Database Software and Spoken Language Analysis

Chadia CHIOUKH[1], **Jijel University ,Algeria.** chioukh.chadia@gmail.com

Received: 01/05/2021 **Accepted: 20/06/2021** **Published:30/06/2021**

Abstract:

The introduction of Computer Assisted Language Learning (CALL), with its comprehensive, innovative database software, has assisted researchers in the realm of corpus linguistics to analyse in-depth learners' corpora. The latter, being written or spoken, can be scrutinised from diverse perspectives hinging on the researcher's prime aims. The process of recording, transcribing and analysing spoken language might be intricate. Given that, many Master two students at the department of English Language of Mohammed Seddik Ben Yahia University- Jijel show reluctance to use database applications to treat their spoken corpora, as it requires both the analysis and the transcription of audio files into written one; a process that might be tiresome and time-consuming if conducted manually. Therefore, the present paper aspires to introduce some widely used software that may facilitate their research process and expose some practical applications that might be used in the data analysis. More importantly, the paper aims at displaying some web-based tools put into practice to probe some speaking skill features, namely fluency and Complexity. The massive range of applications accredited to analyse language complexity is textinspector.com and lextutor.ca, whereas the Praat application is deemed one of the best software implemented to investigate speech fluency.

Keywords*: Computer Assisted Language Learning (CALL), Corpus Linguistics, Spoken Corpora, Speaking skill, Complexity, fluency.*

ملخص:

لقد ساعد تعلم اللغه بمساعده الكمبيوتر مع برامج قواعد البيانات الشامل والمبتكر الباحثين في مجال اللسانيات من تحليل بيانات خطاب المتعلمين بصفه أكثر تعمق. اذ يمكن دراستها، سواء أكانت مكتوبة أو منطوقة، من وجهات نظر مختلفه بناءا على اهداف الباحثين. قد تكون عملية تسجيل اللغة المنطوقة وكتابتها وتحليلها معقدة. بالنظر إلى أن العديد من طلبة الماستر في قسم اللغة الإنجليزية بجامعة محمد الصديق بن يحي- جيجل، يظهرون عزوفًا عن استخدام تطبيقات قواعد البيانات لمعالجة بياناتهم المنطوقة، حيث يتطلب ذلك تحليلا ونسخا الملفات الصوتية إلى ملف مكتوب ؛ عملية قد تكون مرهقة وتستغرق وقتًا طويلاً إذا تم إجراؤها يدويًا. لذا يطمح هذا المقال إلى التعريف ببعض البرامج المستخدمة على نطاق واسع والتي قد تسهل عمليةالبحث. والأهم من ذلك، يهدف هذا المقال إلى عرض بعض الأدوات المستندة إلى الشبكة العنكبوتية التي تساعد في دراسة بعض ميزات مهارة التحدث ، وهي الطلاقة والتعقيد. التطبيقات المعتمدة لتحليل تعقيد اللغة والمعروضة في هذا المقال هما textinspector.com و lextutor.ca .

أما تطبيق برات يعتبر من اهم التطبيقات المعتمدة لدراسة طلاقة الكلام

الكلمات المفتاحية: تعلم اللغة بمساعدة الكمبيوتر- معالجه البيانات اللغوية- البيانات الشفوية- مهاره التحدث- تعقيد الكلام- طلاقة الكلام.

[1] **Corresponding author:** Chadia CHIOUKH, **e-mail address:** chioukh.chadia@gmail.com

1. Introduction

The massive range of applications and data-based language tools has incredibly facilitated the process of recording, transcribing and analysing spoken language output. Researchers interested in analysing learners' spoken language are fortunately introduced with many free online tools that might save their time and effort. For instance, researchers attempting to study English as a foreign language (EFL) learners speaking skill might be offered many practical applications to record the necessary data, of which is the PRAAT application. The latter is widely implemented to analyse speech fluency. They may also access some free online web-based language tools such as textinspector.com and lextutor.com to analyse speech complexity. These websites are impressive as they provide statistics relating to copied and pasted texts. More importantly, these online language tools probe in-depth some texts' features such as frequencies and vocabulary.

2. Computer-Assisted Language Learning (CALL)

Computer-assisted Language Learning (CALL) uses computers in the processes of learning and teaching a second and a foreign language (Tavakoli, 2012). CALL pertains to an interdisciplinary approach in which computers are used "as an aid to the presentation, reinforcement, and assessment of material to be learned, usually including a substantial interactive element" (Tavakoli, 2012, p.78).

3. Corpus Linguistics

The advent of a wide range of computers and software has immensely shaped linguistic studies. The ability to electronically storing chunks of naturally produced language led to the evolution of what is known today as corpus linguistics, as elucidated in the words of Breyer (2011): "The ability to store large amounts of language data electronically and to access and retrieve this data through a software interface has paved the way for the emergence of corpus linguistics." (p.1). Corpus linguistics, as a discipline, then aims at storing language data in digital format or corpora.

4. Spoken Corpora

Spoken Corpora are a spectrum of speech recordings that have been transcribed into written texts to form a database or a corpus (Caines, McCarthy and O'Keeffe, 2016, p.348). As a term, it should not be confounded by the speech corpora concept. As termed by Baker, Hardie and McEnery (2006), a spoken corpus is "a corpus consisting entirely of transcribed speech. This could be from a range of sources: spontaneous informal conversations, radio phone-ins, meetings, debates, classroom situations etc." (p.148). Given that it is time-consuming to compile and transcribe, spoken corpora did not thrive than written corpora until the last twenty years (Caines et al., 2016). Transcribing speech produced by a second (L2) or a foreign learner (FL) and compile it into a database may make it tremendously challenging for many researchers.

Collecting data on spoken corpora of second or foreign language learners is deemed valuable in the second language acquisition (SLA) realm. Spoken corpora analysis provides insights into linguistic features, knowledge and usage of learners' language and discerns how different these are produced by native speakers (Yoon, 2020). Put otherwise, spoken corpora can be used to probe some aspects in learners' language that are due to first language (L1) influence. Spoken corpora can also reflect aspects of development processes (Ellis and Barkhuizen, 2005, p.336). Likewise, they are significant as they demonstrate repeated frequencies and patterning in the language (Mauranen, 2004, p.102). Given their naturally occurring production in courses of interaction, spoken corpora provide teachers, researchers, and educationists interested in noticing patterning a large number of instances for observation in a speech.

Spoken corpora are accredited to provide linguistic materials that are high in terms of authenticity (Mauranen, 2004, p.103). Bearing in mind that they are produced spontaneously by an L2 or an FL learner, spoken corpora represent rich linguistic data for scrutinising many language aspects relating to lexis, grammar…etc. (Meyer, 2009). They demonstrate prevailing and recurrent features in target language speech texts, typical of L2 or FL learners' language. Considering their genuine production, spoken corpora exhibit abundant examples in which ellipsis, word-order acquisition, part of speech error occur. (Mackey& Gass, 2005). Their prominence relates to the fact that they give access to researchers attempting to probe second language acquisition (SLA) process.

5. Steps of Conducting Spoken Data Analysis

Researchers interested in conducting spoken corpus analysis need first to record speech and gather relevant data as termed by Meyer (2009): "Collecting data involves recording speech, gathering written texts, obtaining permission from speakers and writers to use their texts, and keeping careful records about the texts collected and the individuals from whom they were obtained" (p.55).

5.1. Data Collection Stage

As highlighted previously, the researcher needs to record speech as a preliminary step to data gathering, ask for permission from speakers to use their texts. The recorded speech has to be manually transcribed and then saved for computerisation and annotation (Meyer, 2009, p. 55). Nevertheless, before being transcribed, the recorded speech has to be as natural as possible. Doing this implies recording L2 and FL learners in a natural speaking environment, as Meyer (2009) suggested. The quality of the recordings is another issue that the researcher needs to consider when starting data collection (Meyer, 2009).

5.2. Data Computerisation Stage

While it is now an easy process to prepare written texts for computerisation in a corpus, there is little hope for making spoken text compilation and transcription easier. It will be challenging for the foreseeable future to find people who might be willing to be recorded, make recordings, and have the recordings transcribed. While digitising spoken samples and using specialised software to transcribe them has advantages, the transcriber must still listen to speech segments many times to achieve an accurate transcription. Advances in speech recognition technology may enable the transcription of certain types of speech (for example, speeches) to be automated (Meyer, 2009); Du Bois (2006).

5.3. Data Annotation Stage

Corpus annotation refers to "the practice of adding interpretive, linguistic information to an electronic corpus of spoken and/or written language data." (Garside, Leech, & McEnery, 1997, p. 2). Put differently, researchers interested in constructing spoken corpora should consider interpreting the data by providing necessary information. Gardside et al. (1997) further explained:

> Annotation can also refer to the end-product of this process: the
> linguistic symbols
> which are attached to, linked with, or interspersed with the electronic
> representation
> of the language material. A typical and familiar case of Corpus
> annotation is
> grammatical tagging (also called world-class tagging, part of speech
> tagging or POS
> tagging) (p.2).

Annotating data is vital for corpus analysts, as it represents their corpus electronically.

4.4. Analysing Spoken Data

Researchers interested in analysing spoken data that is already compiled, transcribed, and annotated may probe participants' accuracy, fluency, and complexity as fundamental components of the speaking skill. Within the frame of this article, light is shed on some prominent existing database software and web-based language tools that may facilitate the process of analysis for researchers in the field.

4.4.1. Fluency, Complexity and Database Software and Online Web-based Tools

Researchers can use a wide range of applications and software to do an automated speech analysis, more precisely speaking fluency and Complexity components.

4.4.1.1. Fluency: an Operational Definition

Fluency, as termed by Tavakoli (2013): "might be the rapid, smooth, effortless, accurate, lucid, and efficient translation of thought or communicative intention into language under the temporal constraints of online processing." (p.135). Similarly, Ellis& Barkhuizen (2005) defined fluency as: "the production of language in real-time without undue pausing or hesitation" (p.39). It is manifested in the frequency of pauses in speech, the use of fillers in pauses, length of runs (number of syllables separating pauses) (Ellis& Barkhuizen, 2005, p39).

Fluency measurements are twofold: *temporal variables* and *hesitation phenomena* (Ellis &Barkhuizen, 2005, p. 156). The former pertains to speech speed, while the latter relates to dysfluencies.

- *Temporal Variables:* number of pauses, pause length, length of run.
- *Hesitation Phenomena:* reformulation, false start, repetitions and replacements.

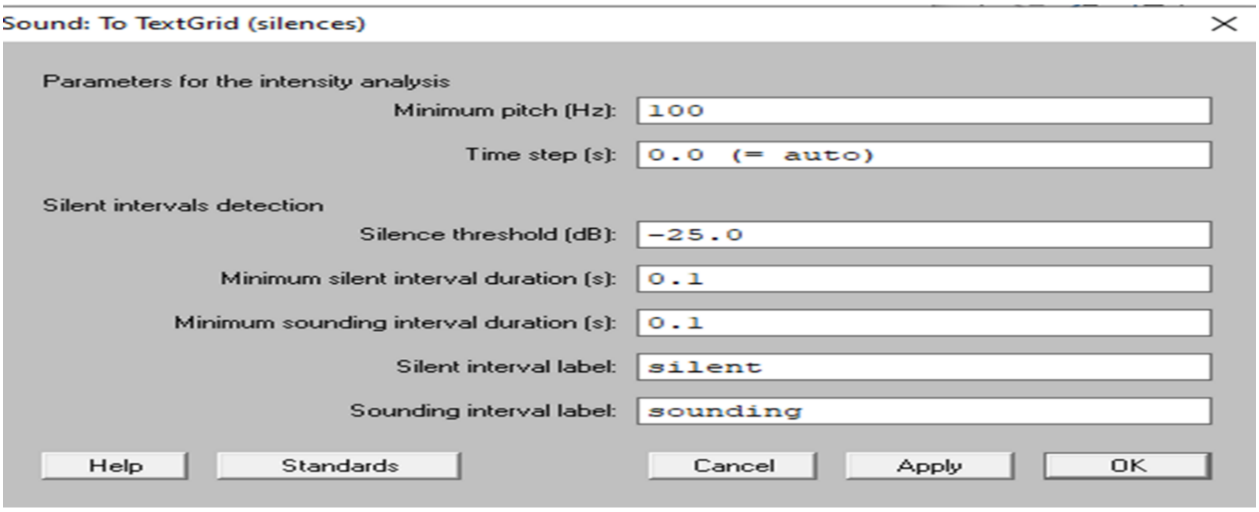
As an index of fluency, temporal variables can be automated using some web-based applications and software, of which is PRAAT.

4.4.1.2.Fluency and PRAAT Software

PRAAT software is a speech analysis programme designed "for phonetic and acoustic analysis" (Boutsen and Dvorak, 2016, p. 78). Temporal variables explained above can be measured via PRAAT application as it displays:

- ➢ the frequency and the length of pauses in recorded speech using Textgrid (silence in milliseconds),
- ➢ the use of fillers in pauses,
- ➢ the length of runs (number of syllables separating pauses).

The following figure is an illustration of how the PRAAT application can identify silence intervals and pause lengths.



Sound: To TextGrid (silences)

Parameters for the intensity analysis

Minimum pitch (Hz): 100

Time step (s): 0.0 (= auto)

Silent intervals detection

Silence threshold (dB): -25.0

Minimum silent interval duration (s): 0.1

Minimum sounding interval duration (s): 0.1

Silent interval label: silent

Sounding interval label: sounding

Figure1. PRAAT Analysis of Length Pauses

Researchers can use the PRAAT application to record speech and analyse communication breakdown by measuring silent and filled pauses and speed fluency. Researchers can either create a speech file from scratch or read an already recorded one (Boersma and Weenink,2021). The following is a figure that demonstrates a sample of a PRAAT speech recording.

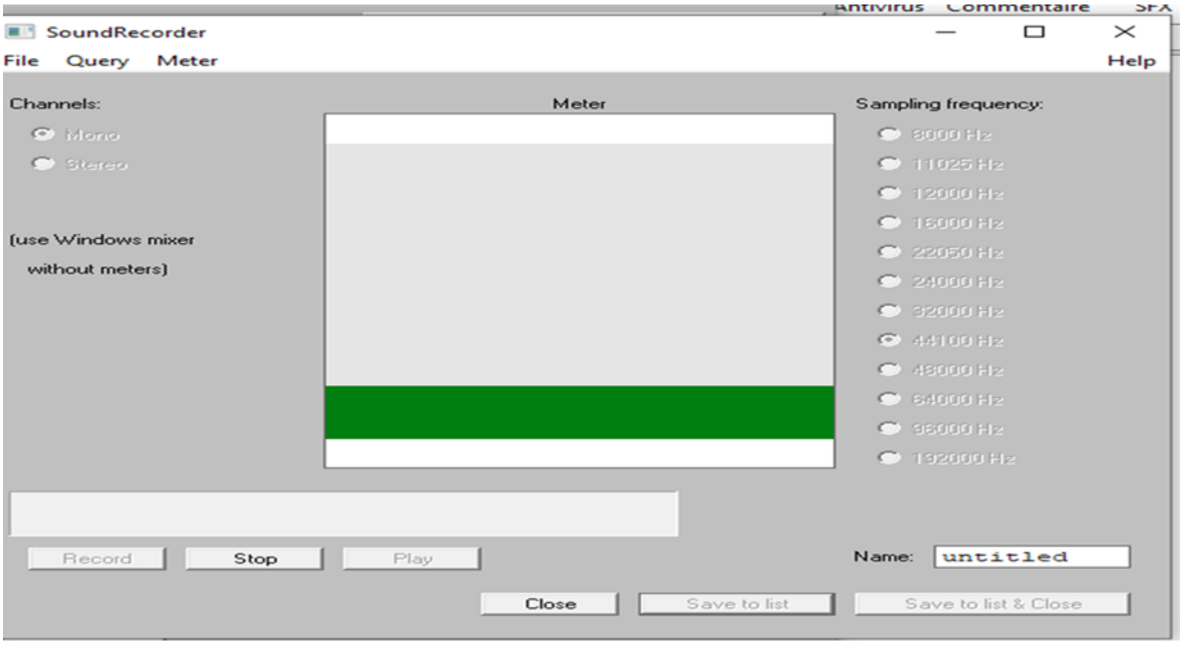
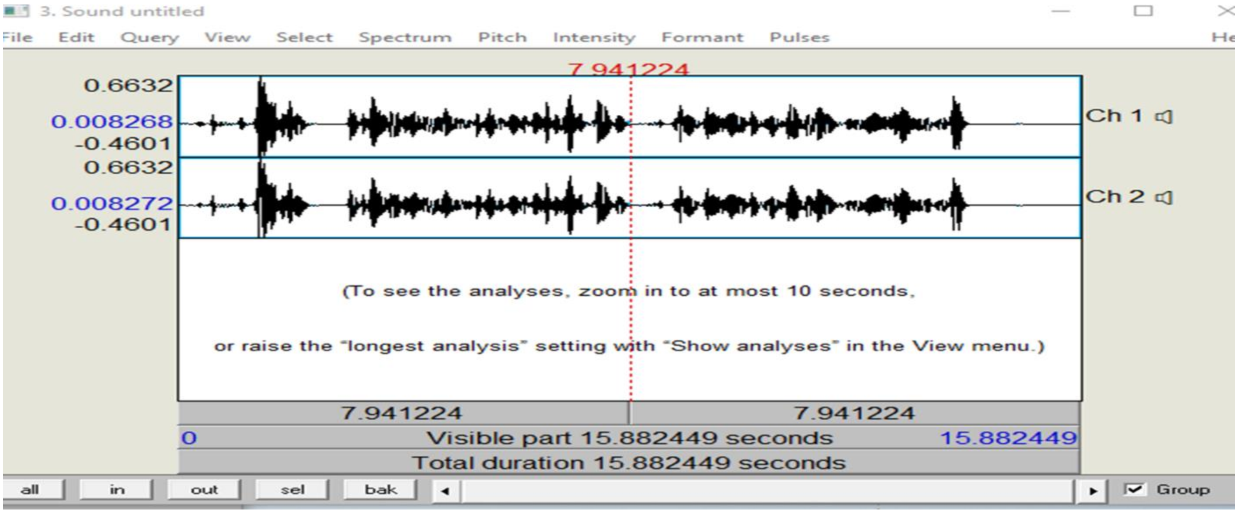


Figure 2. Speech Recording via PRAAT

PRAAT application can be used to probe features of speech such as prosodic aspects (suprasegmental), of which are pitch, tone, stress and rhythm. It analyses accurately pitch evaluation and pitch variation. It also analyses stress and speech intensity (Rupley, Rasinski, Nichols& Paige, 2020,p.102). The following figure represents a process of PRAAT speech analysis.



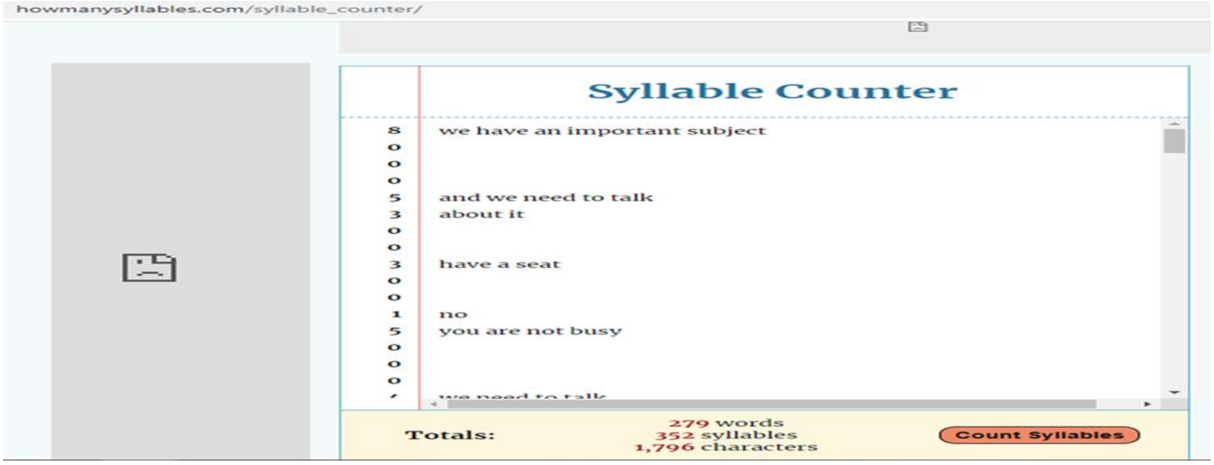
(To see the analyses, zoom in to at most 10 seconds,

or raise the "longest analysis" setting with "Show analyses" in the View menu.)

Figure 3. Speech Analysis via PRAAT

4.4.1.3. Web-based Tools to Analyse Speech Rate as a Fluency Index

To analyse speech rate, researchers need to compute the number of syllables produced within a given time (e.g., 1 minute) in a transcribed text. Hence, the website https://www.howmanysyllables.com ("How Many Syllables", 2021) is an online tool that accurately counts the number of syllables by copying and pasting the transcribed text. It is practical as it saves researchers' time and energy by counting the number of words and syllables in long texts in an automated process. The advantage of such a web-based tool is that the results are shown within few seconds irrespective of the text's length. The following figure is an illustration of how the website functions:



howmanysyllables.com/syllable_counter/

Syllable Counter

```
8   we have an important subject
0
0
0
5   and we need to talk
3   about it
0
0
3   have a seat
0
0
1   no
5   you are not busy
0
0
0
    we need to talk
```

Totals: 279 words
352 syllables
1,796 characters

*Figure 4.*Spoken text analysis using Textinspector.com

4.4.2. Complexity: an Operational Definition

Ellis& Barkhuizen (2005) elucidated that complexity refers to the extent to which learners produce elaborated language. Pallotti (2009) explains that speech complexity demonstrates when an L2 user can produce linguistically, and thus cognitively, more demanding linguistic material (e.g. longer units with more complex embedding elements). Complexity can be either linguistic or lexical (Michel, 2017); while the first relates to the length of utterances, subordination, coordination, and the extent and sophistication of grammatical forms, the second pertains to the text diversity, sophistication, density and variation (Michel, 2017, pp.6-7). In the scope of this paper, two software and web-based language tools that can facilitate the study of the transcribed text's lexical complexity are reviewed.

4.4.2.1. Lexical Complexity and Web-based Tools

Gass, Behney& Plonsky (2013) clarified that linguistic complexity is demonstrated in L2 learners' use of long utterances, subordination, coordination, and the extent and sophistication of grammatical forms. Lexical complexity is measured in terms of:

- Diversity: the size of lexis; gauged using type-token ratio based measures.

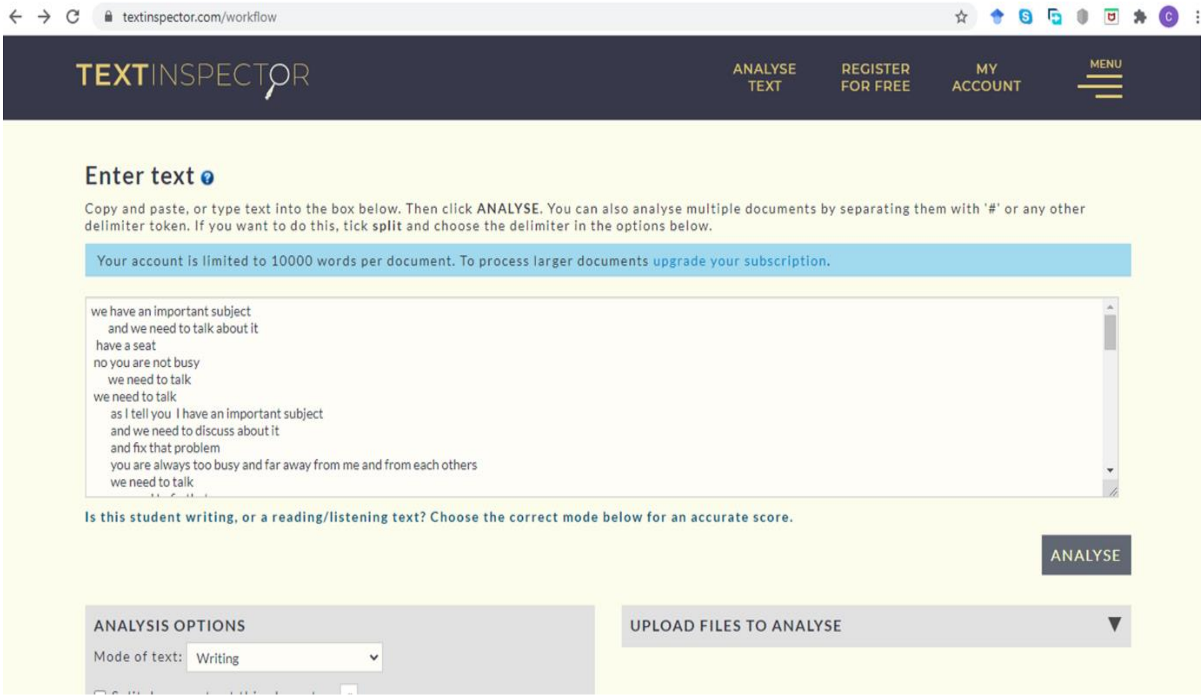
Sophistication: depth of lexis; gauged using frequency measures, for example, of words beyond the 1000 most common words.

- Density: information packaging of lexis; gauged using, for example, the ratio of lexical words per function words.

To measure lexical complexity, https://www.lextutor.ca and https://textinspector.com/ can be used

4.4.2.2.Textinspector.com

Textinspector.com (2021) is a web-based language analysis tool developed by Stephan Bax, an applied linguistic Professor. It provides detailed and thorough information about some texts' features, such as readability, complexity, and lexical diversity. It is trusted by many universities around the world, such as Kings' College London (Textinspector.com 2021). It is used to compute the type-token ratio (which is an index of speech complexity computed by counting the different types of words divided by the total number of words in texts). It is used to analyse text vocabulary to unveil its lexical diversity. It computes the average sentence length and number of syllables in the overall text.



Enter text

Copy and paste, or type text into the box below. Then click **ANALYSE**. You can also analyse multiple documents by separating them with '#' or any other delimiter token. If you want to do this, tick **split** and choose the delimiter in the options below.

Your account is limited to 10000 words per document. To process larger documents upgrade your subscription.

we have an important subject
and we need to talk about it
have a seat
no you are not busy
we need to talk
we need to talk
as I tell you I have an important subject
and we need to discuss about it
and fix that problem
you are always too busy and far away from me and from each others
we need to talk

Is this student writing, or a reading/listening text? Choose the correct mode below for an accurate score.

ANALYSE

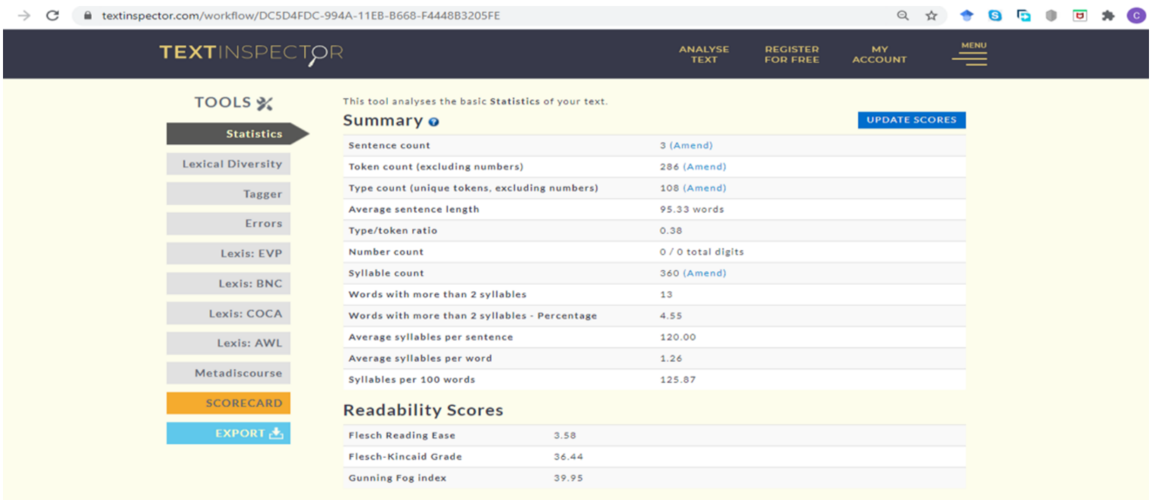
ANALYSIS OPTIONS

Mode of text: Writing

UPLOAD FILES TO ANALYSE ▼

Figure 5. Lexical Complexity Measures via Textinspector.com

As the figure above demonstrates, the researchers paste their transcribed texts to the box available on the website. Alternatively, they can upload long files for analysis. Textinspector.com (2021) web-based language tool, as elucidated by (Fok and Li, 2017.): "contains a number of tools for analysing the frequency of recovery collocations and grade- level determiners for readings. All tools are free for learners and researchers" (p.72). It might be helpful to researchers interested in identifying the progress of learners in terms of using some given lexical items (Fok&Li, 2017).



This tool analyses the basic **Statistics** of your text.

Summary

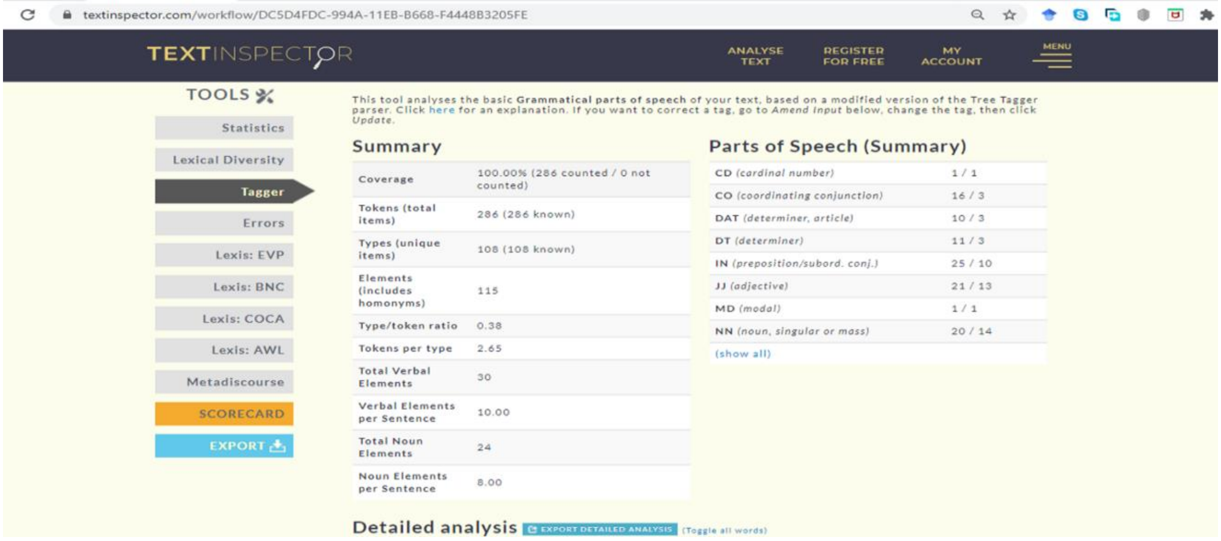
Sentence count	3 (Amend)
Token count (excluding numbers)	286 (Amend)
Type count (unique tokens, excluding numbers)	108 (Amend)
Average sentence length	95.33 words
Type/token ratio	0.38
Number count	0 / 0 total digits
Syllable count	360 (Amend)
Words with more than 2 syllables	13
Words with more than 2 syllables - Percentage	4.55
Average syllables per sentence	120.00
Average syllables per word	1.26
Syllables per 100 words	125.87

Readability Scores

Flesch Reading Ease	3.58
Flesch-Kincaid Grade	36.44
Gunning Fog index	39.95

Figure Lexical Complexity Measures via Textinspector.com

The web-based tool offers an in-depth analysis as to the vocabulary, more precisely its diversity and complexity. Thereby, "This can help us improve our language teaching and learning, help empower students and further our understanding of language." Textinspector.com (2021). It also provides statistics about sentence count, type-token count, average sentence length, the average of syllable or words per sentence, and 100 words. The web-based tool displays parts of speech tagger comprehensively within texts, as the figure below demonstrates (Textinspector.com, 2021).



This tool analyses the basic **Grammatical parts of speech** of your text, based on a modified version of the Tree Tagger parser. Click here for an explanation. If you want to correct a tag, go to *Amend input* below, change the tag, then click *Update*.

Summary

Coverage	100.00% (286 counted / 0 not counted)
Tokens (total items)	286 (286 known)
Types (unique items)	108 (108 known)
Elements (includes homonyms)	115
Type/token ratio	0.38
Tokens per type	2.65
Total Verbal Elements	30
Verbal Elements per Sentence	10.00
Total Noun Elements	24
Noun Elements per Sentence	8.00

Parts of Speech (Summary)

CD *(cardinal number)*	1 / 1
CO *(coordinating conjunction)*	16 / 3
DAT *(determiner, article)*	10 / 3
DT *(determiner)*	11 / 3
IN *(preposition/subord. conj.)*	25 / 10
JJ *(adjective)*	21 / 13
MD *(modal)*	1 / 1
NN *(noun, singular or mass)*	20 / 14
(show all)	

Detailed analysis

Figure 6. Part of Speech Tagger via Textinspector.com

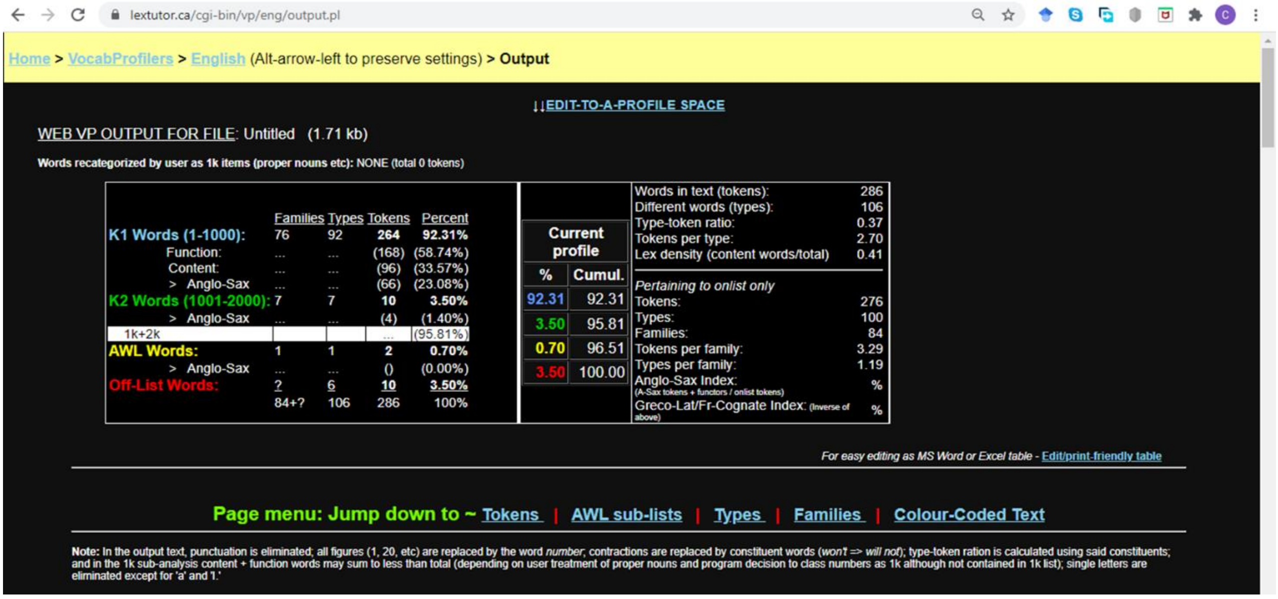
4.4.2.3. lextutor.com

Compleat Lexical Tutor (2021) web-based Lextutor, designed by Cobb in 1995, is an online analysis tool used to analyse language complexity. The lextutor, in the words of Jones and Waller (2015)

> It allows researchers to analyse large amounts of text and produce a database on aspects such as frequency. It also allows producing concordance lines for qualitative analysis of data. Although the aim is to examine language in

> terms of lexis, it can also be used to inform about grammar lexicogrammar and is a very useful piece of software (p.53)

It calculates many lexical complexity aspects such as lexical density, type-token ratio, lexical frequency. The online language tool enables teachers and educators to know their learners' vocabulary and writing level by just pasting their texts on the box and waiting for the statistics to appear in a very comprehensive way. The following is a figure demonstrating the functions of the lextutor web-based tool.


lextutor.ca/cgi-bin/vp/eng/output.pl
Home > VocabProfilers > English (Alt-arrow-left to preserve settings) > Output
↓↓EDIT-TO-A-PROFILE SPACE

WEB VP OUTPUT FOR FILE: Untitled (1.71 kb)

Words recategorized by user as 1k items (proper nouns etc): NONE (total 0 tokens)

	Families	Types	Tokens	Percent
K1 Words (1-1000):	76	92	264	92.31%
Function:	...	...	(168)	(58.74%)
Content:	...	...	(96)	(33.57%)
> Anglo-Sax	...	...	(66)	(23.08%)
K2 Words (1001-2000):	7	7	10	3.50%
> Anglo-Sax	...	...	(4)	(1.40%)
1k+2k			...	(95.81%)
AWL Words:	1	1	2	0.70%
> Anglo-Sax	...	...	()	(0.00%)
Off-List Words:	?	6	10	3.50%
	84+?	106	286	100%

Current profile	
%	Cumul.
92.31	92.31
3.50	95.81
0.70	96.51
3.50	100.00

Words in text (tokens):	286
Different words (types):	106
Type-token ratio:	0.37
Tokens per type:	2.70
Lex density (content words/total)	0.41
Pertaining to onlist only	
Tokens:	276
Types:	100
Families:	84
Tokens per family:	3.29
Types per family:	1.19
Anglo-Sax Index: (A-Sax tokens + functors / onlist tokens)	%
Greco-Lat/Fr-Cognate Index: (Inverse of above)	%

For easy editing as MS Word or Excel table - Edit/print-friendly table

Page menu: Jump down to ~ Tokens | AWL sub-lists | Types | Families | Colour-Coded Text

Note: In the output text, punctuation is eliminated; all figures (1, 20, etc) are replaced by the word *number*; contractions are replaced by constituent words (*won't => will not*); type-token ration is calculated using said constituents; and in the 1k sub-analysis content + function words may sum to less than total (depending on user treatment of proper nouns and program decision to class numbers as 1k although not contained in 1k list); single letters are eliminated except for 'a' and 'I.'

Figure7.Lextutor.com Web-based language Tool Statistics.

Conclusion

The proliferation of database software and web-based language tools has exceedingly assisted researchers in conducting the data gathering and analysis of learners' spoken data. With its wide range of applications, CALL offers EFL researchers ample opportunities to collect data in a less complicated way as it comprises myriad free download applications designed for recording spoken language. Likewise, it paves the way for researchers to analyse some spoken language aspects such as speaking fluency and complexity effortlessly. Nevertheless, no automated software and application are available so far to have automated transcription. The latter has to be manually conducted, i.e., it is still a human-conducted process.

References

1. Baker, P., Hardie, A., & McEnery, T. (2006). A glossary of corpus linguistics. Edinburgh: Edinburgh University Press Ltd.
2. Boersma, P., Weenink, D. (2021). Praat. Retrieved 8 May 2021, from https://praat.en.lo4d.com/windows
3. Boutsen, D, Dvorak, J. D. (2016). MATLAB Primer for Speech Language Pathology and Audiology. San Diego: Plural Publishing, Inc,

4. Breyer, Y.A. (2011). Corpora in Language Teaching and Learning: Potential, Evaluation, Challenges.
5. Caines, A., McCarthy, M.J. & O'Keeffe, A. (2016). Spoken language corpora and pedagogic applications. In F. Farr and L. Murray (Eds) Routledge Handbook of Language Learning and Technology. London: Routledge, pp. 348 - 361.
6. Compleat Lexical Tutor. (2021). Retrieved 8 May 2021, from https://www.lextutor.ca/
7. Du Bois, John W. (2006) VoiceWalker: A discourse transcription utility. The University of California Regents.
8. Ellis, R., & Barkhuizen, G. (2005). Analysing Learner Language. Oxford: Oxford University Press.
9. Fok, W., & Li, V. (2017). Teaching and learning with technology. SingaporeWold Scientific Publishing. CO.Ptc. Ltd.
10. Garside. R, Leech. G, & McEnery.A. (Eds.)(1997), Corpus Annotation: Linguistic Information from Computer Text Corpora, London: Longman.
11. How Many Syllables. (2021). Retrieved 8 May 2021, from https:// www.howmanysyllables.com/
12. Jones, C., & Waller, D.(2015) Corpus linguistics for grammar. Oxon: Routledge.
13. Mackey, A., &Gass, S.M. (2005). Second language research: methodology and design. New Jersey: Lawrence Erlbaum Associates, Inc.
14. Mauranen, A. (2004). Spoken corpus for an ordinary learner. In J.M. Sinclair (Ed) How to Use Corpora in Language Teaching. Amsterdam: John Benjamins Publishing Company.
15. Meyer, C.F. (2009). English corpus linguistics: an introduction. Boston: Cambridge University Press.
16. Michel, M. (2017).Complexity, accuracy and complexity in L2 acquisition.In S. Loewen, M. Sato (Eds.), The Routledge handbook of instructed second language acquisition (50-68). NY.Routledge.
17. Pallotti G. 2009. CAF: Defining, refining and differentiating constructs. Applied Linguistics 30 (4): 590–601. DOI: 10.1093/applin/amp045
18. Rasinski, T., Rupley, W., Paige, D.,& Young.C.(2020). Reading Fluency. Switzerland: MPDI
19. Rupley, W., Nichols, W., Rasinski, T., & Paige, D. (2020). Fluency: Deep Roots in Reading Instruction. Education Sciences, 10(6), 155. DOI: 10.3390/educsci10060155
20. Tavakoli, H. (2012). A Dictionary of research methodology and statistics in applied: linguistics. Tehran: Rahnama Press.
21. Tavakoli, H. (2013). A dictionary of language acquisition a comprehensive overview of critical terms in first and second language acquisition.Tehran: Rahnama Press.
22. Textinspector. (2021). Retrieved 8 May 2021, from https://textinspector.com/
23. Yoon, S. (2020). The Learner Corpora of Spoken English: What Has Been Done and What Should Be Done?. Language Education Institute, Seoul National University. DOI: https://doi.org/10.30961/lr.2020.56.1.29.