23.4.2 sentencoid 2

23.4.2.1 (1) = OK

23.4.3 sentencoid 3

23.4.3.1 “in the twentieth century” = OK (R ≥ 3)

23.5 RW > 0 → Advanced language addition

23.6 The addition of the created sentencoids:

The theory was established + established + in the twentieth century.

The theory was established + ~~established~~ + in the twentieth century.

The theory was established in the twentieth century.

24. Sentence: **It describes, and rules the subatomic world.**

24.1 Type: SLS ($5 \leq W \leq 7$)

24.2 SLS → divide into sentencoids

24.3 The list of the created sentencoids

24.3.1 sentencoid 1: It describes
24.3.2 sentencoid 2: and
24.3.3 sentencoid 3: rules the subatomic world

RW: 0

24.4 Analyses of the created sentencoids

24.4.1 sentencoid 1

24.4.1.1 “It describes” = OK ($R \geq 3$)

24.4.2 sentencoid 2

24.4.2.1 (1) = OK

24.4.3 sentencoid 3

24.4.3.1 “rules the subatomic world” = OK ($R \geq 3$)

24.5 RW = 0 → Simple language addition:

24.6 The addition of the created sentencoids:

It describes + , + and + rules the subatomic world.

It describes, and rules the subatomic world.

3.3 Comparison between the analyses

Sentence 3: Protons have positive electric charges, and electrons have negative electric charges.

According to the grammar of the English language:

Type of sentence:	compound sentence
First clause:	Protons have positive electric charges
Coordinator:	and
Second clause:	electrons have negative electric charges

According to the theory of sentencoids:

Type of sentence:	third level sentence
sentencoid 1:	Protons have positive electric
sentencoid 2:	electric charges
sentencoid 3:	and
sentencoid 4:	electrons have negative electric
sentencoid 5:	electric charges

Sentence 5: The consequence of this fact is that the nucleus is positively charged.

According to the grammar of the English language:

Type of sentence:	complex sentence
First clause:	The consequence of this fact is
Subordinator:	that
Second clause:	the nucleus is positively charged.

According to the theory of sentencoids:

Type of sentence:	third level sentence
5.3.1 sentencoid 1:	The consequence of this
5.3.2 sentencoid 2:	this fact is
5.3.3 sentencoid 3:	that
5.3.4 sentencoid 4:	the nucleus is positively
5.3.5 sentencoid 5:	positively charged

Sentence 7: The subatomic particles differ in mass.

According to the grammar of the English language:

Type of sentence:	simple sentence
First clause:	The subatomic particles differ in mass.

According to the theory of sentencoids:

Type of sentence:	second level sentence
7.3.1 sentencoid 1:	The subatomic particles differ

7.3.2 sentencoid 2: differ in mass

Sentence 8: The mass of each electron is a fraction when we will compare it to the unified atomic mass unit.

According to the grammar of the English language:

Type of sentence:	complex sentence
First clause:	The mass of each electron is a fraction
Subordinator:	when
Second clause:	will compare it to the unified atomic mass unit

According to the theory of sentencoids:

Type of sentence:	fourth level sentence
sentencoid 1:	The mass of each
sentencoid 2:	each electron is a
sentencoid 3:	is a fraction
sentencoid 4:	when
sentencoid 5:	we will compare it
sentencoid 6:	it to the unified
sentencoid 7:	unified atomic mass unit

Sentence 10: These facts lead to the conclusion that the nucleus holds almost all of the mass of every atom.

According to the grammar of the English language:

Type of sentence:	complex sentence
First clause:	These facts lead to the conclusion
Subordinator:	that
Second clause:	the nucleus holds almost all of the mass of every atom.

According to the theory of sentencoids:

Type of sentence:	fourth level sentence
sentencoid 1:	These facts lead to
sentencoid 2:	to the conclusion
sentencoid 3:	that
sentencoid 4:	the nucleus holds almost
sentencoid 5:	almost all of the
sentencoid 6:	of the mass of
sentencoid 7:	mass of every atom.

Sentence 15: The nucleus plays the role of the Sun.

According to the grammar of the English language:

Type of sentence:	simple sentence
First clause:	The nucleus plays the role of the Sun

According to the theory of sentencoids:

Type of sentence:	third level sentence
sentencoid 1:	The nucleus plays the
sentencoid 2:	plays the role of
sentencoid 3:	role of the Sun

Sentence 18: The subatomic world required a correction.

According to the grammar of the English language:

Type of sentence:	simple sentence
First clause:	The subatomic world required a correction.

According to the theory of sentencoids:

Type of sentence:	second level sentence
sentencoid 1:	The subatomic world required
sentencoid 2:	required a correction

Sentence 20: The quantum theory states that it is impossible to give the exact location of an electron.

According to the grammar of the English language:

Type of sentence:	complex sentence
First clause:	The quantum theory states
Subordinator:	that
Second clause:	it is impossible to give the exact location of an electron

According to the theory of sentencoids:

Type of sentence:	fourth level sentence
sentencoid 1:	The quantum theory states
sentencoid 2:	that
sentencoid 3:	it is impossible to
sentencoid 4:	to give the exact
sentencoid 5:	exact location of an
sentencoid 6:	of an electron

Sentence 22: Such calculation is based on probability.

According to the grammar of the English language:

Type of sentence:	simple sentence
First clause:	Such calculation is based on probability.

According to the theory of sentencoids:

Type of sentence:	second level sentence
22.3.1 sentencoid 1:	Such calculation is based
22.3.2 sentencoid 2:	based on probability

Sentence 23: The theory was established in the twentieth century.

According to the grammar of the English language:

Type of sentence:	simple sentence
First clause:	The theory was established in the twentieth century

According to the theory of sentencoids:

Type of sentence:	second level sentence
sentencoid 1:	The theory was established
sentencoid 2:	established
sentencoid 3:	in the twentieth century

Conclusions:

1. Coordinators can connect both independent clauses, and sentencoids,

2. A coordinator always stands alone,

3. A subordinator does not create sentencoids with other words

4 Summary

The last section summarizes the paper. The section confronts the theory with the grammar of the English language, and also suggests directions for the future work.

4.1 The theory of sentencoids, and the grammar of the English language—The differences and similarities

The following subsection confronts the theory of sentencoids with the grammar of the English language. This is done by analyzing the differences, and similarities between them.

The similarities:

1. The status of parenthetical material

A pair of parentheses contains content that can be removed from the sentence. The lack of such content will not affect the meaning of the sentence; the sentence will be still grammatically correct. Consider the following sentence:

(69a) I can say that my cat (a Maine Coon kitten) is the most beautiful animal I have ever seen.

The parenthetical material will be removed:

(69b) I can say that my cat is the most beautiful animal I have ever seen.

From the theory's point of view, we can freely remove parenthetical material too. This is possible, because the text inside a pair of parentheses does not create sentencoids with the text outside parentheses—see the analysis of the sentence 69a.

The sentence 69a will be analyzed in accordance with the theory:

(69a) I can say that my cat (a Maine Coon kitten) is the most beautiful animal I have ever seen.

69a.1 Type: FhLS ($W \geq 15$)

69a.2 FhLS → divide into sentencoids

69a.3 The list of the created sentencoids

69a.3.1 sentencoid 1: I can say
69a.3.2 sentencoid 2: that
69a.3.3 sentencoid 3: my cat is the
69a.3.4 sentencoid 4: is the most beautiful
69a.3.5 sentencoid 5: beautiful animal I have
69a.3.6 sentencoid 6: have ever seen

RW: is, the, beautiful, have

69a.3.3 + 1: a Maine Coon kitten

RW for 69a.3.3 + 1: 0

69a.4 Analyses of the created sentencoids

69a.4.1 sentencoid 1

69a.4.1.1 "I can say" = OK ($R \geq 3$)

69a.4.2 sentencoid 2

69a.4.2.1 (1) = OK

69a.4.3 sentencoid 3

69a.4.3.1 “my cat is the” = OK (R ≥ 3)

69a.4.4 sentencoid 4

69a.4.4.1 “is the most beautiful” = OK (R ≥ 3)

69a.4.5 sentencoid 5

69a.4.5.1 “beautiful animal I have” = OK (R ≥ 3)

69a.4.6 sentencoid 6

69a.4.6.1 “have ever seen” = OK (R ≥ 3)

69a.4.3 + 1

69a.4.3 + 1.1 “a Maine Coon kitten” = OK (R ≥ 3)

69a.5 RW > 0 → Advanced language addition

69a.6 The addition of the created sentencoids:

I can say + that + my cat is the + is the most beautiful + beautiful animal I have + have ever seen.

a Maine Coon kitten

I can say + that + my cat is the + ~~is the~~ most beautiful + ~~beautiful~~ animal I have + ~~have~~ ever seen.

a Maine Coon kitten

I can say that my cat (a Maine Coon kitten) is the most beautiful animal I have ever seen.

2. The status of words that interrupt the flow of a sentence.

Words that interrupt the flow of a sentence are set off from the rest of the sentence by two en-dashes, or two commas. Consider the following example:

(70a) My new friend—a real fan of football—is going to watch some movies.

A non–essential clause is set off by commas, or en–dashes. The clause of course belongs to the sentence.

From the theory's point of view, words that interrupt the flow of a sentence do not create sentencoids with the rest of the sentence; they create a separate group of sentencoids. However, they cannot be removed from the sentence. To illustrate this, the sentence 70a will be analyzed in accordance with the theory:

(70a) My new friend—a real fan of football—is going to watch some movies.

70a.1 Type: TLS ($8 \leq W \leq 14$)

70a.2 TLS → divide into sentencoids

70a.3 The list of the created sentencoids

70a.3.1 sentencoid 1: My new friend is
70a.3.2 sentencoid 2: is going to watch
70a.3.3 sentencoid 3: watch some movies

RW: is, watch

70a.3.1 + 1: a real fan of
70a.3.1 + 2: fan of football

RW for 70a.3.1–2: fan, of

70a.4 Analyses of the created sentencoids

70a.4.1 sentencoid 1

70a.4.1.1 "My new friend is" = OK ($R \geq 3$)

70a.4.2 sentencoid 2

70a.4.2.1 "is going to watch" = OK ($R \geq 3$)

70a.4.3 sentencoid 3

70a.4.3.1 "watch some movies" = OK ($R \geq 3$)

70a.4.3.1 + 1

70a.4.3.1 + 1.1 "a real fan of" = OK ($R \geq 3$)

70a.4.3.1 + 2

70a.4.3.1 + 2.1 "a real fan of" = OK ($R \geq 3$)

70a.5 RW > 0 → Advanced language addition

70a.6 The addition of the created sentencoids:

My new friend is + is going to watch + watch some movies

a real fan of + fan of football

My new friend is + ~~is~~ going to watch + ~~watch~~ some movies

a real fan of + ~~fan of~~ football

My new friend, a real fan of football, is going to watch some movies.

Conclusion: words that interrupt the flow of a sentence create a separate group of sentencoids. They do not share words with other sentencoids.

3. The status of a coordinating conjunction (coordinator):

A coordinating conjunction is used to connect at least two independent clauses in order to form a compound sentence. However, a coordinating conjunction itself does not belong to any independent clause.

From the theory's point of view, a coordinating conjunction also stands alone. It always stands as a sentencoid consisting of only one word. The sentence 71 will be analyzed in accordance with the theory:

(71) Sentence: I like to play chess, but I am not a professionalist.

71.1 Type: TLS ($8 \leq W \leq 14$)

71.2 TLS → divide into sentencoids

71.3 The list of the created sentencoids

71.3.1 sentencoid 1: I like to play
71.3.2 sentencoid 2: play chess
71.3.3 sentencoid 3: but
71.3.4 sentencoid 4: I am not a
71.3.5 sentencoid 5: not a professionalist

RW: play, not, a

71.4 Analyses of the created sentencoids

71.4.1 sentencoid 1

71.4.1.1 "I like to play" = OK ($R \geq 3$)

71.4.2 sentencoid 2

71.4.2.1 "play chess" = OK ($R \geq 3$)

71.4.3 sentencoid 3

71.4.3.1 (1) = OK

71.4.4 sentencoid 4

71.4.4.1 “I am not a” = OK ($R \geq 3$)

71.4.5 sentencoid 5

71.4.5.1 “not a professionalist” = OK ($R \geq 3$)

71.5 $RW > 0 \rightarrow$ Advanced language addition

71.6 The addition of the created sentencoids:

I like to play + play chess + , + but + I am not a + not a professionalist

I like to play + ~~play~~ chess + , + but + I am not a + ~~not a~~ professionalist

I like to play chess, but I am not a professionalist.

The coordinating conjunction “but” stands alone. It is present only in the third sentencoid.

4. The role of commas

Commas divide a sentence (the whole) into clauses (parts). In a complex sentence, the comma signals when the previous clause ends, and when the next clause starts. A list of words is an exception to this.

From the theory’s point of view, commas divide a sentence (the whole) into sentencoids (parts). Each comma introduces an external border—i.e. the comma signals when the previous sentencoids ends, and where the next one starts. A list of words is an exception to this.

5. The presence of a comma in short compound sentences

In a short and balanced compound sentence, the comma is optional. No matter if a comma is used or not, the coordinating conjunction separates two words, or clauses—see the following sentences:

Cars and bicycles.

He swims and she reads.

Robert likes oranges but Tom prefers watermelons.

From the theory’s point of view, in a short and balanced compound sentence, the lack of the comma is not a problem. The coordinating conjunction itself can act like a comma, introducing internal borders—see the short analyses of the following sentences:

(72) Cars and bicycles.

sentencoid 1: Cars
sentencoid 2: and
sentencoid 3: bicycles

(73) He swims and she reads.

sentencoid 1: He swims
sentencoid 2: and
sentencoid 3: she reads

(74) Robert likes oranges but Tom prefers watermelons.

sentencoid 1: Robert likes oranges
sentencoid 2: but
sentencoid 3: Tom prefers watermelons

Differences:

1. The proper use of punctuation marks

In the grammar of the English language, the improper use of a comma is not always a serious problem. It is of course an error, but the sentence in most cases can be still understood. Its meaning is not always lost, because of the error.

From the theory's point of view, the improper use of a punctuation mark like a comma has serious consequences. The created sentencoids will contain different words in comparison with the sentencoids that were created from the original sentence in which punctuation marks were used properly. That is why, the theory puts emphasis on the proper use of punctuation marks.

2. The case of word classes

A word can be replaced by another word when both of them belong to the same word class. All words that can replace the previous word create a set of words. Like all other sets, such set of words contains the empty set. Its presence allows to change a passive construction into the short passive. Absence of the performer, and the preposition "by" can be explained using the concepts of word classes, and the empty set.

3. The New terms

The grammar of the English language uses different clauses in order to describe the structure of a sentence, and to distinguish various kinds of sentences, like a complex sentence, a complex–compound sentence, etc.

In contrast, the theory of sentencoids has replaced the earlier mentioned terms with the term "sentencoid." Because the classification of sentences is no longer based on clauses, terms like simple, complex, compound, and compound–complex sentence are not longer used.

The theory uses terms that are not present in the grammar of the English language. These terms are: an external boundary, an internal boundary, repeated words, a sentencoid, a first, second, third, and fourth level sentence, the process of breaking and re–creating an original sentence.

4. Language addition

The use of language addition is a consequence of the assumption that a sentence can be divided into sentencoids. Language addition makes possible to form the original sentence, by adding created sentencoids.

5. Checking correctness of a sentence

In order to check correctness of the analyzed sentence the process of breaking and re-creating an original sentence is used. The process consists of six stages.

6. The use of equations

The theory of sentencoids uses equations to write the changes that occur during the process of breaking and re-creating an original sentence.

7. The correct placement of a comma

A comma that separates two words must be placed next to the last letter of the previous word, and space must be present between the comma, and the first letter of the next word:

I like apples, grapes, oranges, and watermelons.

The following persons are guilty: Tom, Robert, and Sarah.

When this rule is violated, the comma is placed incorrectly—see the following sentence:

I can speak English , French ,Russian , and Spanish.

From the theory's point of view the incorrect placement of a comma has no effect on the content of the created sentencoids.

4.2 Conclusions

1. Sentences have been classified into four categories. The status of a sentence is based on the quantity of words that a sentence contains.

2. A sentence is divided into parts, called sentencoids, in order to check its correctness.

3. The theory of sentencoids describes the conflict between the classification that is based on quantity of words, and the presence of a proper noun that consists of more than one word. The conflict is caused by the fact that a sentence can be visually classified as a second, third, or a fourth level sentence; however, according to the theory of sentencoids such classification is misleading,

4. In most cases the last word of the previous sentencoid is repeated at the beginning of the next sentencoid.

5. Articles, and the word “of” have received a special status. Each time one of them appears at the end of a sentencoid, it is repeated together with the word that precedes it (repeated words—RW) at the beginning of the next sentencoid.

6. The theory of sentencoids describes two types of borders:

6a External borders that are introduced by exclamations points, interrobangs, periods, and question marks. The function of the external border is to end a sentence, and start the new one,

6b Internal borders that are introduced by commas, en–dashes, parentheses, quotation marks, and semicolons. The function of the internal border is to divide an original sentence into parts called sentencoids.

7. The theory of sentencoids emphasizes the correct use of punctuation marks because of their role in introducing external, and internal borders.

8. Language addition allows to re–create original sentences. They are two types of language addition:

8a simple language addition,

8b advanced language addition.

The value of RW determinates, which type of language addition will be used.

9. A word can be morphed (changed) into another word only within the same word class. All words that can be used to replace the original word, and meet the additional requirement(s) create a specific set of words.

10. Each set of words consists of the empty set. The empty set allows to remove the selected word(s).

4.3 When the theory cannot be applied–the limitations

1. The case of fused sentences

The term “fused sentences” describes:

1) the lack of a comma, and a coordinating conjunction between two independent clauses

2) the lack of a semicolon between two independent clauses.

The lack of is classified as an error. The sentence 75 illustrates the error:

(75) My cat is very clever he is always hungry.

The sentence will be analyzed in accordance with the theory:

(75) My cat is very clever he is always hungry.

75.1 Type: TLS ($8 \leq W \leq 14$)

75.2 TLS → divide into sentencoids

75.3 The list of the created sentencoids

75.3.1 sentencoid 1: My cat is very
75.3.2 sentencoid 2: very clever he is
75.3.3 sentencoid 3: is always hungry

RW: very, is

75.4 Analyses of the created sentencoids

75.4.1 sentencoid 1

75.4.1.1 “My cat is very” = OK ($R \geq 3$)

75.4.2 sentencoid 2

75.4.2.1 “very clever he is” = OK ($R \geq 3$)

75.4.3 sentencoid 3

75.4.3.1 “is always hungry” = OK ($R \geq 3$)

75.5 RW > 0 → Advanced language addition

75.6 The addition of the created sentencoids:

My cat is very + very clever he is + is always hungry

My cat is very + ~~very~~ clever he is + ~~is~~ always hungry

My cat is very clever he is always hungry.

The second sentencoid that contains words from both independent clauses, is classified as a correct sentencoid. This is because the error has created the sentencoid that is present in other sentences within the available language material:

Show him how very clever he is!

s1: Show him how very
s2: very clever he is

This illustrates how very clever he is.

s1: This illustrates how very
s2: very clever he is

However, the sentence 75 is incorrect. To correct this error, a comma, and the coordinating conjunction “but” will be placed:

(76) My cat is very clever, but he is always hungry.

The analysis of the sentence 76 sentence will take place:

(76) My cat is very clever, but he is always hungry.

76.1 Type: TLS ($8 \leq W \leq 14$)

76.2 TLS → divide into sentencoids

76.3 The list of the created sentencoids

76.3.1 sentencoid 1: My cat is very
76.3.2 sentencoid 2: very clever
76.3.3 sentencoid 3: but
76.3.4 sentencoid 4: he is always hungry

RW: very

76.4 Analyses of the created sentencoids

76.4.1 sentencoid 1

76.4.1.1 “My cat is very” = OK ($R \geq 3$)

76.4.2 sentencoid 2

76.4.2.1 “very clever” = OK ($R \geq 3$)

76.4.3 sentencoid 3

76.4.3.1 (1) = OK

76.4.4 sentencoid 4

76.4.4.1 “he is always hungry” = OK ($R \geq 3$)

76.5 RW > 0 → Advanced language addition

76.6 The addition of the created sentencoids:

My cat is very + very clever + , + but + he is always hungry.

My cat is very + ~~very~~ clever + , + but + he is always hungry.

My cat is very clever, but he is always hungry.

The sentence 76 is grammatically correct.

2. The case of the comma splice error

The term “comma splice” is used to name a situation, in which only a comma is placed between two independent clauses. The sentence 77 illustrates the error:

(77) Robert loves books, he almost lives in the library.

77.1 Type: TLS ($8 \leq W \leq 14$)

77.2 TLS → divide into sentencoids

77.3 The list of the created sentencoids

77.3.1 sentencoid 1: Robert loves books
77.3.2 sentencoid 2: he almost lives in
77.3.3 sentencoid 3: in the library

RW: in

77.4 Analyses of the created sentencoids

77.4.1 sentencoid 1

77.4.1.1 “Robert loves books”= OK ($R \geq 3$)

77.4.2 sentencoid 2

77.4.2.1 “he almost lives in” = OK ($R \geq 3$)

77.4.3 sentencoid 3

77.4.3.1 “in the library” = OK ($R \geq 3$)

77.5 RW > 0 → Advanced language addition

77.6 The addition of the created sentencoids:

Robert loves books + , + he almost lives in + in the library.

Robert loves books + , + he almost lives in + ~~in~~ the library.

Robert loves books, he almost lives in the library.

The analysis may suggest that the sentence 77 is grammatically correct. However, the sentence is incorrect, because only a comma was used to join two independent clauses. To correct the error, the comma will be replaced by a semicolon:

(78) Robert loves books; he almost lives in the library.

78.1 Type: TLS ($8 \leq W \leq 14$)

78.2 TLS → divide into sentencoids

78.3 The list of the created sentencoids

78.3.1 sentencoid 1: Robert loves books
78.3.2 sentencoid 2: he almost lives in
78.3.3 sentencoid 3: in the library

RW: in

78.4 Analyses of the created sentencoids

78.4.1 sentencoid 1

78.4.1.1 “Robert loves books”= OK ($R \geq 3$)

78.4.2 sentencoid 2

78.4.2.1 “he almost lives in” = OK ($R \geq 3$)

78.4.3 sentencoid 3

78.4.3.1 “in the library” = OK ($R \geq 3$)

78.5 RW > 0 → Advanced language addition

78.6 The addition of the created sentencoids:

Robert loves books + ; + he almost lives in + in the library.

Robert loves books + ; + he almost lives in + ~~in~~ the library.

Robert loves books; he almost lives in the library.

The sentence 78 is grammatically correct.

4.4 The theory of sentencoids—a tool in the process of learning English

For beginners, it may be more interesting to work on already grammatically correct sentences instead of struggling to create them. This approach, however, has some limitations: the available language material, which serves as a set of data, must be written in accordance with the rules that govern the grammar of the English language. The role of these rules can be explained as follows—the theory of sentencoids uses the "final products" of the grammar of the English language without referring to its basic terms, like: a clause, phrase, compound–complex sentence, etc.

4.5 Suggestions for the future research

The subsection proposes the following areas for future work:

1) the description of other ways that allow to morph countable nouns into uncountable nouns and vice–versa,

2) describe how the morphing of different types of adjectives influences the rest of a sentence,

3) the further exploration of the English language.

4.6 Index of terms

4.7 Contact the author

You can contact the author at:

jhawrot.english@poczta.fm