Speech Analysis Using Praat Software in between Native and Nonnative speakers of English.

Abderrahim Bouderbane
University of Abdalhafid Boussouf Mila, Faculty of Letters and Languages,
Department of English, Algeria.
Abderrahim.bouderbane@gmail.com

Abstract

This research investigates the characteristics of speaking fluently by analyzing its main components like pauses, filled pauses, and hesitations. These components are very important in speaking, and they are used generally to generate ideas, plan what to say next and organize the content. The sample in this research is composed of 20 participants split between 10 native speakers and 10 non-native speakers who were selected randomly. The nature of this research determined the selection of the participants, as they should be proficient in English even non-native speakers of English. This research is to compare samples of recorded speech samples of native and non-native speakers of English to distinguish between pauses, fillers and hesitations. The samples are evaluated using software called Praat. This analysis is based mainly on three main steps which are counting the total time of speaking, then we compare between the pauses' length of native and non-native speakers on an adjustable level of up to *5000 or 6000 H_z*. This is generally the highest level of pauses length in Praat measurements. After that, speech rate and hesitation phenomena (pauses, hesitations and fillers) are compared to using the Pearson product-moment correlation coefficient (r). The results indicate that native speakers engage in genuine real life communication with the support of both language knowledge and topic knowledge and produce speech automatically. Non-native speakers on the other hand, require different proceduralization of information and they produce speech in stages, and they require a lot of time to think about what to say next and produce more hesitation phenomenon.

Keywords: native and non-native speakers, fluency, hesitation phenomenon, Praat, Speech analysis.

1. Measuring Fluency with Praat

The measuring of fluency was always a distinguishing characteristic in between speed of delivery and hesitation phenomenon (Seito, et, al 2018, Suzuki and Kormos, 2019). Such recent researchers account for the scores attributed in the speech rate against the rate of all types of hesitation phenomenon. Other researchers, tried to relate fluency to other factors like problem solving (Peltonen 2017), paired conversational tasks (Williams 2024), temporal and non'temporal features (Thomson, 2015), and even native and non-native speakers (.Guz, 2015.). De Jong was among the pioneers to use Praat softoware to measure speed fluency and breakdown fluency automatically. In a recent article, De Jong et, al (2021) tried to measure automatically aspects of L2 fluency in between Dutch and English speakers. New Praat scripts of filled pauses are tested in this research against old scripts for accuracy. The results indicated promising measures which can be used in the assessment of proficiency of fluency for L2 learners. Old scripts relied heavenly on measuring speed of speech and and silent pausing , and they helped on detecting pauses easily (De Jong and Wempe 2009). However, research pertaining to fluency proficiency is scares in second language contexts, and even in official language tests (IELTS, TOEFL and Cambridge), examiners rely on different scoring rubrics which give direct instructions about do's and don't of language proficiency without no clear cut distinction about the main components of each skill..

2. The Components of Fluency

The testing of fluency is judged from the performance of the learner with the other interlocutors in different social settings. The assessor generally divides what has been said, to understand what is meant, or to find out a criterion for evaluating connected speech. The point is that, this division allows the assessor to diagnose the learners' mistakes and their positions, and even categorize the mistakes into different types, and each category is assessed alone. Many researchers including Götz, S. (2013), Kormos & Dénes, (2004), Tavakoli., Nakatsuhara, & Hunter, (2017) focused on dividing fluency up into hesitation phenomena components which are included in speed of delivery.

2.1. Speed of Delivery

Speed of delivery is identified by the informants and their capacity to produce words per minute. Freed (1995) argues that speed of delivery has something to do with exposure and repetition. Learners may acquire language very easily if they are exposed to it, and the same expressions are used repeatedly either by the learner himself or by other people like classmates and the teacher in the case of classroom environment. It is argued by researchers in second language acquisition that speed of delivery is a characterization of native-like speaking, and that native speakers' language production is automatic, it contains fewer pauses and interruptions. In a study Lennon (1990) tried to prove that language production and speed of delivery are highly related to exposure, he studied the improvement of four German students, who resided in England for a period of one year, Lennon noticed three important indices of improvement in the students' language proficiency which are: the quality of speech, the rate of speech and the fewer number of pauses used to separate units of speaking. Speed of delivery is based on the rate of speech, the decrease of pauses like (ehm, err, ah) and the increase of unit production mainly per minute.

1.2 Hesitation Phenomena

The hesitation phenomena represents a number of factors which influence the production of language and speech rate in general, these factors are: pauses, fillers, hesitations, repetitions, luck of discourse markers and sentence connectives. Skehan (2001) considered these aspects as the most comprehensive picture of fluency performance since it is a combination of what should be measured in fluency.
.

1.3 Pauses

A pause may occur to indicate the end of the turn especially when the utterance ends in low key, and is associated with fillers like: um, er, or uhu. In some other cases pauses are used to plan for what to say next when the idea is in the mind but the learner is still looking for the right words to express it clearly (Bouderbane, 2021). Fulcher (2003:101) explains that pauses are used to add examples, counter-examples, or reasons to support a point of view; he said "Pauses are sometimes used as an oral parenthesis before adding extra information to an argument or point of view, or break up a list of examples." They are actually measured with seconds and they stretch as longer as a second, out of the number of words per minute (Oppenheimer, 2000)

1.4 repetitions

Repetitions occur with repeated syllables, words or phrases the repeated word does not add any propositional content of the utterance. Generally speaking, a pause occurs between the word and its repetition which stands for another missing or unfound word. Heike (1980) draw a distinction between prospective and retrospective repetition. Prospective are classified as those introduced because of perceived upcoming difficulty for the speaker. While retrospective occur when the speaker detected that a problem has already occurred. The repletion here is needed to establish fluency of speech. Skehan (2009) claims that this is a kind of repair fluency as it allows speakers to repairs speaking in correspondence to language proficiency.

1.5 Hesitations

Hesitation in speech is always marked by fillers, pauses and prolongations of words these features are remarkably common in most continuous speaking. These features affect both the processing of speech and the lasting representation of the material. Hesitation is due to the increase in the difficulty in conceptualizing utterances when hesitations like word prolongations are produced very frequently (Schnadt and Corley, 2006).

3. Procedural Fluency

Fluency in not only tied to the number of utterance produced and the number and characteristic of hesitations produced to separate the utterances and to correct mistakes. It includes also the capacity to transform the message from ideas into words, and utterances joined together. Towel, et al (1999) considered proceduralization as part of the linguistic knowledge Planning in speaking is not that easy, the speaker does not have much time to plan what to say next, specifically when the mind is totally engaged in connecting the lexical items together, to construct new knowledge from the acquired background knowledge. As it has been mentioned before, Levelt (1989) explained what happens when the information is processed in the mind using schema knowledge, planning the information, and transforming it into lexical items. At the beginning of the planning stage, the speaker conceptualizes the ideas in an attempt to organize them coherently, and to choose one option to say the right thing in the right way. Then the speaker formulates his schema

knowledge according to the goals and objectives of the intended message. At the end, fluency reflects planning and organization to organize the content coherently.

2 Methods

2.1 Sampling

The sample in this research is composed of 20 participants split between 10 native speakers and 10 non-native speakers who were selected randomly. The nature of this research determined the selection of the participants, as they should be proficient in English even non-native speakers of English. Therefore, 10 Algerian teachers of English as a Foreign Language were selected to analyse their speech extracts. Concerning native speakers, the researchers ceased the

opportunity of being in Spain to attend a conference about Teaching English as A Second Language at the University of Granada. The researcher selected 10 native speakers and they are mixed 3 British, 3 Americans, 2 Canadians, 1 Irish, and 1 Swedish. The researcher asked them general questions about teaching and even testing. The point behind the discussion is to measure the length of the pauses and not topic knowledge.

2.2 Tools

First, the samples are evaluated using software called Praat. This software is mainly used in phonetics to describe pitch, stress and intonation. It can also be used to calculate the total time of speaking, and demonstrate the pauses graphically as they are produced by speakers. Accordingly, this software is very useful when it comes to distinguishing between pauses and their length in the total time of speaking. The demonstration of pauses in Praat is done on wave forms and spectrograms, which show the length of pauses and pitch when fillers are used in pauses. The last option is called 'fillers annotations' and it describes the sounds made, the pitch and the intensity of the sounds. The measurement of one pause is made reliable by Praat as far as the measurement is adjustable at all levels (Styler 2013). The adjustment can start from a *0.00H$_z$* to *5000 or 6000 H$_z$*. This software cannot tell where one word start and where it ends. As such, the researcher needs often to segment sound fillers with such information when using any sort of automated information. This is generally done by creating a textgrid annotation in a text file, which is saved separately. These annotations are composed of different tiers which are marked either by intervals or specific points in the file.

It is also worth mentioning that Praat allows the importation of graphics and scripts of long stretches of discourse which extend up to 20 minutes in total. As though, this gives researchers in the field of computational linguistics enough data to demonstrate their analysis of continuous speech. It can also be adjusted to treat long stretches of discourse in an option called voice analysis. This option can be found under the 'puls' menu. The menu contains a number of measures for quantifying irregularities in the duration (jitter), and amplitude (shimmer) of individual cycles and they are indicated by dark blue puls lines in the waveform signal (Van Lieshout 2003).

In this research, the Praat software is basically used for two main reasons. The first is to count speech rate of the speakers to distinguish in between the total time of speaking and the actual speaking time. The second is to count the pauses produced by the speakers by measuring their length, and their frequency in the total time of speaking.

The results found in the first method of the experiment will be compared with native speech analysis which is achieved in the same way as it is evaluated in the same way as those recorded speeches of non-native speakers. The analysis in this research will focus on the productions speech (pitch, stress and intonation), and hesitation phenomena (fillers, pauses and hesitations). Rosenberg (2009) found a high rate of pitch and intonation in a production study of native speakers. While Sagalowitez (2010) found that pauses and hesitations are more frequent and longer in non native speaking production. However, there are not many studies conducted to distinguish between pauses, fillers and hesitations. As a result, this research is an attempt to compare automatic (machine) speech processing of native and non-native productions.

Second, a correlation between speech rate and hesitation phenomena (pauses, fillers and hesitations) is calculated between the variables on the basis of manual calculation of the Pearson's correlation coefficient.

Table 1. Summary of measures of fluency

De Jong and Perffetti (2011)	Fluency	Phonation time ratio	Speaking time divided by total time	Global
De Jong and Perffetti (2011)	Fluency	Mean Length of pause (F2)	Average length of filled and unfilled pauses (> 200 ms)	Fluency breakdown

| 4 | **De Jong and Perfetti (2011).** | Fluency | Mean length of fluent run (F3) | Average number of syllables in utterances bounded by pauses> 200 ms | Fluency proceduralization |

Before calculating the correlation coefficient between the variables hesitation phenomena and speech rate, a number of other calculations ought to be made to compare the difference between phonation time (free of pauses time), and the total time. The latter is made to calculate the total time of pauses made in speech production. Therefore, a number of fluency will be adapted from the aforementioned researchers in the field of language fluency and speech processing.

Overall, Praat is used in this case to calculate speech rate and the number and length of pauses produced in the speech rate. Then SPSS analysis was used to calculate the correlation between the speech rate with pauses, hesitations, fillers and repetitions. Basically, the statistics we have calculated in this research are the Pearson Correlation coefficient (R),

the standard deviation, the degree of freedom and the t-test.

3   Results

The variance of native and non-native speech rate and pauses results are as the following:
- Native speakers' speech rate is determined in between 4.06 and 5.58 minutes.
- Nonnative speakers' speech rate is determined in between 4.49 and 5.49 minutes.

This is of course counted after the subtraction of the total time of speaking which ranges from 7 to 9 minutes as suggested by the researcher.
- Native speakers pauses' range from .38 to .46/minute
- Nonnative speakers' pauses range from .46 to .68/minute. Table 2 demonstrates the results speech rate and pause results.

Table 2. Summary of measures of fluency

	Native speakers		Non-native speakers	
Participants	N of Pauses	Speaking Time	N of Pauses	Speaking Time
1	.43	4,12	.54	5,16
2	.38	4,17	.48	5,43
3	.40	5,34	.46	4,49
4	.46	4,06	. 66	5,16
5	.44	4,44	.46	5,33
6	.45	4,2	.52	5,25
7	.42	5,17	.55	5,19
8	.44	5,36	.65	5,37
9	.47	5,58	.68	5,17
10	.46	5,1	.51	5,31

A number of measurements and statistical analysis including coefficient correlation R, t-test, Degree of freedom df and standard deviation are calculated and summarized in table 3.

Table 3. summary of comparative results after statistical measures

Native speakers	Nonnative speakers
R_041	R_0.206
df_8	Df_7
t-test_0.53	t-test_0.22
SD_25.87	SD_27.12

4 Analysis

In the results above, we demonstrate the participants' (both native and nonnative) speaking time and number of pause (including hesitations, filled pauses, and fillers) which are calculated by Praat software. These calculations are used here to compare the means of native and non-native participants and conduct a thorough analysis of the speech rate.

The results indicate how strong the correlation of speech rate and pauses between native speakers achievements in the analysis of speech rate. Native speakers speech production is less, and it is accompanied with less frequent hesitation phenomenon including filed pauses and hesitations. Native speakers produce pause which range in between 75Hz, and extends to 150 or 200Hz in Praat measures (Bosker, et., al, 2014). Regarding nonnative speakers, they produce more

speech (after the subtraction of the total time of speaking), and they produce filled pauses and hesitations (SD=27.12) which range from 100Hz up to 400 or 500Hz in Praat measures.

Native speakers speech analysis revealed significant results in terms of the pauses (0.12) repetitions (0.15), fillers (0,17) and hesitations (0.22). This indicates a tendency to use hesitations and fillers more than repetitions. They like to fill speaking with talk in order to think about what to say next.

Nonnative speakers, on the other hand, maintained a lot of difficulties in terms of pauses (0.8), hesitations (0.25), against no noticeable usage for fillers (0.01) and repetitions (0.01).

Nonnative speakers require time to produce fluent stretches of spoken discourse. As a result, other aspect are considered including proceduralization of information (0.12), planning what to say next (0.19) and speed of delivery

(0.09). The speed of delivery is low, this creates a more complex process in language knowledge and topic knowledge to plan what to say next and process it in their cognitive minds. The Praat analysis of nonnative speech excerpts shows variance in the length of waveforms produced. The analysis was set at 5000 Hz as maximum of pitch intensity, while the speakers produced up to 27 pauses/minute. One pause is produced in around 200 ms (milliseconds), and and it occurs 27 times every minute.

5 Discussion

The adaptations in this research are made to characterize fluency in speaking of native speakers, and non-native speakers of English. This characterization is made to establish grounds between aspects of the hesitation phenomena namely: pauses, hesitations, fillers, and repetitions on the one hand. On the other hand, the components of fluency are compared as a result of the analysis of the hesitation phenomena aspects including planning and organization, proceduralization of information, and speed of delivery. The results in table 1 demonstrate the performances of the participants in the hesitation phenomena. The time interval of the pauses and between the pauses indicates how much non-native speakers engage in hesitation phenomena, as though planning what to say and how to say it is intriguing. The rates of pauses, hesitations, repetitions and hesitations (the results are indicated in table 1) of non-native speakers are higher and much more frequent than those produced by native speakers. In fact, native speakers use less fillers because they are considered as a distraction in speaking. For them, fillers like `um`, `err` and even words and expressions like `kind of `actually` and `sort of` distract the flow of speech and ideas, and even the listener may feel distracted the speaker overuses them. However it has been noticed that even native speakers use pauses and this is generally when they talk to non-native speakers. They also tend to speak slowly in simple and plain English, since they know they are interviewed by a non-native speaker of English.

Native speakers` speech is characterized by less repetitions and hesitations. Unlike non-native speakers who produce repetitions and hesitation as an action in planning ideas and preparing what to say next. Generally speaking, the production of pauses and fillers for non-native speakers is a result of repetitions and hesitations. The moment they start speaking and thinking about what to say and how to say it, they produce pauses and fillers to gain time to plan what to say next. Therefore, repetitions and hesitations accompany fillers and pauses to get more time to plan speaking. However, native speakers produce less repetitions and hesitations, not because they don't need pauses and hesitations, but the purpose behind producing pauses is different. Pauses are produced between sentences and when they shift from one idea to another. They are used to make a distinction between sentences and ideas as well and they help listeners consider the last idea very clearly. The production of repetitions and hesitations represents lack of knowledge and expressing the same idea in different ways. Meanwhile, too many pauses and fillers distract the listener's attention and create confusion in knowledge transforming.

6 Conclusion

Praat analysis of native and non-native helped in determining the general characteristics of speaking fluently. This is done by dividing the pauses into long pauses and short pauses, which generally extends from 100 up to 500. In this research non-native speakers produce longer and more frequent pauses when they speak as they need more time to think about what to say next and how to say it. In this case, non-native speakers shift focus in between language knowledge and topic knowledge. They need to construct schema knowledge about the topic they are discussing ideas, and their level of language proficiency hinders speed of delivery to some extent. Native speakers; produce longer stretches of discourse in Praat speech analysis, with less longer pauses. They already possess language knowledge and they deliver speech automatically.

Future research in this context can include other aspects like intonation, stress, pitch, rhythm and rhyme, which Praat can

analyze. This would help researchers in the domain of phonology obtain a clear image of the sounds produced to study their effects on pauses and hesitations.

Author short bio: Abderrahim Bouderbane has PhD in English from Constantine 1 University in Algiers. He published articles, book chapters and participated in many conferences. Currently he is a Lecturer at Mila University Algeria.

References

[1] Brown, A. (2003). Interviewer variation and the co-construction of speaking proficiency. Language Testing 20, 1–25.

[2] Bosker, H. R., Quené, H., Sanders, T., & de Jong, N. H. (2014). The perception of fluency in native and nonnative speech. Language Learning, 64(3), 579–614. https://doi.org/10.1111/lang.12067

[3] De Jong, N. H., and Perfetti, F. (2011. "Linguistic skills and speaking fluency in a second language." Applied Psycholinguistics. doi: http://dx.doi.org/10.1017/S0142716412000069.

[4] Fulcher, G. (2003).Testing second language speaking. Harlow, Pearson Education Limited.

[5] Freed, B. F. (1995). What makes us think that students who study abroad become fluent? In B. F. Freed (Ed.), Second language acquisition in a study abroad context, (pp. 123-148). Philadelphia: John Benjamins

[6] Guz, Zwa (2015). Establishing the fluency gap between native and non-native speech, Reasearch inlanguage , 13(3); pp.230-247DOI: 10.1515/rela-2015-0021

[7] Lennon, P. (1990) 'Investigating fluency in EFL: a quantitative approach,' Language

[8] Learning (40): 387–417.

[9] Levelt, W. J. M. (1989). Speaking: from intention to articulation. MIT Press, Cambridge,MA

[10] Lieshout, V. P (2003). Pratt Short Tutorial; A Basic Introduction. USA: University of Toronto.

[11] Nivja H. de Jong, Jos Pacilly & Willemijn Heeren (2021) PRAAT scripts to measure speed fluency and breakdown fluency in speech automatically, Assessment in Education: Principles, Policy & Practice, 28:4, 456-476, DOI: 10.1080/0969594X.2021.1951162

[12] Schnadt, M. J., & Corley, M. (2006).The influence of lexical, conceptual and planning based factors ondisfluency production.In Proceedings of the twenty-eighth meeting of the cognitive science society.Canada: Vancouver.

[13] Skehan, P. (2001). Tasks and language performance assessment.In M. Bygate, P.

[14] Skehan,& M. Swain (Eds.), Researching pedagogic tasks: Second language

[15] learning, teaching and testing (pp. 167-185). Essex, U.K.: Pearson Education

[16] Limited.

[17] . Peltonen, P (2017) Temporal fluency and problem-solving in interaction: An exploratory study of fluency resources in L2 dialogue System, 70 (2017), pp. 1-13, 10.1016/j.system.2017.08.009

[18] Saito, K., Ilkan, M., Magne, V., Tran, M. N., & Suzuki, S. (2018). Acoustic characteristics and learner profiles of low-, mid- and high-level second language fluency. *Applied Psycholinguistics*, *39*, 593–617.

[19] Styler, W. (2013). Using Praat for Linguistic Research. USA: University of Colorado.

[20] Rosenberg. (2009). Automatic Detection and Classification of Prosodic Events. Ph.D. Thesis.

[21] Segalowitz (2010). Cognitive bases of second language fluency. New York: Routledge.

[22] Suzuki, S. & Kormos, J. (2019). Linguistic Dimensions of Comprehensibility and Perceived Fluency: An Investigation of Complexity, Accuracy, and Fluency in Second Language Argumentative Speech. Studies in Second Language Acquisition, 1-25.

[23] Thomson, R, I (2015). Fluency. M. Reed, J.M. Lewis (Eds.), The handbook of English pronunciation, Wiley (2015), pp. 209-226

[24] Williams, K (2024) Exploring English for academic purposes instructors' perceptions of speech fluency through developing and piloting a rating scale for a paired conversational task, System, Volume 121, https://doi.org/10.1016/j.system.2024.103266.