

A Stylistic Analysis of Effective Public Speech in TED Talks

Assistant Lecturer Haya Abdulsalam Jasim

Ibn Sina University of Medical and Pharmaceutical Sciences, Baghdad, Iraq

Haya.s.aljumaily@ibnsina.edu.iq

Assistant Lecturer Sarah Hatem Jaddoa

Open Educational College, Baghdad, Iraq.

Sarahatem367@gmail.com

دراسة أسلوبية للخطابة العامة التحفيزية في محاضرات TED

م. م. هيا عبد السلام جاسم

جامعة ابن سينا للعلوم الطبية والصيدلانية

م. م. سارة حاتم جدوع

المفتوحة

التربوية

الكلية

Abstract:

Given the growing importance of public speaking skills across real-life, professional, and educational contexts. The current study examines a particular type of effective public speech: TED Talks. It aims to analyse the lexical and syntactic levels and to identify the effect of choosing these elements in making public speech more effective. This study aims to determine whether gender affects the stylistic selection of these items. To perform its analysis, this study employs corpus stylistic analysis to identify the language patterns that make TED talks effective, following the corpus linguistics methodologies described by McIntyre and Walker (2019). It performs the stylistic analysis following Leech and Short's (2007) model. The following study analyses a balanced corpus of 30 popular TED Talks (15 speeches by males and 15 by females). Further, to identify the impact of gender, this study organises its data into two corpora: one dedicated to males and the other to females. It is observed that the language is characterised by some linguistic peculiarities: preference to shorter words (two to four letters), medium sentence length (14.93 words on average), and an equal lexical diversity (STTR = 39.65 per cent). Gender comparisons show that female speakers prefer nouns related to communication, emotive verbs, and possessive pronouns to create contact with others by sharing experience. In contrast, male speakers use much more concrete vocabulary, action verbs, and direct address to the audience. Irrespective of these variations, both genders employ the same methods to balance technical words and provide comprehensible phrases, as well as strategic pronouns to improve audience appeal. The results of such a study add to a perception of the linguistic background of skilful speaking in a crowd and can be applied to communications training, speech training, and teaching speaking in a crowd. Keywords: public speech, corpus stylistics, TED Talks, lexical features, motivational discourse

المستخلص :

تتناول هذه الورقة البحثية العناصر الأسلوبية للخطابة العامة المؤثرة من خلال تحليل مدونة مختارة من محاضرات TED على المستويين المعجمي والنحوي، وكيفية إسهام هذه العناصر في نمط الخطابة التحفيزية، وما إذا كانت الفروق بين الجنسين تؤثر على نمط الخطابة التحفيزية. ونظرًا للأهمية المتزايدة لمهارات الخطابة العامة في المجالات المهنية والتعليمية والحياتية، تُوظف هذه الدراسة علم تحليل المدونات اللغوية لتحديد الأنماط اللغوية التي تجعل محاضرات TED مؤثرة، مما يُوفر معلومات عملية للمتحدثين والمعلمين والمتخصصين في مجال التواصل. بالاعتماد على النموذج الأسلوبية للبيتش وشورت (٢٠٠٧) ومنهجيات علم اللغة الحاسوبي التي وصفها ماكنتاير ووالكر (٢٠١٩)، يحلل هذا البحث مدونة متوازنة

تضم ثلاثين محاضرة من محاضرات TED الشهيرة (خمس عشرة محاضرة ألقتها نساء وخمس عشرة محاضرة ألقها رجال)، ويبلغ مجموع كلماتها حوالي ستين ألف كلمة. لاحظ أن اللغة تتميز ببعض الخصائص اللغوية: تقصيل الكلمات القصيرة (من حرفين إلى أربعة أحرف)، ومتوسط طول الجملة (١٤.٩٣ كلمة في المتوسط)، وتنوع معجمي متساوٍ (نسبة الكلمات القصيرة إلى الطويلة = ٣٩.٦٥٪). تُظهر المقارنات بين الجنسين أن المتحدثات النساء يفضلن الأسماء المرتبطة بالتواصل، والأفعال العاطفية، وضمان الملكية للتواصل فيما بينهن من خلال تبادل الخبرات، بينما يستخدم المتحدثون الرجال مفردات أكثر واقعية، وأفعالاً حركية، ويخاطبون الجمهور مباشرة. بغض النظر عن هذه الاختلافات، يستخدم كلا الجنسين الأساليب نفسها في الموازنة بين المصطلحات التقنية والعبارة الواضحة، بالإضافة إلى استخدام الضمائر المناسبة لتعزيز جاذبية المتحدث للجمهور. تُسهم نتائج هذه الدراسة في فهم الخلفية اللغوية للخطابة الماهرة أمام الجمهور، ويمكن تطبيقها في تدريب مهارات التواصل، وتدريب الخطابة، وتعليم فن الخطابة أمام الجمهور. الكلمات المفتاحية: الخطابة العامة، أسلوبية المدونات اللغوية، محاضرات TED، السمات المعجمية، الخطاب التحفيزي

Introduction:

Language is a communication medium in a variety of forms, both written and spoken, with numerous uses; among them is public speaking. As a result, it may be argued that public speaking is the process of sharing information with an audience through language. There are three categories of public speaking: informative, persuasive, and special-purpose. Informative speech is designed to inform listeners about a specific subject and enrich their knowledge. Meanwhile, the persuasive speech aims to influence the audience's thoughts and perceptions regarding a given topic. The third type, commonly known as special or purpose speech, including a motivational address and being the main focus of the present research, is intended to arouse emotion and encourage the audience to change an undesirable situation (Hollingsworth et al., 2021). Various strategies and linguistic options should be considered to deliver a successful speech. The study of people's speech is of interest to scholars across various fields as a means of identifying linguistic preferences that make this speech effective. In that regard, this paper will take a stylistic approach to the analysis of public speech. As an area of applied linguistics, stylistics is mainly concerned with the style of a given text. According to Leech and Short (2007), style is defined as "the way in which language is used in a given context, by a given person, for a given purpose and so on." (p.9). They regard the style as "the dress of thought." (p.13), the way a particular person expresses his/her thoughts differently from others by choosing specific words, expressions, morphological and syntactic constructions. The field's versatility in embracing many approaches makes it especially appealing for this study. The style of a particular discourse, be it a written or spoken one, is determined by the repetitive decisions that the speaker makes concerning the application of specific strategies or linguistic devices. According to Ghazala (2011), the style is characterized as a language choice made by a specific author under the constraints and possibilities of the language/grammar, i.e., within the overall choices in the syntactic, semantic, phonological, and pragmatic systems (p. 41). Stylistics and linguistics are closely related, as Norgaard et al. (2010) demonstrate; they draw on models and theories to offer a comprehensive understanding of style in a given text. The main issue in stylistics research is defining how and why certain features are presented in a text, and estimating the linguistic characteristics that explain their use. Therefore, to determine the linguistic stylistic characteristics of public speech, especially effective and motivational speech, the present study uses TED Talks as its data. TED, an acronym for Technology, Entertainment, and Design, is a non-profit organization founded in 1984 to host a conference focused on these three areas. It can be accessed at ted.com. This platform covers a wide range of topics. These subjects are divided into sets of A to Z, each with relevant discussions of the corresponding topic. These topics can be filtered by duration, topic, date, and relevance, and the majority of the content is readily available free of charge, making it easy for visitors and researchers to access. The main target of the talks is the people who have undergone any form of stress or hardship and need inspiration or motivation. This is why the talkers are supposed to deliver their talks in a positive, inspirational, and motivational way, and to use their abilities and language effectively to convince the listeners and encourage them to think differently. Therefore, this paper will examine the linguistic and stylistic characteristics of effective oratory, using TED Talks as an example. Corpora are used to conduct such an analysis; as Jones and Waller (2015) state, a corpus is nothing more than a relatively electronically stored, searchable collection of texts. These are written or spoken texts that might vary in length, though, as a rule, will be longer than a single speaking turn or a single written sentence (p. 5) Those studies, which involve managing a large volume of information and keeping and analyzing it electronically, are considered corpus studies (Paltridge, 2012, p. 142). The benefits of corpus use, as McEneaney

et al. (2006) illustrate, include the ability to efficiently manage and organize data, time and effort savings, more accurate and precise data, and the avoidance of errors (p. 6). Corpus studies are considered the synthesis of both quantitative and qualitative analysis; they not only present quantitative results but also explain these patterns qualitatively. Therefore, this paper aims to establish the linguistic stylistic peculiarities of effective oratory using the TED corpus as an exemplar. The corpus software (WordSmith) is used in the analysis. Besides, since linguistic stylistic elements differ across genres, the study also performs a keyword analysis of the TED corpus, aiming to identify the most significant words speakers use most often as signs of this genre's aboutness (effective public speech). The paper will be valuable to researchers in the field of presentation, as it provides essential information on the linguistic and stylistic decisions used in public speech, which can be incredibly helpful in teaching students to deliver the speech.

Related Studies:

This section is further subdivided into three subsections, which respectively represent: studies intended to establish the stylistic properties of given texts; investigations of phenomena in public speech; and linguistic works that have used TED talks as their input. Every section identifies the contributions of previous researchers, thus serving as a basis for the current research. The general purpose of the present work is to use the stylistic approach in the analysis of good speech. Once the terms of style and stylistics are clarified, due to the corpus-based character of the current inquiry, it is relevant to mention the previous corpus-stylistics studies that aimed at extracting linguistic properties of a certain corpus or texts: A stylistic analysis was conducted of the linguistic options used by the female characters in three identified novels by Qais Ghanem, as presented in the work by Al-Nakeeb and Mufleh (2018). The data were structured as corpora, and the analysis was conducted using the corpus tool GraphColl, along with the theory of transitivity. The lexical and lexico-grammatical levels were analyzed quantitatively, whereas qualitative analysis was used in explicating concordance-line data. The results indicated that the female characters used specific adverbs and adjectives quite often. Moreover, the paper by Kang and Yu (2011) introduces a corpus-based stylistics study that examines the language used by tourism providers on their websites. The data were also placed as corpora and analyzed using the WordSmith and AntConc corpus tools. A quantitative and qualitative analysis was conducted, and the findings demonstrated the linguistic stylistic features that define tourism language, setting it apart from other genres. These, among other studies, indicate that a range of stylistic theories/tools can be used to determine the linguistic stylistic characteristics of a given text or corpus. In that regard, the authors believe that no previous study has been conducted to date to analyze effective oratory using the TED Talks corpus as a data source.

Public Speaking:

Having reviewed the literature on TED talks, it is evident that numerous methodologies, tools, and programs have been used to analyze public speech in its diverse forms, both generally and specifically in TED talks. The analysis of public speech has been shown to interest scholars across disciplines, including Kahl (2014). In his research, the author was interested in the significance of teaching public speaking in schools and identified a gap in pedagogical research. The scholar argued that there should be a consideration of every aspect of public speaking, such as content, body language, and teaching on the speech delivery, and not just on the context. The qualitative research focused on analyzing high school students' experiences and views of what their teachers taught them about public speech. The results highlighted the importance of giving adequate training to students in this respect, especially at the secondary level. However, another article by Yang (2019) conducted a linguistic analysis of 20 speech scripts from an English public-speaking competition and explored the most common rhetorical strategies (e.g., simile, metaphor, antithesis). Despite the small sample size limiting generalizability, the research revealed the importance of such gadgets and found that they made people more lively, captivating, and convincing in their speech. Besides, Agbo (2022) analyzed the linguistic elements of the Apostle Paul in his first letter, which is an exhortatory public speech intended to strengthen the community in the face of adversity, and applied Bitzer's rhetorical theory, focusing on the concept of the so-called Logos as a persuasion tool. These findings revealed that rhetorical theory is useful for analyzing persuasion in oral speech. Another study by Tan (2023) aimed to develop a theoretical framework for improving public-speaking instruction through memetics, a theory of mass transmission. Through a qualitative analysis, this paper concluded that applying this theory to the art of teaching public speaking would increase students' awareness of the significance of understanding their audiences' cultural backgrounds, thereby fostering interactive, engaging public speech. In turn, this subsection shows that researchers have analyzed public speech in various ways. In particular, the next studies dealt with the examples of TED talks as representative data on the speech in society:

TED Talks:

Di Carlo (2014) studied how the rhetorical devices (Ethos, Pathos, and Logos) are used by the same speaker, Shereen El-Feki, in two contexts (one a TED talk and the other an academic symposium). It was found that there was a prevalence of Ethos and Pathos, likely intended to make an impression on the audience and to arouse their feelings. The findings of this study cannot be generalized because it was conducted with only one speaker. However, it gave a point of departure in studying how rhetoric is employed in general speech. The study of the application of the rhetorical strategy Pathos (appeal to emotions) was conducted by Di Carlo (2015) through a qualitative analysis of Dr. Jill Bolte Taylor's TED talk. The focus on a single talk was insufficient to achieve the study's outcomes; however, it was possible to learn that speakers resorted to personal, emotional, and evocative language, which made the speech more interactive, intimate, and memorable. In a similar vein, Alnasser (2023) investigated persuasive language techniques in public speeches from Saudi TEDx talks. Using the Political Discourse Analysis approach as a framing theory and the three rhetorical devices (Pathos, Ethos, and Logos), the analysis was conducted across 10 videos. The findings revealed that Saudi speakers commonly use Pathos, which aligns with social conventions that emphasize family, faith, and togetherness. However, they are cautious when promoting ideas that could bring about societal transformation. Further, Alnasser (2023) applied a political discourse analysis approach to analyse the persuasive language in Saudi TEDx talks. The analysis was conducted using the three rhetorical devices (Pathos, Ethos, and Logos) across ten selected videos. The result of such analysis confirmed the social conventions about the Saudi community by emphasising faith, family, togetherness, and being cautious or avoiding dealing with ideas that encourage societal transformation. The research paper is important to the literature because it demonstrates that persuasive strategies vary according to the speaker's ideology. Some of its weaknesses are a limited sample size and the ideological and cultural influences of the speakers. Salahuddin et al. (2023) applied the TED Talks Corpus to research the use of figurative language (e.g., similes and as) in the speech of the population, using both quantitative analysis and qualitative methods. The findings showed that speakers used these figurative devices more often than others to explain and highlight their words. This research advanced literature on the analysis of figurative language by demonstrating that these techniques are not only used in literary contexts but can also be employed in speech to engage and persuade audiences. In addition to articles commenting on rhetorical language, Salahuddin et al. (2023) conducted a stylistic review of TED Talks. The current research used the Influential English TED Talks Corpus as a data source and corpus tools and stylistics theories to analyse it. The analysis indicated that speakers in TED talks employ various devices, such as metaphors and similes, to make their speeches sound more engaging. Besides, the speakers employed both formal and semi-formal styles to emphasise the same purpose. Meanwhile, Marchenko and Minenko (2020), analysed TED talks by focusing on linguistic and extralinguistic aspects. They examined the data through qualitative analysis of selected transcriptions and video recordings, using the classification offered by D.M. Kaysina. Their results showed that they use a mix of linguistic and extralinguistic gestures and tone of voice to achieve better speech. Other research has also helped in the field of public speaking by clarifying the use of certain grammatical features. The study by Wang (2012) compared the use of word patterns in short and long talks, referring to the TED corpus as an example of a short talk and the BASE corpus as an example of a long talk. The study found that the difference in pronoun use between the two corpora was significant, as demonstrated through quantitative analysis using corpus tools (AntConc, RANGE, and KF-Ngram) and further supported by qualitative analysis. Besides, the grammar items that were shared between the two corpora were mostly used by the speakers. Di Carlo (2015) was more interested in analysing the use of subjective adjectives in TED talks using a mixed-methods approach. The result indicated that more vivid and descriptive adjectives were employed by the speakers to ensure the interesting and the involvement of the audiences. Azlia (2022) examined how male and female speakers differ in their use of discourse markers in public speeches, drawing on TED Talk data. Using 49 transcribed videos from the motivational and personal-growth sections, the researcher conducted quantitative and qualitative research in accordance with Hyland's framework. The findings suggested that female speakers used engagement discourse markers more widely than male speakers, and that the latter discourse was more direct and focused on markers. Even though some past investigations viewed TED talks as generalizations of public speech, other studies analyzed particular types of talk. Altamimi (2021) examined the linguistic processes in terrorist-related negotiations to analyze the specialized discourse of this field. Similarly, Mattiello (2019) analyzed 49 TED Talks focused on cancer. In both qualitative and quantitative analyses, the same study found that TED speakers used specific linguistic strategies to make cancer-related content more engaging, more accessible, and more positively presented. In

conclusion, it remains apparent that the majority of the studies mentioned above that have explored TED talks have focused on rhetorical and linguistic analysis. Even though stylistic analysis is used, no work has comprehensively explored the linguistic and stylistic components of this genre. This current study aims to determine the linguistic and stylistic aspects of public speech at the lexical and syntactic levels and to examine their role in the effectiveness of the speech. It also aims to examine the effect of gender in forming this kind of speech.

Methodology:

This paper examines the language style of the chosen texts using models of stylistic analysis grounded in computational linguistics. The analysis is performed using Leech and Short's (2007) model to examine the speeches' style features. It is complemented by the methodology developed by McIntyre and Walker (2019) for analyzing texts using corpus analysis tools. This analysis is carried out based on the following:

1. Theoretical framework of analysis. The model used in the analysis is based on the analysis of Leech and Short (2007) that separates style into four major types, two of which are used in the current analysis:

- Lexical Categories: Dynamics of Vocabulary used, complexity, and frequency.
- Grammatical Categories: The sentence structure and the grammatical structure used are analyzed.
- McIntyre and Walker's (2019) approach supports the analysis by providing clear steps for using analytical software to analyse the text.

2. Data collection: the selected speeches were chosen following certain characteristics, including the type of speech, situation and the number of views (the most popular talks).

3. Text preparation: digital preparation was made to organise the data and present a readable text that can be analysed using corpus tools.

4. Corpus tools: This paper employed WordSmith software to perform the corpus analysis, including the word and sentence length, the mean number of words per sentence, which helps in identifying the linguistic complexity and readability of the text. This software presents the Token Type Ratio (TTR) analysis, which is used to assess lexical diversity in texts and identify word frequency. Further, this study performs the Semantic word distribution analysis to the data, which means categorization of words based on semantic type: noun, verb, adjective, adverb, depending on the list of Leech and Short (2007).

5. Keyword analysis: Recognition of the most common words through comparing texts via WordSmith by using a statistical method of keyword analysis.

- Application of Log-Likelihood (LL) and Bayesian Information Criterion (BIC): Returning the words that are more or less frequent than expected to be. This could be done by using the approximate Bayes factors BIC as a measure for the effect size, at 6.000 degrees with the associated p-value 0.00001 and LL 15.13 to present strong evidence against the null hypothesis, which states that the difference in frequency of a particular item between the two corpora occurs by chance. In addition to using log-ratio to reduce the number of items in the keyword lists, any item with a log-ratio of 2 or higher is considered a key.

Keyword classification:

- Positive words: These are words that have more frequency and show the unique characteristics of style.
- Negative words: Less common Words, which help to detect the stylistic variations between texts.

6. Syntactic and semantic relationship analysis: Syntactic pattern analysis: The patterns of pronouns, verbs, and nouns in language are studied to know the complexity of language structures.

Data Selection

This paper considers the linguistic and stylistic characteristics of TED talks using corpus-based analysis. The official TED.com site was used to gather data based on specific criteria to create a balanced corpus.

Selection Criteria

The following parameters were used to select TED Talks:

- Themes: Themes chosen were in the humanities and the creative areas (Beauty, Art, Language, Storytelling, Media, Creativity, Performance, and Literature). This subject matter allows meaningful comparisons among the speakers while maintaining the subject area's cohesiveness.
- Length: 12 to 18 minutes, which is the average length of TED talks.
- Language: English alone, to ensure that the language remains homogeneous and that translations do not have any influence on the stylistic elements.
- Gender Balance: 15 female and 15 male speakers equally represented.

▪ Popularity: Favoritism of the extensively viewed talks that have already proved to resonate with the mass audience, as represented by the number of views on the TED site.

Based on these criteria, 30 TED Talks (15 by women and 15 by men) totaling about 70,975 words were chosen. The chosen talks were all delivered after 2006 and were relevant to modern-day public speaking practices. The official TED.com site provides transcripts; hence, these were manually collected and checked to implement. As TED Talks are openly distributed resources from speakers willing to be shared, their use is associated with very few ethical and legal issues. Under the TED company's terms of use, they can be used.

Compiling the Corpora

The data were organized into three major corpora (as represented in Figure 1).

- Combined TED talks Corpus (CTC): Full Collection of 30 transcripts.
- Female Speakers Corpus (FSC): 15 transcripts of female speakers.
- Male Speakers Corpus (MSC): 15 transcripts of male speakers.

Figure 1. Detailed information for the corpora used in this study

N	text file	file size	tokens (running words) in text
	Overall	777,062	141,950
1	TED corpus 30 talks.txt	388,533	70,975
2	TED corpus for Female talks .txt	198,206	35,713
3	TED corpus for Male talks.txt	190,323	35,262

Sampling and Annotation

The transcripts were all electronically copied and kept in plain-text (.txt) format, which would later be processed using corpus analysis programs. This entailed an extensive cleaning process to prepare the data for computation.

- Cleaning: Elimination of date and notes on the audience response, regularization of punctuation, and regular formatting.
- Annotation: Part-of-speech tagging using TagAnt 1.2 software (Anthony, 2015) and, hence, improvement of analytical functions.

The characteristics common to all corpora include being specialized in research, oral discourse expressed in writing, monologic, synchronic, and based on a similar temporal range.

Data Analysis and Findings

Words and Sentences Length

A study of word length in TED Talks, conducted with WordSmith, provides insights into the linguistic style used primarily by TED speakers. The shorter word length may be because speakers want to capture the audience as much as possible and use the least lexical content to reduce cognitive load during listening and create a more effective rhetorical effect in oral presentation. Besides, the large percentages of common and shorter words produce more memorable material. The small gender gap in the length of words means that the strategies of effective speaking as a speaker are not gender specific, as both male and female speakers focus on the availability of vocabulary to make their message sound approachable and promote the notion of the speaker using words of two to four letters in all the corpora. Figure 2. Shows the word length of the study corpora. Figure 2. Word Length in the Corpora used in this study

N	1-letter words	2-letter words	3-letter words	4-letter words	5-letter words	6-letter words	7-letter words	8-letter words	9-letter words	10-letter words	11-letter words	12-letter words	13-letter words	14-letter words	15-letter words	16-letter words
	7,180	25,080	29,900	28,788	16,496	11,358	8,862	5,504	3,902	2,534	1,188	602	354	84	80	30
1	3,590	12,540	14,950	14,394	8,248	5,679	4,431	2,752	1,951	1,267	594	301	177	42	40	15
2	1,692	6,380	7,251	7,052	4,171	2,914	2,279	1,447	1,056	758	358	195	116	22	15	4
3	1,898	6,160	7,699	7,342	4,077	2,765	2,152	1,305	895	509	236	106	61	20	25	11

Sentence Length Analysis

A look at sentence length shows patterns that are consistent across the TED Talk corpora. Overall, the Combined TED Talks Corpus contains 9,466 sentences, averaging 14.93 words per sentence. The length of sentences of the female speakers (15.62 words) is longer than that of the male speakers (14.29 words). These are lengths that would be in the normal readability range of Sanyal (2006), indicating that TED speakers manage the complexities/accessibility ratio well. The length of the sentence is moderate and serves multiple functions: it is

not too long; it is detailed enough to explain complex ideas; and it sounds like a conversation (Scotto di Carlo, 2014). The minimal gender differences suggests that both male and female speakers adhere to similar range of spoken comprehensibility, which is 12-18 words to retention maximum in oral presentations (Feng et al., 2010). Figure 3. Shows the mean and sentence length in the study corpora. Figure 3. Mean and sentence Length in words.

	N	sentences	mean (in words)	std.dev.
		9,466	14.93	11.81
1		4,733	14.93	11.81
2		2,276	15.62	11.21
3		2,457	14.29	12.32

Lexical Diversity

A part-of-speech analysis of TED Talks provides useful insights into their stylistic characteristics. TagAnt software and POS tagging reveal differences in patterns between the genders in the corpora. The Female Speakers Corpus has nouns (NN), personal pronouns (PP), adverbs (RB), and prepositions/subordinating conjunctions (IN) as its predominant parts of speech. Meanwhile, the male corpus employs nouns (NN) more frequently than any other part of speech, followed by determiners (DT), adverbs (RB), and adjectives (JJ). The analysis of the current study indicated that TED Talk speakers, regardless of gender, employ a nominal style of speech. Using this style helps them deliver their speech in a more understandable, memorable way by focusing on specific ideas and concepts. Further, female speakers prefer a conversational style when addressing the audience. The ratio of nominal elements and the use of first-person pronouns are considered by Biber and Conrad (2019) as signs of a persuasive attitude that combines informational messages with interpersonal communication. The two corpora show a high degree of adverb use, and thus, speakers attempt to qualify their words and make them stronger to create a more powerful rhetorical effect. This observation is similar to what has recently been observed in speakers of TED talks, where adverbial modification is habitually used by orators to make their arguments easier to remember and to highlight the most important points (see Scotto di Carlo, 2014). Although there are no significant differences between the two gender corpora in the overall pattern, it is evident that female speakers use more personal pronouns and adverbs than male speakers do. In contrast, the latter use more determiners and coordinating conjunctions, suggesting minor gender differences in communicative strategies. However, at the same time, they retain the main stylistic characteristics of effective oratorical practice as a whole. Figures 4. And 5 shows the POS of each corpus. Meanwhile, Appendix 1. Show the meaning of each abbreviation presented in these figures.

Figure 4. POS in Female Speaker Corpus

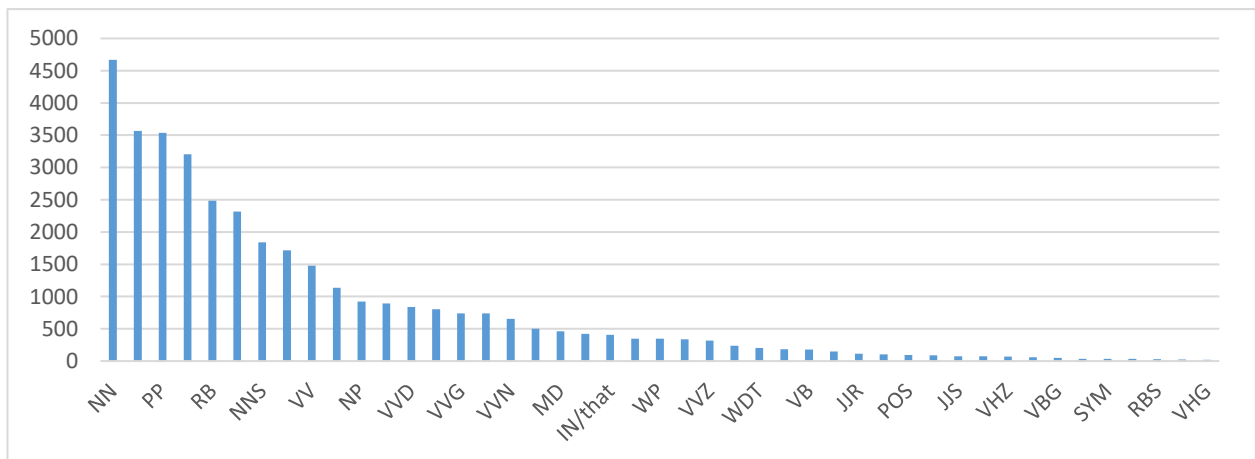


Figure 5. POS in Male Speaker Corpus



Token Type Ratio

The Type-Token Ratio (TTR) is useful for analyzing the lexical diversity of TED Talks and for understanding their vocabulary richness. This research has determined both standard TTR and Standardized TTR (STTR) in the three corpora using WordSmith 8.0. As shown in Figure 6, the merged TED Talks Corpus has a Type-Token Ratio (TTR) of 10.42 percent and a Standardized Type-Token Ratio (STTR) of 39.65 percent, indicating moderate lexical diversity. When comparing the two genders, a gendered comparison shows that the Female Speakers Corpus has greater lexical variance, with TTR of 14.17% and STTR of 40.64%, whereas the Male Speakers Corpus has slightly less constrained figures (TTR = 12.77%, STTR = 38.59%). Such numbers indicate that TED speakers use the middle register - neither too monotonous nor too rich. The median STTR of about 40 percent implies that speakers have reached a compromise between accessibility and lexical richness, providing enough variety to maintain audience attention and supporting understanding by creating a repetition strategy. The relatively higher level of lexical diversity in the speech of female speakers can be explained by their use of more diverse rhetorical strategies and a greater vocabulary to express complex concepts in a more vivid way. This moderate lexical deviation serves the rhetorical purpose of TED Talks, as it allows for the integration of more specific terminology and supports important ideas through effective repetition.

Figure 6. The Token-Type Ratio in the Study Corpora

N	text file	file size	types (distinct words)	type/token ratio (TTR)	standardised TTR
	Overall	777,062	7,365	5.21%	39.63%
1	TED corpus 30 talks.txt	388,533	7,365	10.42%	39.65%
2	TED corpus for Female talks.txt	198,206	5,038	14.17%	40.64%
3	TED corpus for Male talks.txt	190,323	4,483	12.77%	38.59%

Analysis of Content Words Usage:

Verbs :

Analyzing the verb allocation in TED talks provides abundant information on the lexical-level stylistic features of verb usage. By comparing the frequency of the most commonly used verbs across gender-specific corpora, it can be shown that speakers, regardless of gender, used a diverse range of verbs to maintain audience attention and convey complex information. Both corpora share similarity in that they are the most frequent in the frequency lists, comprising auxiliary and linking verbs. The different versions of the verb to be (is, was, 3s, are, 3 r, 3 m, were, been, being) take an eminent place in these data sets. We can explain this fact by the expository and explanatory character of TED presentations, where speakers indicate relations, define terms, and characterize states or conditions. Moreover, the commonness of the copular verb be supports the finding of Biber and Conrad (2019) that informative language tends to provide explicit links between the subjects and the complements, which involve copular verb forms.

Regarding tense patterns, both female and male speakers use both present (is, are, do, want, etc.) and past (was, were, had, said, did) tenses, with a slight tendency toward the past tense. This moderate approach enables the speakers to relate previous experiences or studies to their current implications, a typical feature of effective knowledge dissemination in the art of speaking to the masses. There is significant variation in the distribution of modes and action verbs between the gender corpora. Females are more likely to use verbs of desire and emotion quite often (e.g., want, feel) than male speakers, who are more prone to use perceptual and communicative verbs (e.g., see, say). This trend suggests slight differences in rhetorical style: female speakers might be more direct to elicit emotional appeal, whilst male speakers might prefer observation-based methods of persuasion. Cognitive verbs are also common in both corpora (e.g., know, think), which characterizes the intellectual nature of TED Talks. The abundance of dynamic (go, come, get, make) and stative (know, want, feel) verbs creates a balanced discourse that explains concepts and describes experiences simultaneously, thereby enhancing listeners' interest.

Nouns :

The analysis of the nouns used within TED talks demonstrates that even the more successful speakers of the TED talks strike the right balance between abstract and concrete nouns and retain the human-based focus of the discourse that contributes to better audience-speaker interaction and understanding. However, there are slight gender variations in the choice of the theme.

The noun people takes the first place in both corpora (189 times in the female corpus, 108 in the male corpus), which is also a human-oriented concern in excellent oratory. According to Leech and Short's (2007) classification, the Female Speakers Corpus shows an equal distribution of abstract nouns (world, time, life, way, compassion) and concrete nouns (people, books, stories, women). Abstract and concrete nouns, respectively, make up about 46 and 43 percent of frequent nouns. The balance suggests that female speakers can successfully intertwine conceptual information with concrete examples to help the audience better understand it.

In the meantime, concrete nouns (people, music, thing, story, kids, school, brain) are favored by the Male Speakers Corpus at about 52 per cent., whereas abstract nouns (time, way, kind, world, life) are favored at 41 per cent. It means that male speakers can use more object-oriented or experience-oriented explanations in their presentations.

There is an important thematic difference in the corpora; female speakers have higher frequencies of nouns that are connected with communication and human connection (books, stories, compassion, self), and male speakers with nouns associated with creative expression and arrangement of the environments (music, art, laughter, school, education). There are no high-frequency proper names in either corpus, suggesting that TED speakers generally speak in broad terms rather than naming specific items. Nevertheless, all collective nouns (family, school) are used in both corpora, which can help speakers build rapport and create a shared experience with their listeners.

Adverbs:

A comparison of the adverb patterns in each of the corpora indicates that there are general tendencies in the use of adverbs, which suggests that no matter what the gender, adverb strategies are applicable in effective oratory. However, even slight differences in the preference towards specific adverbs can indicate slight differences in rhetorical strategy. The analysis showed that not and n't were the most popular adverbs in both corpora, followed by so and just, which were predominantly used to organize the idea and tone of the speech. In terms of semantic classification, as explained by Biber and other people (1999), both corpora show similar tendencies in the distribution of adverbial forms. It is worth mentioning that the use of linking adverbs, including so, then, and also, is conspicuously placed, as they serve to weave conceptual threads, hence, to direct the cognitive mapping of the readership. The occurrence of so is high (252 in the female corpus and 263 in the male corpus), which is evidence of its necessity to provide a logical flow and state conclusions or implications. The use of stance adverbs (really, actually, basically) is also very common in both corpora, allowing speakers to express their attitude toward the content and convey authenticity. There is a slight difference in the use of the word really (102 and 81 instances respectively) by female speakers, which may indicate a tendency to focus on genuineness or trust

Additive/restrictive adverbs (e.g., even, only) also play a major role, enabling speakers to focus the audience's attention on certain ideas. The word just is commonly used in both corpora and is commonly used to make difficult ideas easier to understand or to emphasize the important points. Place and time adverbs (here, there, up, down, now, then, never) place the concepts spatially and temporally, making the abstract concepts more concrete for the audience. Adverbs are the most commonly used device among TED speakers to explain ideas, emphasize important points, and create a conversational tone. **Adjectives:** The percentage of adjectives in the total number of tokens is lower than that of other content words, which indicates their selectivity in use. In the two-gender corpora, speakers use adjectives as a resource, not as a decorative device; instead, adjectives are mainly used to make abstract ideas more concrete so the audience can easily understand them. In both corpora, general descriptive adjectives prevail, as the character of TED Talks is to explain something, and the speaker must be able to clarify concepts to various audiences. A significant difference emerges in the use of evaluative adjectives. Comparative and superlative forms are also commonly used by female speakers (more, most, better), indicating that they prefer to introduce ideas in a series or in reference to various options. Accordingly, the male speaker often employed absolute evaluative adjectives, such as good, right, and wrong, to convey a more assertive rhetorical style. Regarding the use of attributive versus predicative words, both gender corpora are more inclined toward attributive adjectives (that is, modifying nouns directly), reflecting the informative nature of TED Talks, in which characterization helps people understand the speech.

Pronouns:

Pronouns make up a significant part of all words used in both corpora, and this fact brings out the importance of using pronouns in delivering effective speech. According to the analysis, personal pronouns are the most common in both corporations. The first-person singular pronoun I is the most frequently used pronoun in TED

Talk transcripts, which is atypical of other genres of discourse. This increased usage aligns with the fact that the genre depends on personal stories and the sharing of ideas. One more remarkable difference concerns the use of second-person pronouns. Male speakers use 'you' more than female ones, suggesting they are more likely to address the audience directly. On the other hand, the female speakers prefer to use possessive pronouns like my and our, a decision that can indicate the focus on the development of personal relations and a common experience basis. The personal pronoun we use in both corpora demonstrates how the speakers are trying to create a sense of belonging to a community of listeners. This rhetorical approach is inclusive and, consequently, fosters a collective view and strengthens group identity, which is probably very important to persuasive speech. The gender-based corpora show a similar distribution of third-person pronouns (he, she, they, them). However, the use of she is even less frequent among female speakers, whereas males use he more frequently, indicating a slight tendency to invoke gendered exemplars in the discourse.

Keywords Analysis:

The lexical analysis of the corpus of TED Talks will reveal the linguistic features that make effective use of language in public discourse, compared to other daily life situations. This paper used WordSmith Tools, the Log-Likelihood (LL) statistical tool, and the Bayesian Information Criterion (BIC) statistical tool to determine those words that appear to be statistically significantly more common in TED Talks. The top of the keyword register is taken up by keywords that intuitively fit the motivational speech theme of the population. The first-person singular is the most common keyword (LL = 605.00, BIC = 586.27), which is used to convey credibility and connection with the speaker, that is, an individualized narrative style. This highlights the individualism of good motivational speech.

The keywords can be subdivided into a few semantic categories, proving the key characteristics of the genre:

- Audience Engagement Terms: Terms that directly address the audience or engage the audience, such as YOUre, (LL = 298.90, BIC = 280.17) or words in the form of a pronoun, including we, our, ourselves, which are statistically significant. This testifies to the effectiveness of speakers who make their speech a conversation in this way, therefore, creating the illusion of immediacy when a vast number of people listen to a monologue.
- Emotional and Experiential Terms: Emotional terms, such as COMPASSION (LL = -117.10, BIC = -98.37) and VULNERability (LL = -301.34, BIC = -282.61), highlight the importance of empathy in maintaining audience attention. Competent speakers use genuine self-disclosure rather than merely conveying information, thereby building an emotional connection with listeners.
- Narrative and Knowledge Framing: Words with keywords related to storytelling predominate (STORIES (LL = 222.90, BIC = 204.18), STORY (LL = 165.94, BIC = 147.21), and BOOKS (LL = 188.47, BIC = 169.75)). This trend is indicative of the narrative frameworks adopted by skilled orators to make complex concepts understandable and memorable to different audiences.
- Innovation and Creative Domains: The vocabulary associated with innovation and creativity includes deception, music, art, power, and creativity. This denotes a focus on new ideas and interdisciplinary investigation, which define the genre.

The “aboutness” of this genre can thus be described as through transformation (personal tales that bring wider understanding), accessibility (complex ideas made understandable), and inspiration. Such word choices make public motivational speaking stand out from the rest of the discourse and play a significant role in its success in spreading knowledge and interacting with audiences. Figure 7 shows the keywords in the combined TED talk corpus.

Figure 7. List of Keywords to be used in Combined TED Talk Corpus.

N	Key word	Freq.	%	Texts	RC. Freq.	Rc. %	BIC	Log_L	Log_R	P
1	I'M	212	0.30%	1	39,455	0.03%	586.27	605.00	3.36	0.0000000000
2	LAUGHTER	54	0.08%	1	737		404.26	422.99	7.13	0.0000000000
3	BOTH	27	0.04%	1	27		314.62	333.34	10.90	0.0000000000
4	COMPASSION	43	0.06%	1	901		282.61	301.34	6.52	0.0000000000
5	YOU'RE	130	0.18%	1	33,112	0.02%	280.17	298.90	2.91	0.0000000000
6	DECEPTION	34	0.05%	1	284		279.31	298.04	7.84	0.0000000000
7	WORLD	125	0.18%	1	34,451	0.03%	251.72	270.45	2.80	0.0000000000
8	THAT'S	177	0.25%	1	77,046	0.06%	232.16	250.89	2.14	0.0000000000
9	MUSIC	84	0.12%	1	15,023	0.01%	227.02	245.75	3.42	0.0000000000
10	STORIES	69	0.10%	1	10,439		204.18	222.90	3.66	0.0000000000
11	BOOKS	57	0.08%	1	8,274		169.75	188.47	3.72	0.0000000000
12	WE'RE	97	0.14%	1	34,289	0.03%	150.81	169.54	2.44	0.0000000000
13	STORY	96	0.14%	1	34,342	0.03%	147.21	165.94	2.42	0.0000000000
14	ART	43	0.06%	1	5,013		140.95	159.68	4.04	0.0000000000
15	POWER	55	0.08%	1	10,564		135.11	153.84	3.32	0.0000000000
16	CREATIVITY	23	0.03%	1	728		124.10	142.83	5.92	0.0000000000
17	PERCENT	28	0.04%	1	1,621		122.49	141.22	5.05	0.0000000000
18	OURSELVES	36	0.05%	1	3,967		118.76	137.49	4.12	0.0000000000
19	CLASSICAL	18	0.03%	1	277		118.12	136.85	6.96	0.0000000000
20	BOOK	53	0.07%	1	12,309		111.65	130.38	3.04	0.0000000000
21	AFRICA	22	0.03%	1	838		110.00	128.73	5.65	0.0000000000
22	POWERFUL	28	0.04%	1	2,122		108.05	126.78	4.66	0.0000000000
23	BLAH	32	0.05%	1	3,489		104.12	122.85	4.14	0.0000000000
24	NONVERBALS	10	0.01%	1	11		103.39	122.12	10.77	0.0000000000
25	FOUR	51	0.07%	1	12,558		101.52	120.25	2.96	0.0000000000
26	THEY'RE	84	0.12%	1	36,687	0.03%	99.90	118.63	2.13	0.0000000000
27	STORYTELLER	11	0.02%	1	31		99.29	118.02	9.41	0.0000000000
28	TRUTH	44	0.06%	1	9,017		99.24	117.97	3.23	0.0000000000
29	VULNERABILITY	17	0.02%	1	379		98.37	117.10	6.43	0.0000000000
30	AFRICAN	19	0.03%	1	744		91.43	110.16	5.61	0.0000000000
31	INTROVERTS	14	0.02%	1	188		91.37	110.10	7.16	0.0000000000
32	NICO	9	0.01%	1	15		85.59	104.32	10.17	0.0000000000
33	DESIGN	25	0.04%	1	2,530		80.75	99.48	4.24	0.0000000000
34	QUOTE	8	0.01%	1	8		80.04	98.77	10.90	0.0000000000
35	OLÉ	8	0.01%	1	8		80.04	98.77	10.90	0.0000000000
36	CREATIVE	25	0.04%	1	2,619		79.12	97.85	4.19	0.0000000000
37	MYSTERY	17	0.02%	1	704		78.00	96.73	5.53	0.0000000000
38	RANDOLPH	8	0.01%	1	11		76.36	95.09	10.44	0.0000000000
39	HUMILIATION	13	0.02%	1	259		73.64	92.37	6.59	0.0000000000
40	SEVEN	24	0.03%	1	2,614		73.45	92.18	4.14	0.0000000000
41	LEGACY	13	0.02%	1	265		73.07	91.79	6.55	0.0000000000
42	NETRA	7		1	7		67.69	86.42	10.90	0.0000000000
43	BLIH	7		1	8		66.38	85.10	10.71	0.0000000000

Conclusion :

This paper examines the linguistic and stylistic characteristics of motivational speeches in a corpus-based analysis of TED Talks. The evaluation indicates that the main features of effective motivational discourse are lexical and syntactic peculiarities, such as the use of shorter words (234 letters) to acquire instant understanding, moderate sentence structure (14.93 words), clarity/depth ratio, and moderate lexical heterogeneity. Female speakers use more communication-related nouns, emotive verbs, possessive pronouns, and men speakers use concrete and action verbs and direct address to the audience. Strategic speakers will adopt gender-bound strategies to connect with the audience by balancing abstract ideas with concrete demonstrations and using strategic pronouns to build credibility and connection. The study fills an important gap in literature by offering an empirical, gender-sensitive study of the linguistic patterns of public speaking. The corpus method provides measurable data on real language patterns, enabling efficient knowledge communication and appealing to the audience. The results can be used by a wide audience of people interested in building effective public speeches, communication coaches designing a training programme, and speakers to improve their inspirational influence through evidence-based linguistic techniques.

References :

- Alnasser, A. A. (2023). *Argumentation and persuasion in Saudi TEDx talks : A political discourse analysis approach* (Doctoral dissertation, University of Leeds). White Rose ETheses Online. <https://etheses.whiterose.ac.uk/33624/>
- Altamimi, S. (2021). Disseminating knowledge: A discourse analysis of terrorism in TED talks. *Heliyon*, 7(2), e06312. <https://doi.org/10.1016/j.heliyon.2021.e06312>
- Al-Nakeeb, O. A. M. S., & Mufleh, B. A. H. (2018). Collocations and collocational networks of characters: A corpus-based feminist stylistic analysis. *Language in India*, 18(9). <https://www.languageinindia.com/>

- Azlia, S. C. (2022). Interactional discourse of male and female motivational speech in TED Talks: A corpus-based study. *Journal of Literature, Linguistics and Culture Studies*, 11(1), 42–49. <https://doi.org/10.15294/rainbow.v11i1.54777>
- Biber, D., & Conrad, S. (2019). *Register, genre, and style*. Cambridge University Press.
- Di Carlo, G. S. (2015). Stance in TED talks: Strategic use of subjective adjectives in online popularisation. *Iberica*, 29, 201–222. <http://search.ebscohost.com/login.aspx?direct=true&db=ufh&AN=109298699&site=ehost-live>
- Di Carlo, G. S. (2014). New trends in knowledge dissemination: TED Talks. *Acta Scientiarum Language and Culture*, 36(2), 121–130. <https://doi.org/10.4025/actascilangcult.v36i2.22619>
- Di Carlo, G. S. (2015). Pathos as a communicative strategy for online knowledge dissemination: The case of TED talks. *3L: Language, Linguistics, Literature*, 21(1), 23–34. <https://doi.org/10.17576/3L-2015-2101-03>
- Feng, L., Jansche, M., Huenerfauth, M., & Elhadad, N. (2010). A comparison of features for automatic readability assessment. In *Proceedings of the 27th International Conference on Computational Linguistics*, (pp.27. Association for Computational Linguistics. <https://aclanthology.org/C10-2032.pdf>
- Ghazala, H. S. (2011). *Cognitive stylistics and the translator*. Sayyab Books Ltd. <https://doi.org/10.31235/osf.io/hjnx2>
- Jones, C., & Waller, D. (2015). *Corpus linguistics for grammar: A guide for research*. Routledge. <https://doi.org/10.4324/9781315713779>
- Kahl, D. H. (2014). High school public speaking curriculum: Assessment through student voice. *Qualitative Research Reports in Communication*, 15(1), 51–58. <https://doi.org/10.1080/17459435.2014.955592>
- Kang, N., & Yu, Q. (2011). Corpus-based stylistic analysis of tourism English. *Journal of Language Teaching & Research*, 2(1).
- Leech, G., & Short, M. (2007). *Style in fiction: A linguistic introduction to English fictional prose*. Pearson Education.
- Marchenko, V., & Minenko, Y. (2020). Linguistic and extralinguistic characteristics of Ted Talks As multimodal English text. *Advanced Linguistics*, 0(5), 27–32. <https://doi.org/10.20535/2617-5339.2020.5.205688>
- Mattiello, E. (2019). A corpus-based analysis of scientific TED Talks: Explaining cancer-related topics to non-experts. *Discourse, Context and Media*, 28, 60–68. <https://doi.org/10.1016/j.dcm.2018.09.004>
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus linguistics: the basics*. In *Corpus-based Language Studies: An Advanced Resource Book*. Routledge
- Nørgaard, N., Montoro, R., & Busse, B. (2010). *Key terms in stylistics*. Continuum International Publishing Group.
- Paltridge, B. (2012). *Discourse Analysis: An Introduction* (2nd ed). Bloomsbury Academic.
- Sanyal, J. (2006). *Indlish: The book for every English-speaking Indian*. Viva Books Ltd
- Salahuddin, M., Abdulaziz, M., Ahmad, S., & Qadir, U. (2023). The use of figurative language in public speaking: An analysis of simile structure in english ted talk. *Russian Law Journal*, 11(5), 1237–1247. <https://openscience.uz/index.php/sciedu/article/view/4847>
- Salahuddin, M., Abdulaziz, M., Qadir, U., & Mastoor, F. (2023). Stylistic analysis of influential monologues: An exploration of prominent stylistic features in English TED talks. *Russian Law Journal*, 11(5), 308–327. <https://cyberleninka.ru/article/n/stylistic-analysis-of-influential-monologues-an-exploration-of-prominent-stylistic-features-in-english-ted-talks>
- Scotto di Carlo, G. (2014). The role of proximity in online popularizations: The case of TED talks. *Discourse Studies*, 16(5), 591–606. <https://journals.sagepub.com/doi/abs/10.1177/1461445614538565>
- Tan, G. (2023). Study on English Public Speaking Teaching from the Perspective of Memetics. *SHS Web of Conferences*, 168, 02028. <https://doi.org/10.1051/shsconf/202316802028>
- Wang, Y.-C. (2012). An exploration of vocabulary knowledge in English short talks- A Corpus-driven approach. *International Journal of English Linguistics*, 2(4). <https://doi.org/10.5539/ijel.v2n4p33>
- Hollingsworth, S. E., Weinland, K., Hanrahan, S., Walker, M., Elwood, T., & Linsenmeyer, M. (2021). *Introduction to speech communication*. Oklahoma State University Libraries. <https://doi.org/10.22488/okstate.23.000003>
- Yang, M. (2019). The application of English rhetorical devices in speechwriting : Taking the preliminary of

Kunming University in 2018 “FLTRP Cup” English public speaking contest as an example. *Advances in Economics, Business and Management Research*, 96, 399–403. <https://doi.org/10.2991/icemse-19.2019.93>

Appendix 1. TagAnt Abbreviations List

TreeTagger Tag Set (58 tags)

POS Tag	Description	Example	POS Tag	Description	Example
CC	coordinating conjunction	<i>and, but, or, &</i>	VB	verb <i>be</i> , base form	<i>be</i>
CD	cardinal number	<i>1, three</i>	VBD	verb <i>be</i> , past	<i>was/were</i>
DT	determiner	<i>the</i>	VBG	verb <i>be</i> , gerund/participle	<i>being</i>
EX	existential there	<i>there is</i>	VBN	verb <i>be</i> , past participle	<i>been</i>
FW	foreign word	<i>d'œuvre</i>	VBZ	verb <i>be</i> , pres, 3rd p. sing	<i>is</i>
IN	preposition/subord. conj.	<i>in, of, like, after, whether</i>	VBP	verb <i>be</i> , pres non-3rd p.	<i>am/are</i>
IN/that	complementizer	<i>that</i>	VD	verb <i>do</i> , base form	<i>do</i>
JJ	adjective	<i>green</i>	VDD	verb <i>do</i> , past	<i>did</i>
JJR	adjective, comparative	<i>greener</i>	VDG	verb <i>do</i> gerund/participle	<i>doing</i>
JJS	adjective, superlative	<i>greenest</i>	VDN	verb <i>do</i> , past participle	<i>done</i>
LS	list marker	<i>(1),</i>	VDZ	verb <i>do</i> , pres, 3rd per. sing	<i>does</i>
MD	modal	<i>could, will</i>	VDP	verb <i>do</i> , pres, non-3rd per.	<i>do</i>
NN	noun, singular or mass	<i>table</i>	VH	verb <i>have</i> , base form	<i>have</i>
NNS	noun plural	<i>tables</i>	VHD	verb <i>have</i> , past	<i>had</i>
NP	proper noun, singular	<i>John</i>	VHG	verb <i>have</i> , gerund/participle	<i>having</i>
NPS	proper noun, plural	<i>Vikings</i>	VHN	verb <i>have</i> , past participle	<i>had</i>
PDT	predeterminer	<i>both the boys</i>	VHZ	verb <i>have</i> , pres 3rd per. sing	<i>has</i>
POS	possessive ending	<i>friend's</i>	VHP	verb <i>have</i> , pres non-3rd per.	<i>have</i>
PP	personal pronoun	<i>I, he, it</i>	VV	verb, base form	<i>take</i>
PP\$	possessive pronoun	<i>my, his</i>	VVD	verb, past tense	<i>took</i>
RB	adverb	<i>however, usually, here, not</i>	VVG	verb, gerund/participle	<i>taking</i>
RBR	adverb, comparative	<i>better</i>	VVN	verb, past participle	<i>taken</i>
RBS	adverb, superlative	<i>best</i>	VVP	verb, present, non-3rd p.	<i>take</i>
RP	particle	<i>give up</i>	VVZ	verb, present 3d p. sing.	<i>takes</i>
SENT	end punctuation	<i>?, !, .</i>	WDT	wh-determiner	<i>which</i>
SYM	symbol	<i>@, +, *, ^, , =</i>	WP	wh-pronoun	<i>who, what</i>
TO	<i>to</i>	<i>to go, to him</i>	WP\$	possessive wh-pronoun	<i>whose</i>
UH	interjection	<i>uhhuhhhh</i>	WRB	wh-abverb	<i>where, when</i>
			:	general joiner	<i>;; , --</i>
			\$	currency symbol	<i>\$, £</i>

Adapted from: <https://courses.washington.edu/hypertext/csar-v02/penntable.html>