Tigim

Overview

Who are we?

Tigim is an Irish company developing a linguistic analysis tool (beta) to measure the accessibility of language and content used by companies in seconds. We use a data-driven approach to demonstrate where and how there can be disconnect with customers who have varying cognitive, literacy and language needs.

The name, Tigim, is a play on an Irish phrase that means 'I understand'.

Mission

It is our mission to give a voice to the unheard so that everyone has the opportunity to access products and services without restrictions.

Goal

To be the data-driven solution that supports companies on their path to achieving a communication style that is understandable and accessible to all.

How we do it

We believe the key to knowing if information will be understood is to focus on (1) the audience and (2) the purpose of the content.

In our analysis, companies set out:

> 1: The profile of the audience and their abilities

> 2. The targets they want to achieve based on the audience and purpose of document

This allows us to measure if the content is at an acceptable level of understanding for the end user.

Anything that will impact on understanding, based on the profiles given, will be flagged with suggestions on how to improve it.

The audience profile and targets can be customised for different types of content and end users.

Tigim

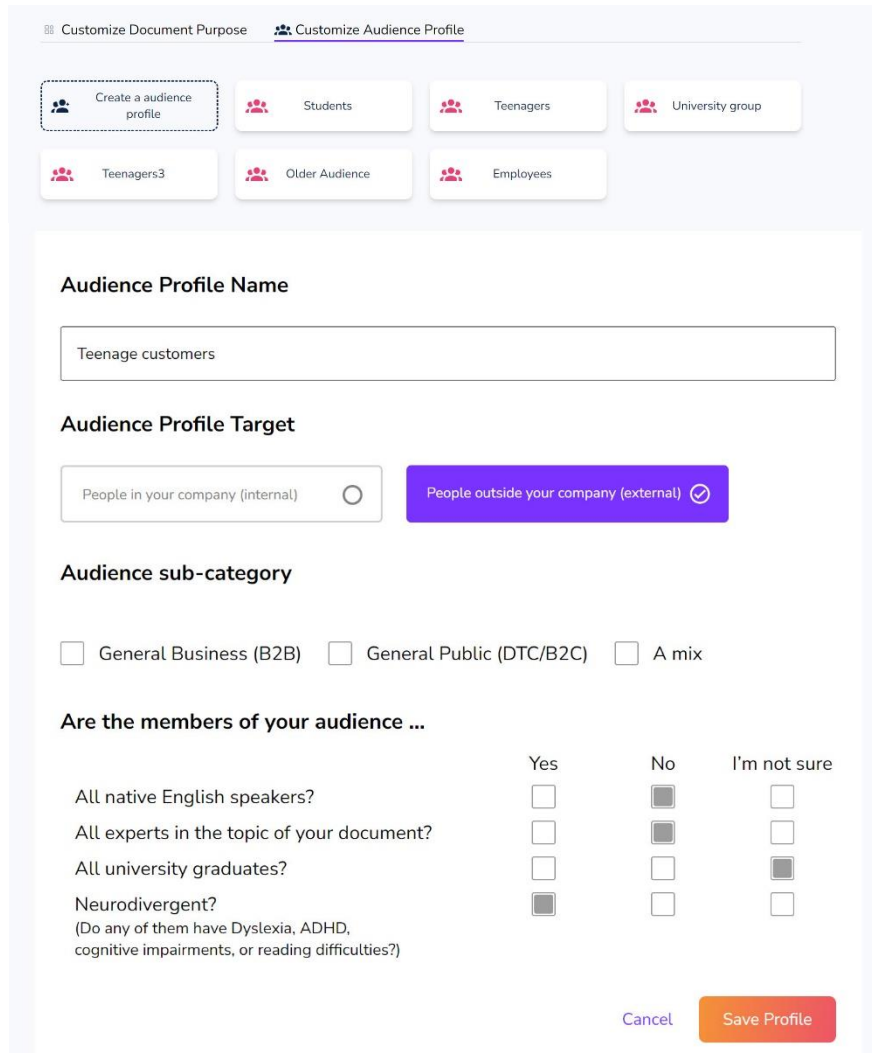**What you can do**

Features coming soon

- ✓ Abstract vs literal language
- ✓ Double negatives
- ✓ Metaphors and Idioms
- ✓ Adding own word watch lists

![Tigim logo]

1. Create an audience profile

Audience profiles can be created, stored and edited for easy access and customisation.

Customising an Audience Profile

Suggestions given:

Based on the profile and assumed abilities of the group (as it may not always be possible to know), we suggest a level of difficulty and English to aim for.

Suggestions given on what level to aim for based on audience profile

2. Customise targets

Customising targets for the purpose of the document and audience

3. Measure Targets

When the audience profile and targets are set, content can be added to analyse. The results are returned in seconds.

Difficulty rating	Language level	Reading time	Difficult Words
5 / 6	C1 Advanced	Less than 1 minute	8 Words 10.26%
			Difficult Sentences
			4 Sentences 57.14%
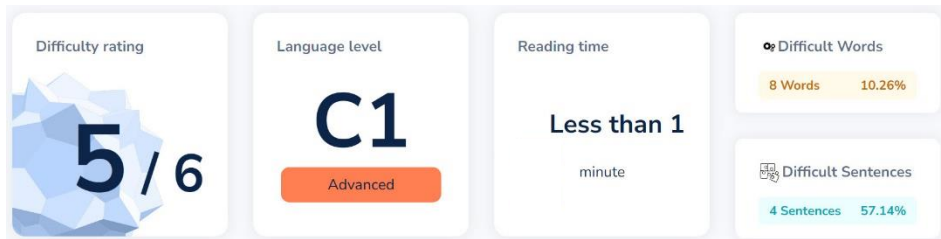
Results | Compliments | Recommendations 17 | Difficult words | Difficult sentences

	Feature	Target	Score	Result	Value
	Processing level	Low	Moderate	Higher than expected	22 pts
	Readability level	High	Low	Lower than expected	37.10 pts
	Word Complexity level	Low	Moderate	Higher than expected	5 pts
	Use of Jargon	Low	High	Higher than expected	7.29%
	Diversity of words	Low	Low	Appropriate	88 pts
	Engagement level	High	Low	Lower than expected	3 pts
	Empathy level	Moderate/High	Low	Lower than expected	0.00%
	Formality level	Low	High	More formal than expected	71 pts
	Directness level	High	High	Appropriate	0.60
	Sentiment level	High	Moderate	Lower than expected	7.81%
	Sentiment type	Positive	Positive	Appropriate	+0.22 pts
	Style	Spoken	Academic	More formal than expected	

*Results showing what targets have been achieved and what needs attention*

4. Measure Word Levels

4.1 English Level

The Common European Frameowrk of Reference (CEFR) is used to categorise and measure the level of words in the English language.

These are categorised based on what level words would be learned by learners of English. Not all words have been categorised as they are likely to be subject or industry-specific words that learners would not engage with in an educating setting. This means that for non-native English speakers, the chances of understanding words at an advanced level are reduced if they have not achieved fluency.

Beginner level words start at A1, moving to more advanced words that would be known by fluent speakers at a C2 level.
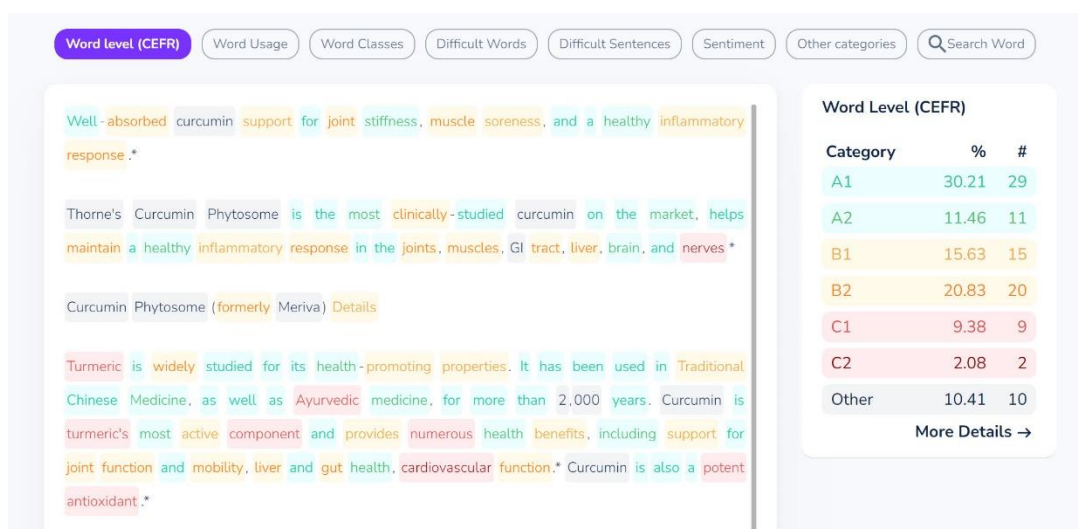
Categorising words based on the Common European Framework of Reference (CEFR)

The 'Other' category includes words that are no classified, jargon, acronyms, numbers and proper names. These are further analysed and segmented.
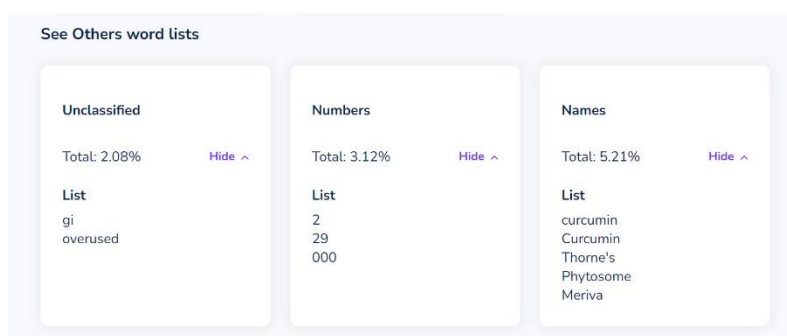
4.2 Word Frequency Level

This categorises words based on how often they are used in everyday life. Looking at the frequency level of a word tells us how difficult that word would be to understand. Words that are commonly used (or most frequent) are assumed to be easier as people would encounter them more often. Less frequent words have a higher degree of difficulty. There are 25 ranges that are based on the classification of words from large language databases; the BNC (British National Corpus) and the COCA (Corpus of Contemporary American English). We have classified them into 6 levels:

- Level 1: contains the 2,000 most frequent word families in the English language that are likely to be used every day, and so it continues with each level.
- Level 2: contains the 2,001 – 4,000 most frequent word families.
- Level 3: contains the 4,001 – 6,000 most frequent word families.
- Level 4: contains the 6,001 – 8,000 most frequent word families.
- Level 5: contains the 8,001 – 10,000 most frequent word families.
- Level 6: contains the 10,001 – 25,000 less frequent word families.

Reference: Nation, I.S.P. (2017). The BNC/COCA Level 6 word family lists (Version 1.0.0)
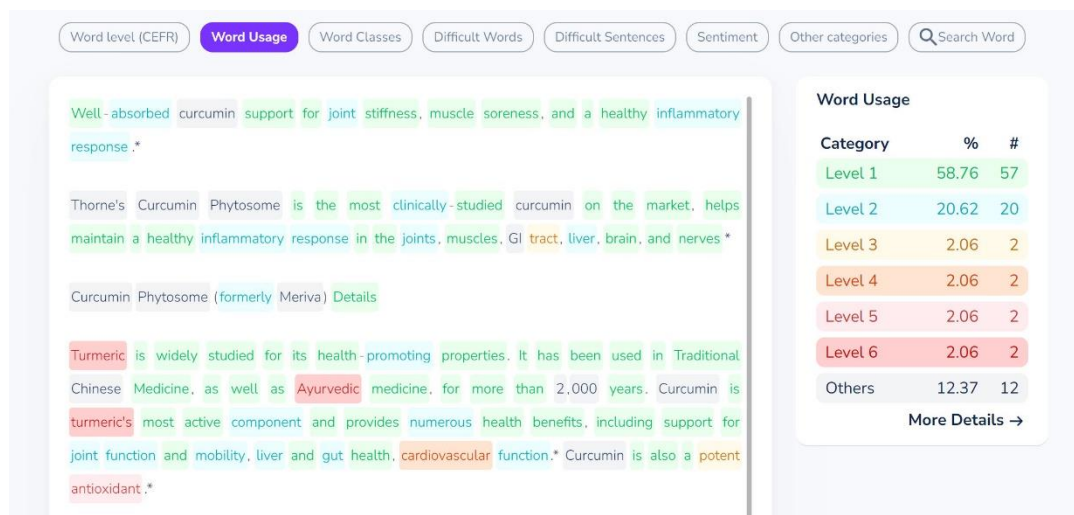
Categorising words based on their frequency of use. Level 1 contains the 2,000 most common word families.

Note: The 'Others' category contains words that have not been classified the National Corpora. This includes numbers, acronyms, abbreviations and words that are industry or subject-specific. We use our own modelling to categorise words that have not been included in the corpora.
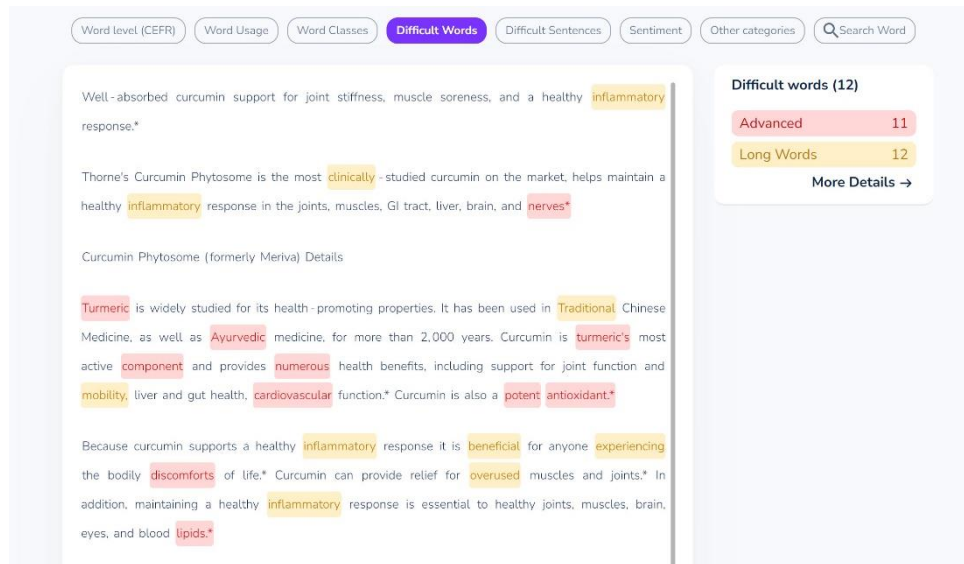
4.3 Difficult Words

Words can add difficulty to a text based on being advanced level words or containing more than 3 syllables.

Having a text with a high proportion of long words (with many syllables) impacts readability and cognitive load.

We show all words that can make the text more challenging for readers with different levels of literacy.

It is possible for words to be both long and advanced.

Highlighting words that can cause difficulty based on level of English and number of syllables.

For each of these words, we make suggestions to use easier words by showing alternative words that are more commonly used.

Providing alternative suggestions to use more commonly used words

4.4 Acronyms/Abbreviations

Any detected acronyms and abbreviations are detected and stored in 'Other'. These can be checked to ensure there is an explanation for each.

4.5 Context-specific/Jargon words

Words that have not been categorised are detected as industry or subject-specific terminology.

An opportunity is given to inform the system of analysis that these words are not jargon if required for a particular industry. On subsequent analyses, they will not be detected as jargon.
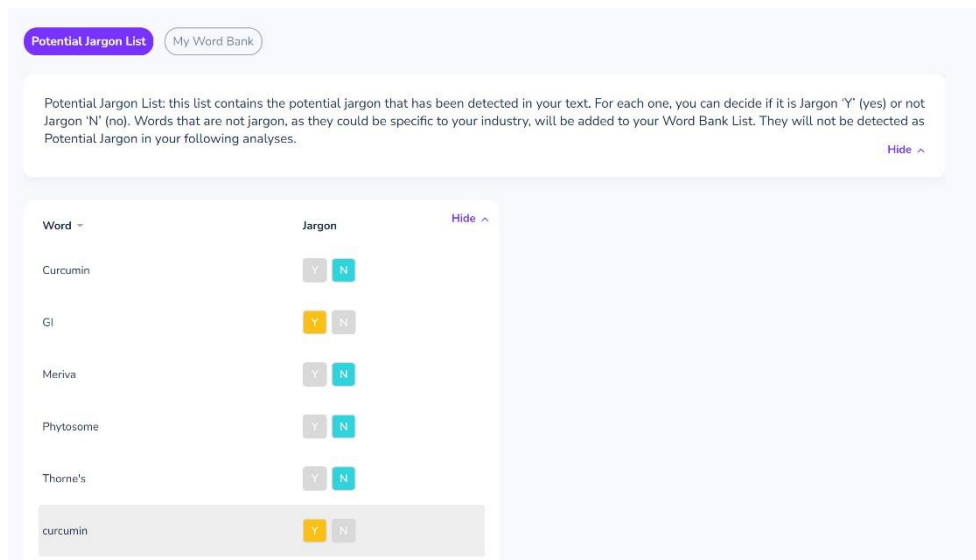
Informing the system on what words can be considered as jargon or not
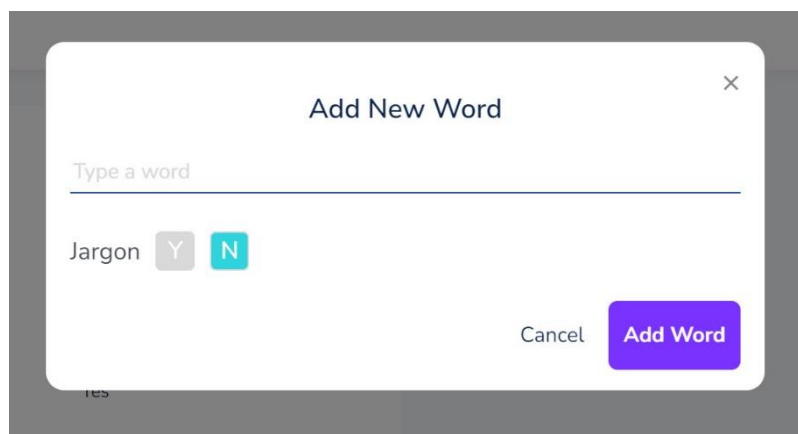
Adding company-specific terminology that should not be classed as jargon

Companies can create their own word lists with necessary words for their industry.

Building industry-specific words lists

5. Measure Sentence Structures

5.1 Difficult sentences

We detect 5 different types of grammatical structures that can increase cognitive load.
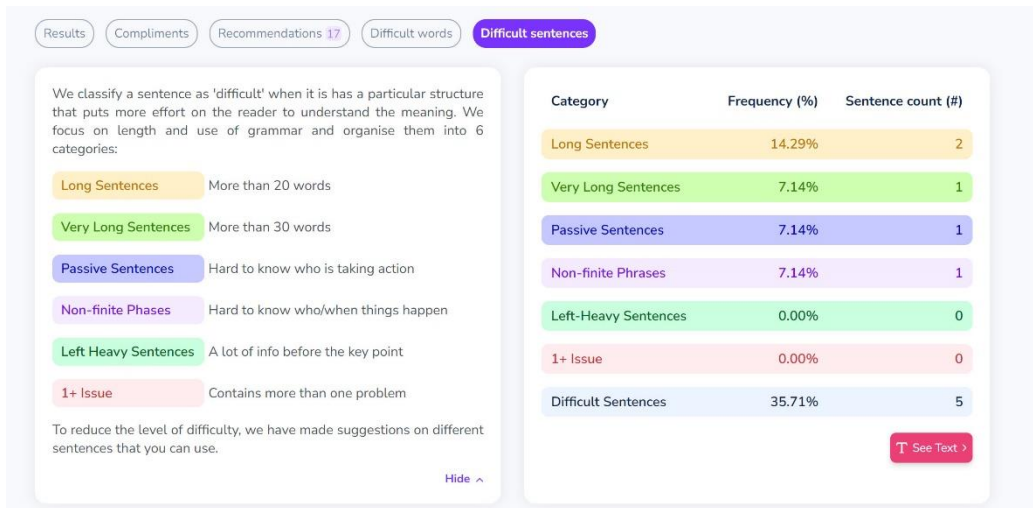
Categorising the frequency of difficult sentences in a text

Each sentence is highlighted in the text with suggestions on how to improve it given in the recommendation section.
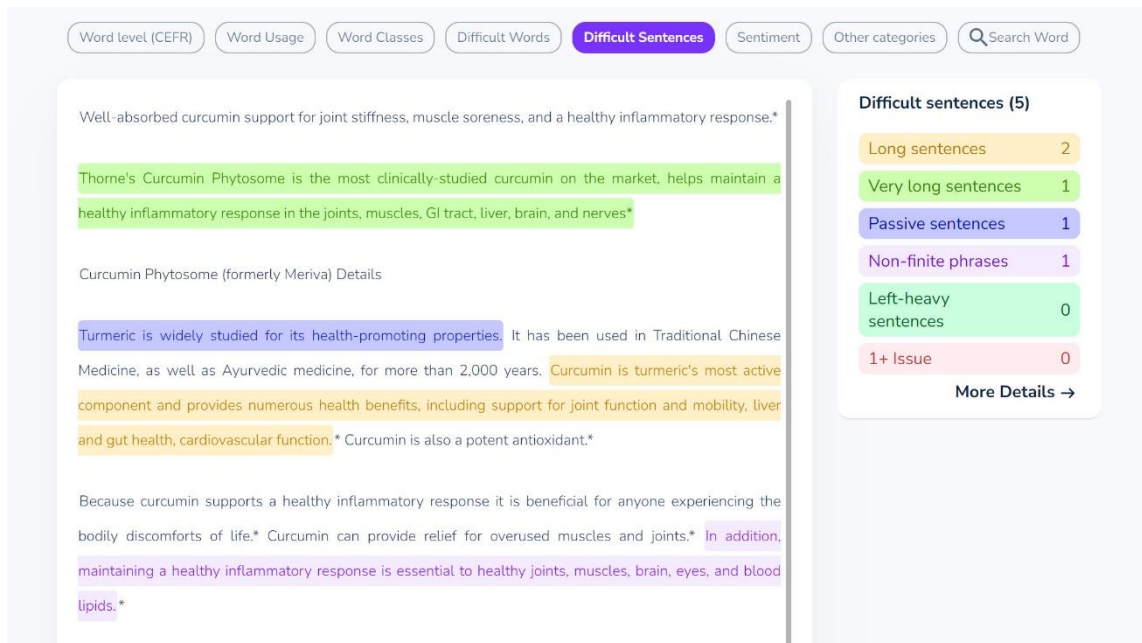
Highlighting the difficult sentences within the text

5.2 Use of Tense

We look at how verb tenses are used and their frequency. This can used to show how much of the text refers to past, present or future events.

This category is called 'Temporal Focus'. It can have an impact on the audience if the best tense is not used to achieve the purpose of the text.
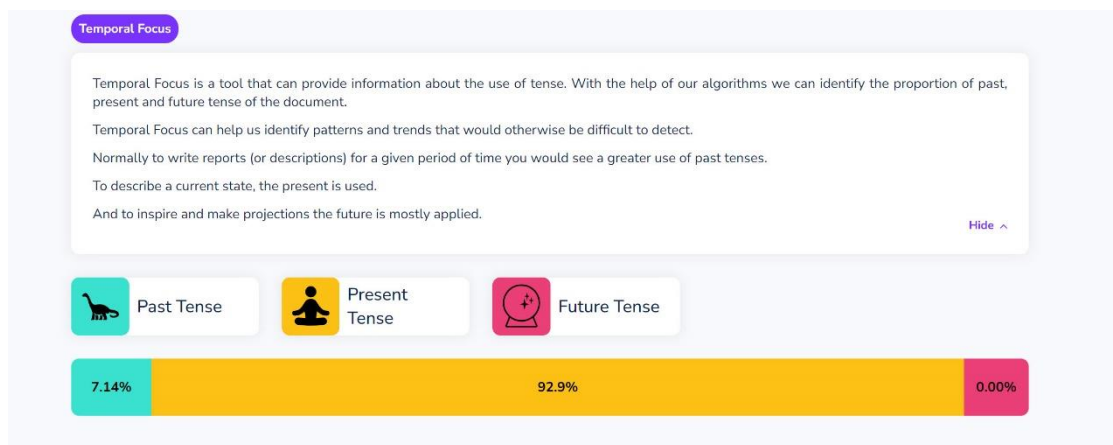
The proportion of verb tenses signalling events in the past, present and future.

6. Measure Structural Composition

Looking at the structure of the text.

Category	Frequency
Characters	4155
Letters	3396
Words (tokens)	577
Unique Words	292
Lexemes (unique base forms)	289
Sentences	88
T-units	93
Clauses	113
Dependent clauses	17
Complex noun phrases	118
Bullet points	0
Headings	2
Paragraphs	5

per Word

Category	Frequency
Letters	5.89
Syllables	2.05

per Sentence

Category	Frequency
Words (tokens)	6.56
T-units	1.06
Clauses	1.28

Analysing the structural composition of the text

7. Compare Documents

After implementing any changes to make the text easier to understand for the audience, comparisons can be made in seconds.

This can help to see how things have changed in a quantifiable way. As well as highlight what additional improvements need to be made to meet targets.

Document Comparison

Before Text 59 ⇄ **After** Text 58

Feature	Before	After	Change	Comments
Difficulty rating	5 / 6	4 / 6	⬇ 1 level	Awesome! You just made it 1 level easier by lowering your difficulty ranking!
English level (CEFR)	C1	B2	⬇ 1 level	Great! Your English text level just got 1 level easier.
Reading time	2 min	1 min	⬇ 1 min	Fantastic work! You've cut down your document reading time by 1 min!

Brain processing

Feature	Before	After	Change	Target
Processing level	Moderate (23)	Low (15)	⬇ 34.8%	✓ Low. Congratulations, your score its on target. ▾
Noun/Verb ratio	High (1.74)	High (2.34)	⬆ 34.9%	⚠ Low. Your score is too high, try to get it lower.
Exclusion Word	High (18.0%)	Low (0.00%)	⬇ 100%	✓ Low. Congratulations, your score its on target.
Long Sentences	Low (5)	Low (1)	⬇ 80.0%	✓ Low. Congratulations, your score its on target.
Very Long Sentences	Low (0)	Low (0)	⊝ No	✓ Low. Congratulations, your score its on target.
Passive Sentences	Low (4)	Low (0)	⬇ 100%	✓ Low. Congratulations, your score its on target.
Non-finite Phrases	Low (9)	Low (2)	⬇ 77.8%	✓ Low. Congratulations, your score its on target.
Left-heavy Sentences	Low (1)	Low (1)	⊝ No	✓ Low. Congratulations, your score its on target.
+1 Issue	Low (4)	Low (1)	⬇ 75.0%	✓ Low. Congratulations, your score its on target.
Difficult Sentences	Low (13)	Low (3)	⬇ 76.9%	✓ Low. Congratulations, your score its on target.
Readability level	Low (20.9)	Moderate (57.7)	⬆ 176%	❗ High. Try to higher your score a little bit more. ▾
Avg. sentence length	Low (6.56)	Low (7.73)	⬆ 17.9%	✓ Low. Congratulations, your score its on target.
Total Sentences	88	37		

Comparing two documents to see changes made relative to targets

13

Tigim

We are constantly working to add features that will help to measure if content is understandable in an objective and quantifiable way.

At present, we analyse English. However, we will extend this level of analysis into other languages.

We are always open to feedback and ideas on how to improve.

Our goal is to be the data-driven solution that supports companies on their path to achieving a communication style that is understandable and accessible to all.

Features coming soon

- ✓ Abstract vs literal language
- ✓ Double negatives
- ✓ Metaphors and Idioms
- ✓ Adding own word watch lists

Thank you.

If you have any questions or suggestions, please feel free to me, Niamh Kelly, CEO, at Tigim.

I would love to talk and share ideas.

Niamh Kelly

niamh@tigim.co