Proceedings of the 5th Faculty of Industrial Technology
International Congress Bandung, Indonesia,
September 17-18, 2025
ISSN 2962 – 1798

Developing a Practical Multi-Tool Pipeline for Analysing Synchronous Online Discussions

Corry Caromawati1[a,1]and *Nur Fitrianti* Fahrudin [a]

[a] Information Systems Department, Institut Teknologi Nasional Bandung, Indonesia

Abstract. In classroom-based and computer-mediated communication research, learner talk analysis has traditionally relied on manual transcription and coding, which are labor-intensive and difficult to scale. This methodological paper proposes a systematic, technology-assisted pipeline for analyzing students' English use as a manifestation of willingness to communicate (WTC) during synchronous online breakout room discussions. Drawing on 15 sessions from a first-year English course, the study examined learner talk through multiple dimensions: turn frequency, lexical richness, speaking duration, and communicative functions, using Python, PRAAT, NVivo, and spreadsheet tools. A tailored Python script, LexiTurn, was developed to filter non-target elements and generate accurate token and type counts per turn. Outputs from all tools were integrated in Excel to produce a coherent, multidimensional dataset. The workflow demonstrates how targeted tool selection and basic coding skills, or collaboration with programmers, can enhance methodological rigor, efficiency, and replicability. Beyond its methodological contribution, the study offers pedagogical implications by illustrating how discourse analytics can help teachers identify participation gaps and promote equitable interaction. The proposed pipeline and LexiTurn will be made publicly available to support future classroom discourse research.

1 Introduction

Learners' willingness to communicate in English as a second language (L2 WTC) is widely recognized as a dynamic construct that reflects their readiness to communicate using the target language when free to do so [1, 2]. In most WTC studies, researchers such as MacIntyre and Wang [3] and Ardiansyah, Wijayanto [4] have relied heavily on self-reported L2 WTC scales to capture learners' perceived readiness to use the L2. These measures offer valuable insights into learners' attitudes and intentions but are inherently subjective.

[1] Corresponding author: corry@itenas.ac.id

Since L2 WTC is considered a prerequisite construct of actual L2 use, some researchers have argued that examining learners' observable language production can provide a more direct behavioral manifestation of their WTC. In this perspective, students' real-time use of English—particularly during spontaneous interactions—serves as tangible evidence of their willingness to communicate [5-8]. Importantly, such observable use of the second language also plays a central role in second language acquisition itself [9]. As Swain's output hypothesis [10] and subsequent interaction-based theories emphasize, opportunities to produce and negotiate meaning in the L2 are critical for internalizing linguistic forms, developing fluency, and advancing communicative competence. Thus, tracking students' actual L2 use not only captures their WTC but also reflects the very processes through which acquisition occurs.

In the context of online learning, particularly synchronous platforms like Zoom, tracking how learners use English in real-time interactions can serve as an important behavioral indicator of their L2 WTC. Numerous studies have attempted to quantify L2 WTC by measuring the amount of target language produced, such as turn frequencies [6, 11, 12], and to qualify it through the communicative functions served by utterances (e.g., initiating, responding, elaborating) [6, 7, 13]. These studies suggest that both the quantity and quality of English use offer crucial insights into learners' moment-to-moment engagement and investment in L2 communication.

However, in many classroom-based and computer-mediated communication studies, the data analysis relied heavily on manual transcription and coding [14]. While informative, these approaches are often labor-intensive, time-consuming, and not easily replicable at scale. In addition, tools for quantifying language output (e.g., Microsoft Word or Excel) and coding themes (e.g., spreadsheets) are limited in handling larger multimodal datasets, especially when audio, transcription, and thematic labels must be combined. The need for a replicable, efficient, and systematic method of analyzing language use in synchronous online discussions is therefore clear.

From a mixed methods perspective, such analytic approaches must also uphold methodological rigor, meaning that the data collection and analysis procedures should be clearly aligned with the research questions, transparent, replicable, and appropriate for the nature of the data [15, 16]. As these scholars argue, rigor in mixed methods research is not only about accuracy but also about integrating qualitative and quantitative strands in a coherent, complementary way to strengthen the study's interpretive power.

This methodological study addresses that gap by offering a comprehensive analysis pipeline for investigating students' WTC in Zoom breakout room discussions. Drawing from turn-by-turn transcriptions and audio recordings, the study quantifies individual learners' English use (token and type counts, duration per turn) and qualifies their communicative purposes through thematic analysis. To handle these complex tasks, the study evaluates the strengths and limitations of various tools (e.g., Excel, Google Sheets, NVivo, Python-based scripting), ultimately proposing a systematic method to perform this multiple analysis process. The study contributes a replicable methodological framework for researchers and educators seeking to examine L2 WTC through synchronous digital interactions.

2 Methodology

This study examined a method implemented in a study on Indonesian university students' English language use in Zoom breakout room discussions to investigate their Willingness to Communicate (WTC) during online instruction. The analysis was based on a corpus of transcribed student interactions, with a combination of quantitative and qualitative analyses applied to measure linguistic engagement and communicative functions.

2.1 Dataset

The dataset comprised turn-by-turn transcriptions and audio recordings from 15 breakout room (BR) sessions conducted across six synchronous online classes over one semester. These BRs were coded according to the week of the meeting and the BR number. For instance, a BR recorded during the first meeting and labeled as the first group in that session was coded as "BR 1.1" (Week 1, Breakout Room 1). This coding system was applied consistently across all recordings to enable efficient cross-referencing and data organization. The codes facilitated systematic tracking of patterns in language use and communicative behaviors across different groups and time points throughout the semester.

Each transcription captured approximately 25–45 minutes of interaction and followed the detailed multimodal transcription convention by Rosenbaun, Rafaeli [17], which integrates elements from Jefferson [18] and Rintel [19]. The transcripts included paralinguistic and multimodal cues such as stress, intonation, overlap, pauses, and indications of gestures or chatbox use, supporting the nuanced analysis of verbal and nonverbal aspects of student communication.

Each breakout room yielded a minimum of 50 turns. A detailed table presenting the total and per-room number of speaker turns is provided in Table 1 to illustrate the scope of the data. Given the amount and density of the data, coupled with the fine-grained level of analysis required (e.g., word-level counting, turn-by-turn segmentation, function labelling), a systematic, accurate, and practical approach to data processing was necessary. Manual processing alone would be error-prone and inefficient. Thus, a range of analytical tools was incorporated.

Table 1. The breakdown of the number of turns in each breakout room

Breakout room Codes	Durations (mins)	Number of turns
1.1	25	397
1.2	25	422
2.1	45	121
2.2	45	356
3.1	35	254
3.2	35	503
4.1	40	243
4.2	40	254
4.3	40	143
5.1	35	220
5.2	35	163
5.3	35	50
6.1	30	271
6.2	30	398
6.3	30	358
TOTAL TURNS		4153

2.2 Analytical Needs

The primary goal was to analyze students' willingness to communicate in English through observable behaviors in each turn. To do so, the study identified five major analytical needs:

1. Quantifying English language use per speaker turn, including:
   - Number of turns per student
   - Number of English words per turn
   - Exclusion of non-target language content such as Indonesian words, hesitation markers (e.g., *uh, um*), filler expressions, and noise (i.e. symbols created in the transcription)
2. Measuring speaking duration per turn to reflect temporal engagement in interaction.
3. Labelling communicative functions per turn to capture students' discourse-level contributions, such as initiating, elaborating, inviting, responding, and concluding.
4. Managing and integrating multiple layers of analysis results using an efficient and organized system.

ISSN 2962 – 1798

2.3 Tools Exploration

Given the volume and complexity of the dataset, particularly its multimodal and turn-by-turn nature, the researcher explored a range of technological tools to determine the most suitable ones for analyzing students' English use, communicative engagement, and interactional patterns. The exploration process was informed by three main considerations: *analytical needs*, *practicality*, and *budget constraints*.

Analytically, the study required tools that could (1) quantify English language use (turn frequency, word count per turn, token type per turn), (2) measure speaking duration per turn to reflect temporal engagement, (3) support thematic coding to identify communicative functions (e.g., initiating, elaborating, inviting, concluding), and (4) support efficient book-keeping and visualization of findings.

Practically, the tools needed to be user-friendly, especially for handling large textual and audio datasets. The researcher trialed multiple options for each analytical task, ranging from free, readily accessible tools like Microsoft Excel and Word, to more advanced and specialized software such as NVivo [20], ELAN [21], Audacity [22], ChatGPT [23], and PRAAT [24]. For instance, while Excel supported initial sorting and basic quantification, it lacked accuracy and efficiency in calculating word-level details or conducting advanced qualitative coding. NVivo, although commercial, was found more appropriate for managing thematic coding due to its streamlined labeling and retrieval features. Similarly, PRAAT was selected for measuring the duration of student speech per turn, given the researcher's familiarity with its interface and its precision in acoustic analysis.

Ultimately, the final combination of tools used in this study balanced cost-effectiveness with methodological robustness. The decision-making process and rationale for tool selection are discussed in detail in the *Results and Discussion* section, where the strengths and limitations of each tool in meeting the analytical demands of the study are evaluated based on actual implementation.

3 Result and Discussion

The analysis of students' speech data required a range of tools to accommodate the dataset's extensive volume and micro-level analytical demands—such as identifying lexical choices, temporal features of speech, and communicative functions. Each tool was selected based on its ability to support specific analytical tasks while ensuring accuracy, practicality, and efficiency.

3.1 Tool Selection to Execute the Analysis

A custom Python-scripted tool, named LexiTurn, was developed and implemented in Google Colab to process the transcription files. This tool was specifically designed to filter out unwanted items—including background noise markers, proper nouns, non-word tokens (e.g., "uh," "erm"), and Indonesian lexical items—so that the resulting word counts accurately represented students' spoken English. Beyond raw counts, LexiTurn also generated lists of unique words and calculated token types, providing measures of lexical richness and enabling further layers of linguistic analysis. By automating these otherwise time-consuming steps, the tool ensured greater accuracy, consistency, and replicability in handling large datasets.

Next, PRAAT was used to measure the duration of each turn. Audio recordings were visualized as spectrograms, allowing the researcher to isolate and play back individual segments. The tool's precise time-stamp interface enabled accurate measurement of speaking time, which was manually entered into the corresponding Excel sheets. This step aligned temporal data with lexical data from the previous stage, adding a crucial time-based dimension to the analysis. Figure 3 illustrates this process, showing a sample analysis of a student's utterance: "They said it's an extension from an older game, but I never played it." The duration of this utterance is displayed in the PRAAT interface, visible in the time markings below the window, which was zoomed out to provide a clear view of the full segment.

Table 2 summarizes the analytical tasks, tools considered, and final tools selected in the data analysis process. For turn-by-turn quantification of students' spoken English—measuring the number of words or tokens per student turn—LexiTurn in Google Colab was chosen over manual approaches using Microsoft Word or Excel, which were inefficient and error-prone for handling large volumes of data. Similarly, PRAAT was selected for measuring speech duration per turn due to its spectrographic precision and user familiarity, outperforming other tools like Audacity and ELAN.

Table 2. Summary of Analytical Tools Selection and Rationale for Each Task

Analytical Task	Tools Considered	Final Tool Selected	Rationale
Turn-by-turn quantification (e.g., number of words/tokens per turn)	Microsoft Word, Excel, Python scripts	Python scripted tool (LexiTurn)	Excel was initially used but found inefficient and prone to human error for high-volume data. Python allowed automated, replicable processing.
Measuring duration of student speech per turn	Audacity, PRAAT, ELAN	PRAAT	PRAAT was selected for its precision in identifying start and end points of speech and measuring speaking time at the turn level. The researcher had prior familiarity with its interface.
Thematic analysis (e.g., identifying communicative functions such as initiating, elaborating, inviting, concluding)	MS Word , Excel, NVivo, ChatGPT	NVivo	NVivo was selected for its efficient coding and data visualization capabilities. Excel was initially trialed but quickly became unmanageable. ChatGPT provided inconsistent output.

Data documentation and integration across tools	Excel, MS Word	Excel	Despite limitations, Excel was retained for central book-keeping—logging codes, timestamps, word counts, and thematic categories in one place.

For the thematic analysis of students' communicative functions, NVivo emerged as the most suitable tool. Although manual coding in Word and Excel was trialled, these approaches were unsustainable for iterative thematic work. ChatGPT (2022 version) was also explored during the exploratory phase, offering quick suggestions for labelling and grouping themes. However, its outputs lacked consistency, transparency, and replicability—key requirements in rigorous qualitative analysis. As a result, NVivo was ultimately selected for its structured coding system and visualisation capabilities. Finally, Excel served as the central platform for documentation, integrating various outputs—such as turn-level word counts, speech durations, and thematic codes—despite its limited analytical functionality.

3.2 Tool Implementation and Affordances in Data Analysis

The final analytical workflow employed a set of complementary tools, each addressing a specific component of the dataset—text, audio, and qualitative codes. This multimodal integration allowed for a comprehensive examination of students' English use and communicative behaviour. This workflow is illustrated in Figure 1 below.

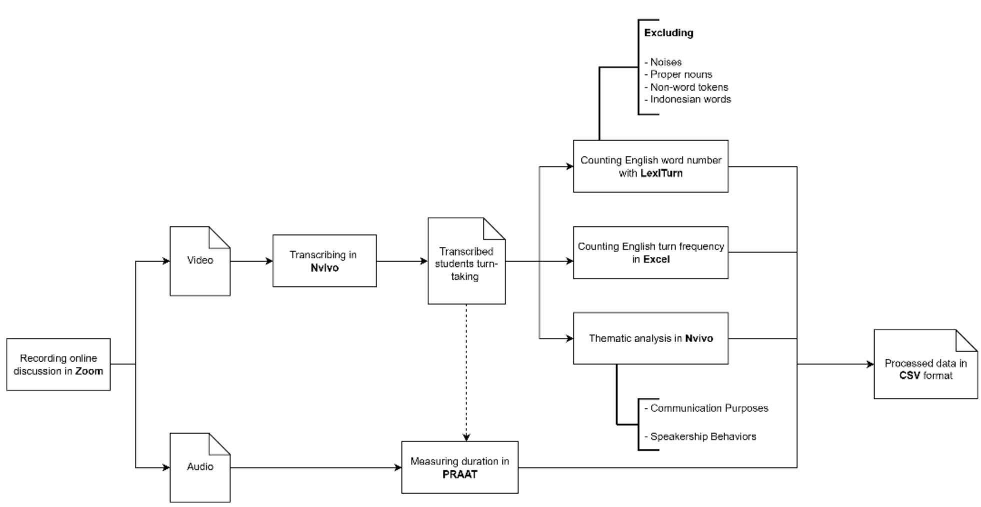


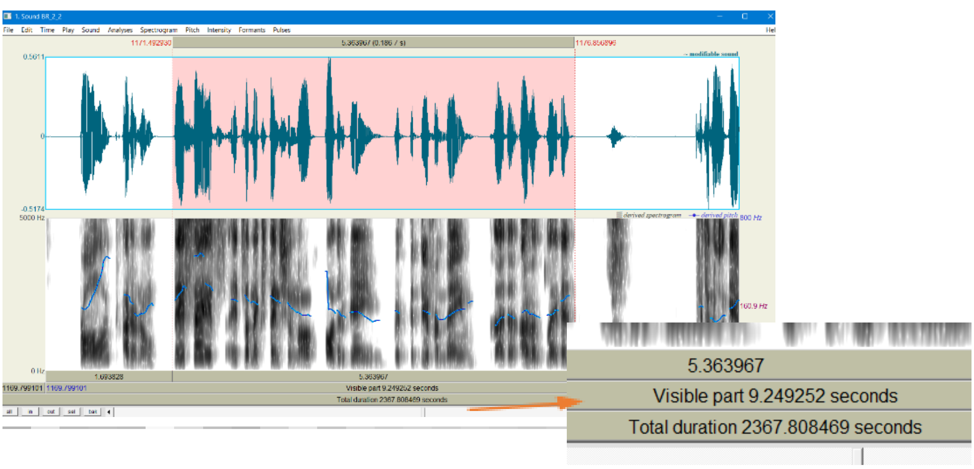
Fig. 1. Workflow of Integrated Tools for Multilayered Analyses of Students' Talk in LexiTurn

The analysis began with LexiTurn, a Python-scripted tool developed and deployed in Google Colab for this study. LexiTurn was designed to process `.csv` transcription files systematically. It filtered out background noise markers, proper nouns, non-word tokens (e.g., “uh,” “erm”), and Indonesian words, ensuring that the resulting word counts accurately reflected students’ spoken English. In addition, it listed unique words and calculated token types per speaker turn, offering indicators of lexical richness and enabling subsequent layers of linguistic analysis. By automating these processes, LexiTurn provided a replicable, efficient, and scalable solution to token-level quantification, which would otherwise be labor-intensive and error-prone if conducted manually. The pseudocode underpinning LexiTurn’s design is presented in Figure 2, illustrating its core filtering and counting functions.

```
BEGIN

IMPORT re, nltk
DEFINE stopwords = [list of filler words like "uh", "oh", "((laugh))", etc.]
DEFINE punctuations = string containing all punctuation symbols

GET all text from dfData["Konten"] → contentAll

INITIALIZE empty lists:
   listPunct         // cleaned text without punctuation
   listUniq          // text with only unique words
   totalWords        // total number of words (excluding stopwords)
   uniqueTokenCount  // number of unique words
   tokenFrequency    // word frequency per sentence

FOR each sentence in contentAll:
   1. REMOVE stopwords
   2. REMOVE punctuation
   3. CONVERT to lowercase
   4. APPEND cleaned text to listPunct
   5. SPLIT text into words → COUNT total → ADD to totalWords
   6. REMOVE duplicate words → JOIN into string → ADD to listUniq
   7. COUNT number of unique words → ADD to uniqueTokenCount
   8. TOKENIZE words using nltk → CALCULATE frequency → ADD to tokenFrequency

ADD all results as new columns in dfData:
   - "listPunct"
   - "listUniq"
   - "totalKata" (totalWords)
   - "Uniq Token" (uniqueTokenCount)
   - "Token List" (tokenFrequency)

END
```

ISSN 2962 – 1798

PRAAT was then used to measure the duration of each turn. Audio recordings were visualized as spectrograms, allowing the researcher to isolate and play back individual segments. The tool's precise time-stamp interface enabled accurate measurement of speaking time, which was manually aligned with the lexical counts produced by LexiTurn. This integration added a crucial temporal dimension to the dataset.



For the qualitative component, NVivo was applied to a representative subset of three BR sessions selected to reflect diverse interactional dynamics. NVivo enabled systematic, iterative coding of communicative functions, leading to the emergence of six primary themes: social, on-task, off-task, technical, language-related, and classroom management. While a detailed account of this process is discussed in a forthcoming publication, the final themes were integrated into the main dataset using Excel's drop-down validation feature. This enabled the researcher to label each turn thematically across all 15 sessions.

The final stage involved generating descriptive reports and visualizations within Excel. The unified dataset—now incorporating lexical counts from LexiTurn, turn durations from PRAAT, and thematic codes from NVivo—allowed for efficient production of summary tables and comparative visuals. These outputs illuminated patterns such as fluctuations in English use across sessions or shifts in communicative focus over time.

ISSN 2962 – 1798

A key strength of this workflow was its inter-tool integration. While each tool served a specialized purpose, Excel acted as the central hub—bringing together outputs from LexiTurn, PRAAT, and NVivo for triangulated analysis. This design maintained methodological transparency while enabling both breadth and depth of interpretation.

It is worth noting that tools such as LexiTurn and PRAAT—though offering superior precision—came with steeper learning curves. However, their benefits in terms of automation, replicability, and granularity justified the initial investment in training. In contrast, while Excel and Word were user-friendly, they proved insufficient for high-volume or complex analysis, reinforcing the need for task-specific solutions in applied linguistics research.

In sum, the selection and implementation of these tools were informed by both empirical testing and practical considerations. The final toolchain enabled scalable, reliable, and nuanced analysis of students' spoken English in online collaborative settings. These methodological choices—particularly the development and application of LexiTurn—were central to producing valid and interpretable findings.

4 Conclusion

This study introduced a systematic pipeline for analyzing students' English use in synchronous online breakout room discussions. Drawing on 15 sessions across six class meetings, the analysis examined learners' language use through turn frequency, lexical richness, speaking duration, and communicative functions. The workflow integrated tools such as PRAAT, NVivo, Excel, and LexiTurn—a Python-scripted tool developed for this study—each selected according to analytical needs, practicality, and budgetary considerations.

A key finding is the advantage of linguists acquiring basic coding skills or collaborating with programmers. Tailored tools such as LexiTurn can streamline analysis, improve accuracy, and enhance replicability. We are currently developing and preparing to publish LexiTurn so that other researchers working on turn-taking can adopt and adapt it for their own projects.

Methodologically, the study contributes to applied linguistics and WTC research by offering a replicable framework for multi-layered classroom discourse analysis. It demonstrates how integrating qualitative and quantitative tools can enrich investigations of language use in digital interactional settings. Pedagogically, the findings also suggest that teachers can benefit from using similar analytical approaches to track learners' engagement, identify participation gaps, and design more targeted classroom interventions.

Future studies should strengthen trustworthiness through intercoder checks, leverage emerging AI-based tools for automatic thematic analysis, and incorporate specialized software for turn-taking segmentation to capture interactional dynamics more effectively.

References

1. McCroskey, J.C. and J.E. Baer, *Willingness to communicate: The construct and its measurement.* 1985.
2. MacIntyre, P., et al., *Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation.* The Modern Language Journal, 1998. **82**(4): p. 545-562.
3. MacIntyre, P. and L. Wang, *Willingness to communicate in the L2 about meaningful photos: Application of the pyramid model of WTC.* Language Teaching Research, 2021. **25**(6): p. 878-898.
4. Ardiansyah, S.A., A. Wijayanto, and A. Asib, *Dynamics of Students' willingness to Communicate in English during an Online Discussion.* International Journal of English Language Studies, 2020. **2**(5): p. 11-20.
5. Suksawas, W., *A sociocultural study of EFL learners' willingness to communicate*, in *Education*. 2011, University of Wollongong.
6. Zhong, Q.M., *Understanding Chinese learners' willingness to communicate in a New Zealand ESL classroom: A multiple case study drawing on the theory of planned behavior.* System, 2013. **41**(3): p. 740-751.
7. Cao, Y. and J. Philp, *Interactional context and willingness to communicate: A comparison of behavior in whole class, group and dyadic interaction.* System, 2006. **34**(4): p. 480-493.
8. Ducker, N.T., *Bridging the gap between willingness to communicate and learner talk.* The Modern Language Journal, 2022. **106**(1): p. 216-244.
9. Larsen-Freeman, D., *Reflecting on the cognitive–social debate in second language acquisition.* The Modern Language Journal, 2007. **91**: p. 773-787.
10. Swain, M., *The output hypothesis and beyond: Mediating acquisition through collaborative dialogue.* Sociocultural theory and second language learning, 2000. **97**(1): p. 97-114.
11. Syed, H. and I. Kuzborska, *Dynamics of factors underlying willingness to communicate in a second language.* The language learning journal, 2020. **48**(4): p. 481-500.
12. Yashima, T., P. MacIntyre, and M. Ikeda, *Situated willingness to communicate in an L2: Interplay of individual characteristics and context.* Language Teaching Research, 2018. **22**(1): p. 115-137.
13. Nematizadeh, S. and Y. Cao, *Investigating willingness to communicate in synchronous group discussion tasks: One step closer towards authentic communication.* International Review of Applied Linguistics in Language Teaching, 2023(0).
14. Satar, H.M. and N. Özdener, *The effects of synchronous CMC on speaking proficiency and anxiety: Text versus voice chat.* The Modern Language Journal, 2008. **92**(4): p. 595-613.
15. Harrison, R.L., T.M. Reilly, and J.W. Creswell, *Methodological rigor in mixed methods: An application in management studies.* Journal of mixed methods research, 2020. **14**(4): p. 473-495.
16. Johnson, J.L., D. Adkins, and S. Chauvin, *A review of the quality indicators of rigor in qualitative research.* American journal of pharmaceutical education, 2020. **84**(1): p. 7120.

17. Rosenbaun, L., S. Rafaeli, and D. Kurzon, *Participation frameworks in multiparty video chats cross-modal exchanges in public Google Hangouts.* Journal of Pragmatics, 2016. **94**: p. 29-46.
18. Jefferson, G., *Glossary of transcript symbols with an introduction.* Pragmatics and Beyond New Series, 2004. **125**: p. 13-34.
19. Rintel, S., *Video calling in long-distance relationships: The opportunistic use of audio/video distortions as a relational resource.* The Electronic Journal of Communication/La Revue Electronic de Communication (EJC/REC), 2013. **23**.
20. International, Q., *NVivo*. 2022.
21. Psycholinguistics, M.P.I.f. 2022.
22. Team, T.A., *Audacity*. 2022.
23. OpenAI. *ChatGPT*. 2022; November 30 Version:[Available from: https://chat.openai.com/.
24. Boersma, P. and D. Weenink, *Praat: Doing phonetics by computer*. 2020.