Collecting, processing and analysing audio observations of teachers' engaging messages in secondary education

Samuel Falcón Pulido

Researcher at the

University of Las Palmas de Gran Canaria

Context

Some background on the framework

Data collection

Devices

Naturalistic procedure

Aspects to consider

Data processing

Transcription

Text filtering with lexical rules (traditional approach)

Text filtering with AI models (new approach)

Data analysis

Statistical analysis: Two-part models and how to handle zero-inflated data

Audio analysis: Paralinguistic features of the messages

Applicability and other possibilities with AI

Applicability of the procedure to other variables

Analysis of open-ended questions with AI

Context

Some background on the framework

Data collection

Devices

Naturalistic procedure

Aspects to consider

Data processing

Transcription

Text filtering with lexical rules (traditional approach)

Text filtering with AI models (new approach)

Data analysis

Statistical analysis: Two-part models and how to handle zero-inflated data

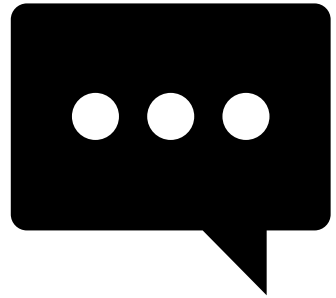
Audio analysis: Paralinguistic features of the messages

Applicability and other possibilities with AI

Applicability of the procedure to other variables

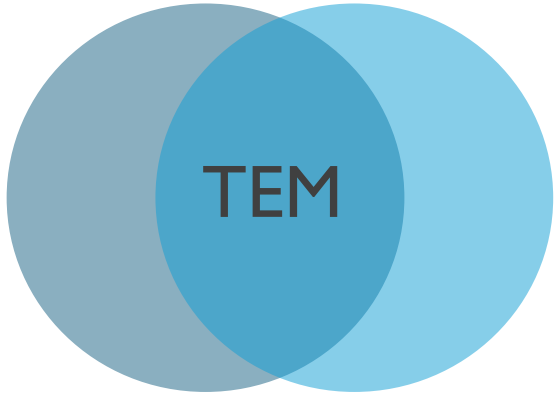
Analysis of open-ended questions with AI

What are teachers engaging messages?



Teachers' engaging messages (TEM) are those explicitly directed towards students with the purpose of engaging them in school tasks (study, do homework, etc.).

They have two dimensions based on how the message is framed and the motivation to which the teacher appeals.



Frame

Message Framing Theory (Rothman & Salovey, 1997)

- *Gain-framed* messages: benefits of engaging in a task
- *Loss-framed* messages: disadvantages of not engaging in a task

Appeal

Self-Determination Theory (Deci & Ryan, 2016)

- *Extrinsic* motivation: rewards, punishments
- *Introjected* motivation: own or other feelings
- *Identified* motivation: future value of studies
- *Intrinsic* motivation: pleasure of engaging

Eight categories in total

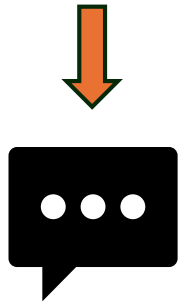
Why focus on them?

Mainly because previous studies have shown some interesting results regarding both teachers and students.

For instance, depending on the autonomy they perceive in their work, we can identify different profiles of teachers who use different types of engaging messages.

And also, the type of message delivered to students has different influences on students' performance, motivation and well-being.

Why use observations?



These previous studies only measured engaging messages through student self-reports, when it is not the best way to study teachers' verbal behaviour.

To close this gap, we set out to obtain direct observations of the messages through classroom audio-recordings.

In addition, we implemented different tools to optimise audio processing and thus save coding time.

Any questions?

What are teachers’ engaging messages

Why did we start using observations

Context

Some background on the framework

Data collection

Devices

Naturalistic procedure

Aspects to consider

Data processing

Transcription

Text filtering with lexical rules (traditional approach)

Text filtering with AI models (new approach)

Data analysis

Statistical analysis: Two-part models and how to handle zero-inflated data

Audio analysis: Paralinguistic features of the messages

Applicability and other possibilities with AI

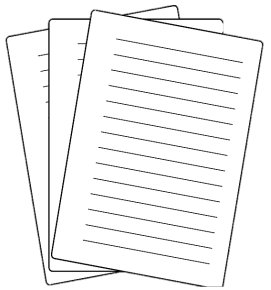
Applicability of the procedure to other variables

Analysis of open-ended questions with AI

Most of the information I will explain from now can be found in more detail in the following papers:

Falcon, S., Admiraal, W., & Leon, J. (2023). Teachers' engaging messages and the relationship with students' performance and teachers' enthusiasm. *Learning and Instruction*, *86*, 101750. https://doi.org/10.1016/j.learninstruc.2023.101750

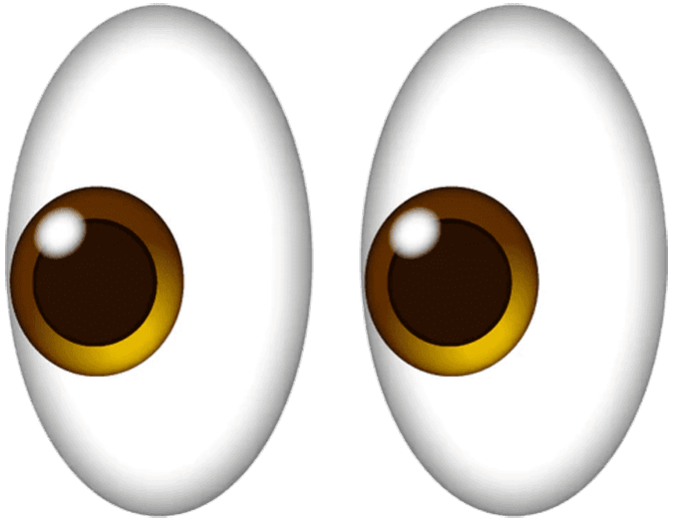
Falcon, S., Alonso, J. B., & Leon, J. (2023). Teachers' engaging messages, students' motivation to learn and academic performance: The moderating role of emotional intensity in speech. *Teaching and Teacher Education*, *136*, 104375. https://doi.org/https://doi.org/10.1016/j.tate.2023.104375



Naturalistic setting

In a controlled environment, behaviours can be influenced or altered by the awareness of being observed.*[1]

For that reason, our research has always focused on studying teachers in their usual working environment, such as the classroom, so that we can get a truer picture of how they communicate with students.

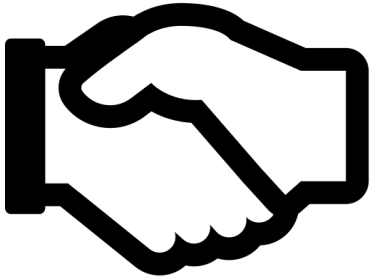


How is the procedure?

We contact the schools to explain to the management teams and teachers what the research is about and what they have to do to participate.

Specifically, we ask them to audio-record at least eight lessons per term in groups from Grade 8 to High School.

With each of the teachers interested in participating, we signed a consent form where they allow us to record their voices and we commit ourselves not to disseminate it, thus complying with the Data Protection Act. *[2]



Devices

To be able to carry out the observations without disturbing the normal functioning of the classrooms and to collect only the voices of the teachers, we ask them to use one of the following two devices:

A small recorder

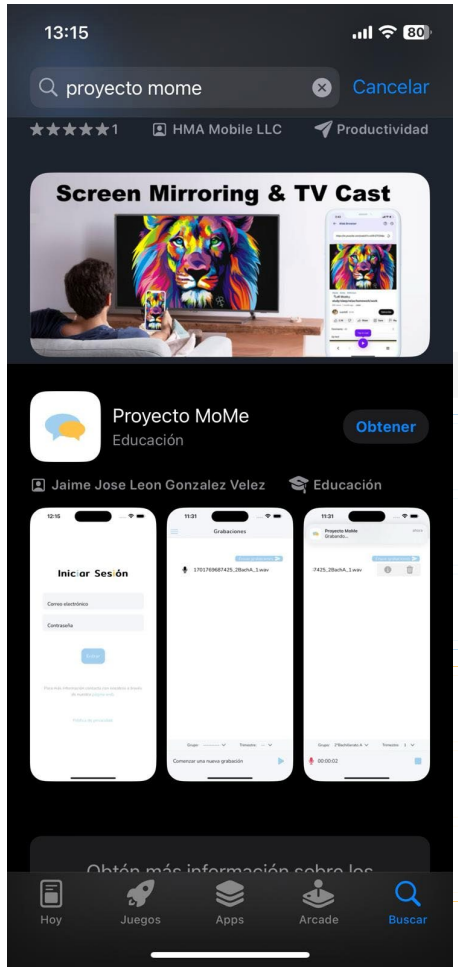
Their cell phones

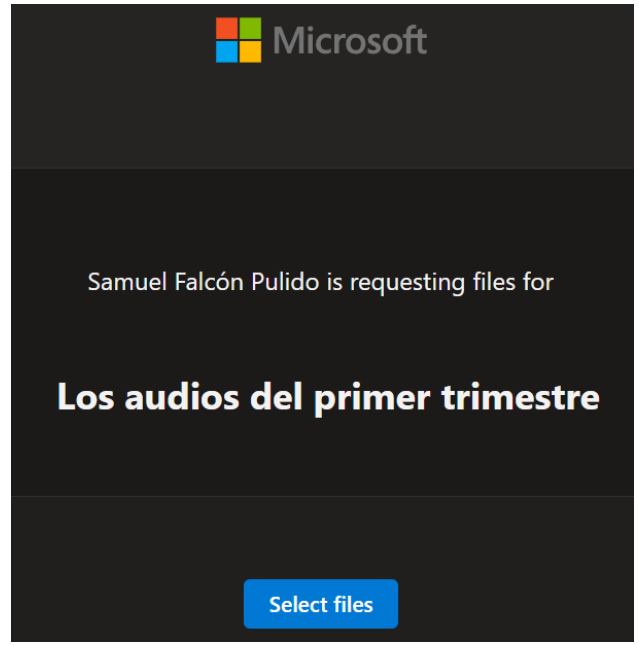
Audio sending

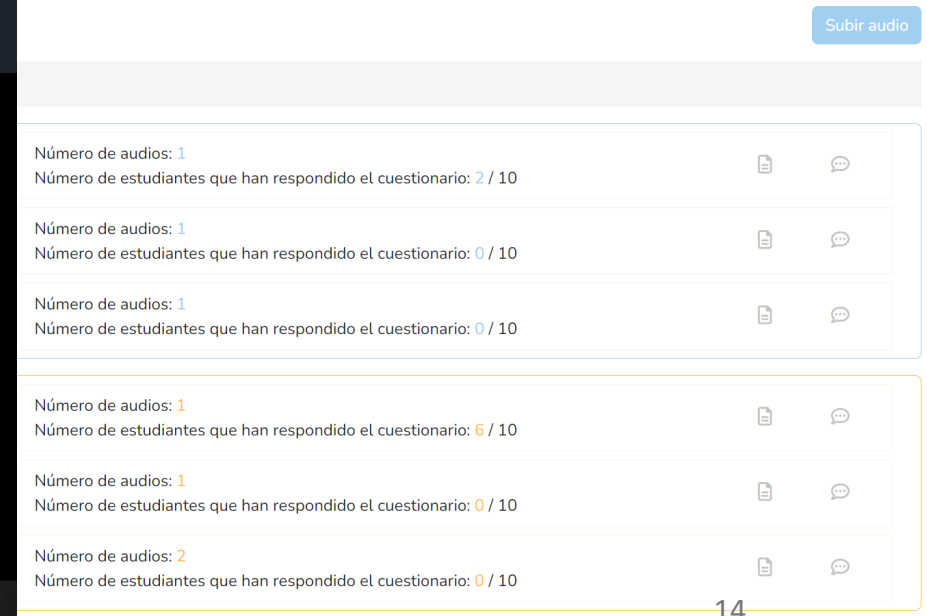
Once they have finished recording themselves, we ask them to send us the audios.

When we started, we asked them to send the audios via a OneDrive link.

Today we have developed a cross-platform application (web and cell phones) so that they can send us the data more easily and consult other data on their participation.

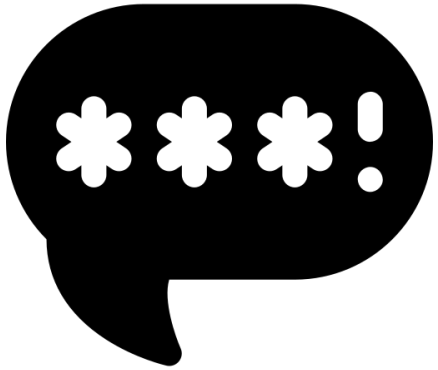






Some considerations

*1 Do teachers change their behaviour knowing they are being recorded? Maybe in the first 5 minutes of the first lesson, but no. Believe me, we've heard all sorts of things....



*2 Do you only record the teachers' voices or also the students' voices? On rare occasions, such as when the teacher is giving feedback to the students, you can hear their voices, but most of the time only the teacher's voice is being recorded because the recorder is close to their mouth.

More considerations

And now, some things that may be obvious but that you should always keep in mind in these kind of data collection:

The simpler the procedure for teachers, the better for us.

Audio quality is key.

Any questions?

Procedure

Devices

Audio sending

Other considerations

Context

Some background on the framework

Data collection

Devices

Naturalistic procedure

Aspects to consider

Data processing

Transcription

Text filtering with lexical rules (traditional approach)

Text filtering with AI models (new approach)

Data analysis

Statistical analysis: Two-part models and how to handle zero-inflated data

Audio analysis: Paralinguistic features of the messages

Applicability and other possibilities with AI

Applicability of the procedure to other variables

Analysis of open-ended questions with AI

How to make the audios more manageable?

Different approximations to obtain observations from audio-recordings:

Some researchers listen to the audios to extract information

This can be very costly and time-consuming

Many researchers transcribe the audio into text for analysis

This may cause a loss of information*[3], but the identification process is less costly

How to transcribe the audios?

You can use software as a service (SaaS), which allows users to connect to and use cloud-based apps:

Microsoft's Azure service	Pros:	Cons:
Google's service	Easy to use	Sending data to third parties
Amazon's service	Good quality	It costs money (not that much)

You can use open-source software, where you download the models and run them locally

Whisper from OpenAI	Pros:	Cons:
	Good quality	It can be difficult to install (I recommend a technician on the team)
	Data stays in your PC	Requires powerful PC
	Free	

How to transcribe the audios?

Audio

Whisper transcript

```
00:00:00        Pour ce matin, mathématiques, nous avons un problème.
00:00:12        Un petit problème.
00:00:14        Donc pour l'instant, je vous le distribue.
```

We have the transcripts, now what?

After transcribing tens or hundred of hours of audio, we can end up with thousands of pages of text

Reading throughout all the transcript searching for engaging messages can be really slow and may lead to decision fatigue

For this reason, over the years we have developed two methods to extract information

How to extract information from the transcripts?

Method A – Filtering based in lexical rules

Easy to follow
Free

Not the fastest

Method B – LLM (AI)

Faster
State-of-the-art

Experimental
It requires an IT expert

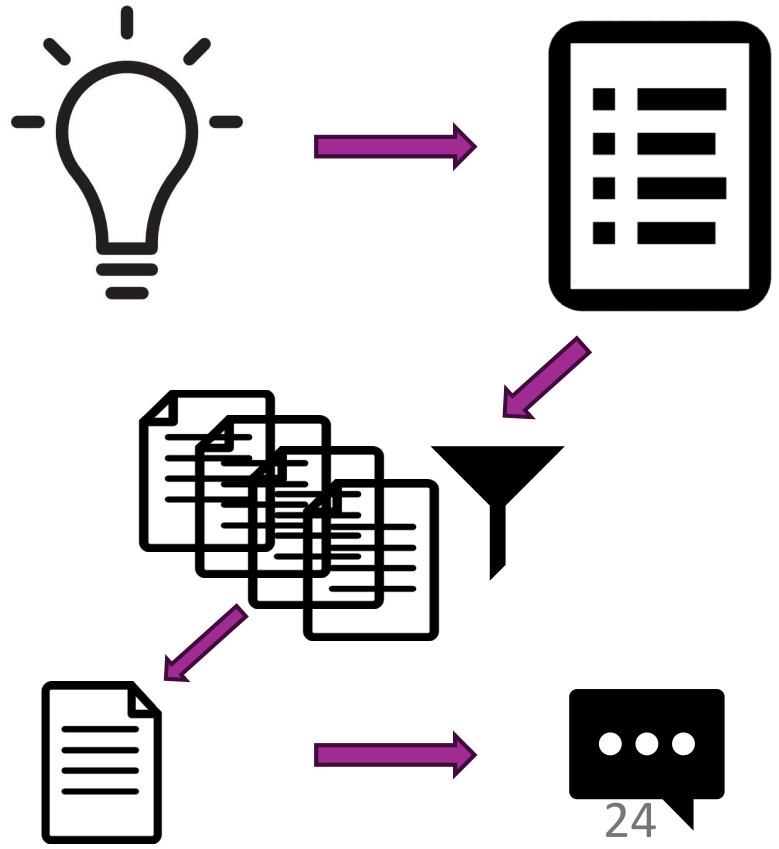
Method A – Without real examples

You want to gather observations on engaging messages, but you don´t have real examples

1º Use the theory or the scale with which these messages are measured to create a list of words that you think will be in or around engaging messages.

2º Filter the original transcripts to obtain only sentences containing those words

3º Code the filtered transcripts that contain 80% less information than the original ones and save time and resources (based on our own results)



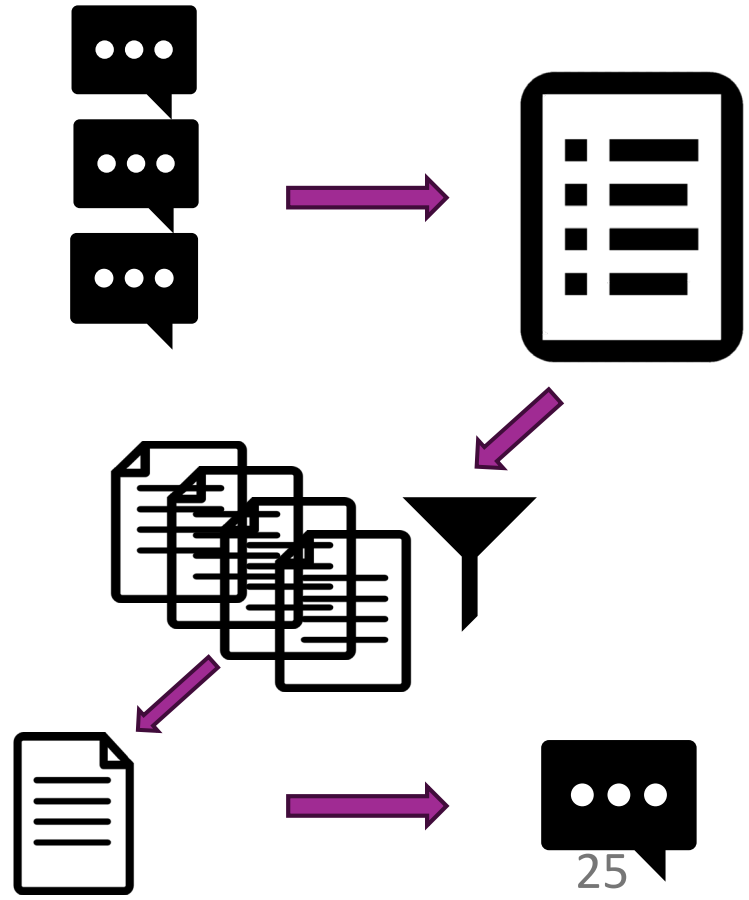
Method A – With real examples

You want to gather observations on engaging messages, and you have real examples

1º Use the real examples to create a list of words that very common in the messages, but less common in the rest of the text.

2º Filter the original transcripts to obtain only sentences containing those words

3º Code the filtered transcripts that contain **90%** less information than the original ones and save time and resources (based on our own results)



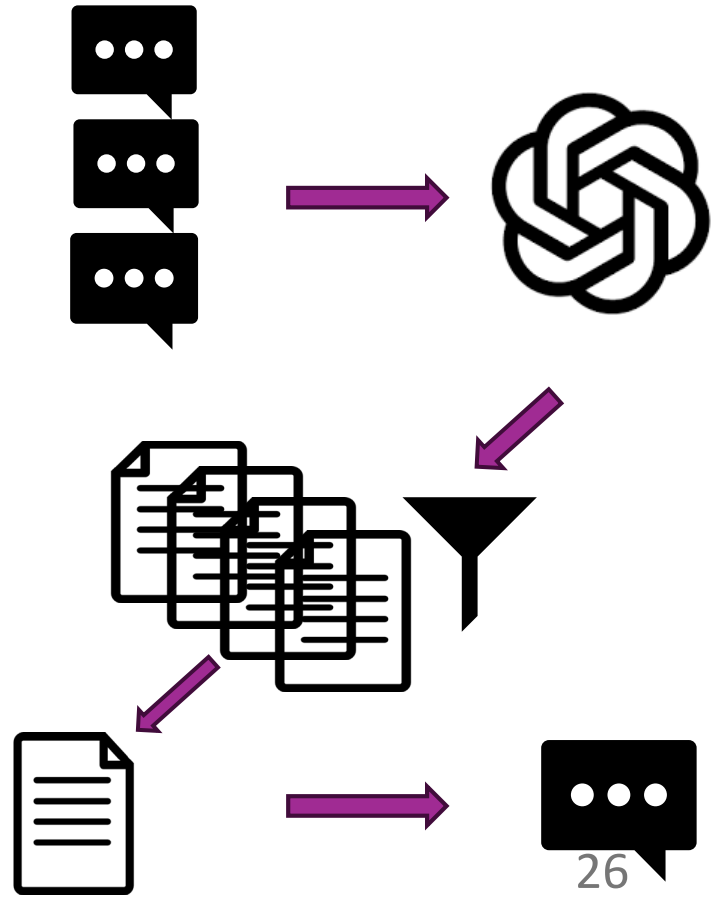
Method B – Only with real examples

You want to gather observations on engaging messages, and you have real examples

1º Use the real examples to train an LMM on real positive and negative examples of what you want to study

2º Analyse the original transcripts with the trained model to obtain only sentences similar to the real examples

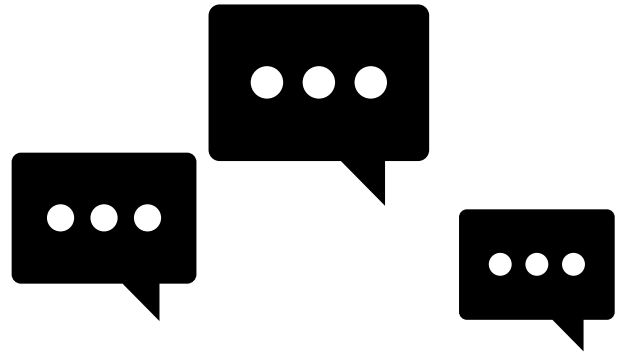
3º Code the filtered transcripts that contain **95%** less information than the original ones and save time and resources (based on our own results)



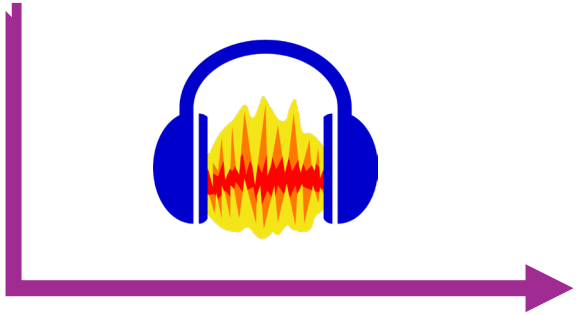
We have the messages, now what?

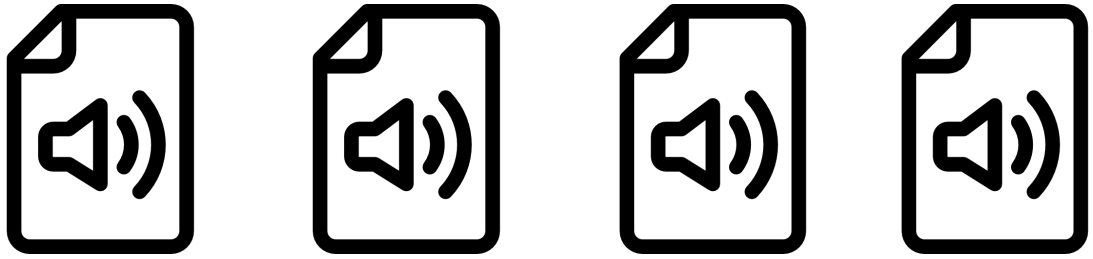
Audio-recorded lessons

Method A
Method B



Observations of teachers' engaging messages





*[3] audio clip of each engaging message

00:00:00	Pour ce matin, mathématiques, nous avons un problème.
00:00:12	Un petit problème.
00:00:14	Donc pour l'instant, je vous le distribue.

Limitations of these methods

These methods don’t come without limitations

Both of them are faster than listening to the whole audio recordings or reading all the transcripts. However, we may miss some engaging messages

We believe that these methods strike a balance between speed and reliability

Any questions?

Transcription

Method A without examples

Method A with examples

Method B with examples

Audio clip creation

Context

Some background on the framework

Data collection

Devices

Naturalistic procedure

Aspects to consider

Data processing

Transcription

Text filtering with lexical rules (traditional approach)

Text filtering with AI models (new approach)

Data analysis

Statistical analysis: Two-part models and how to handle zero-inflated data

Audio analysis: Paralinguistic features of the messages

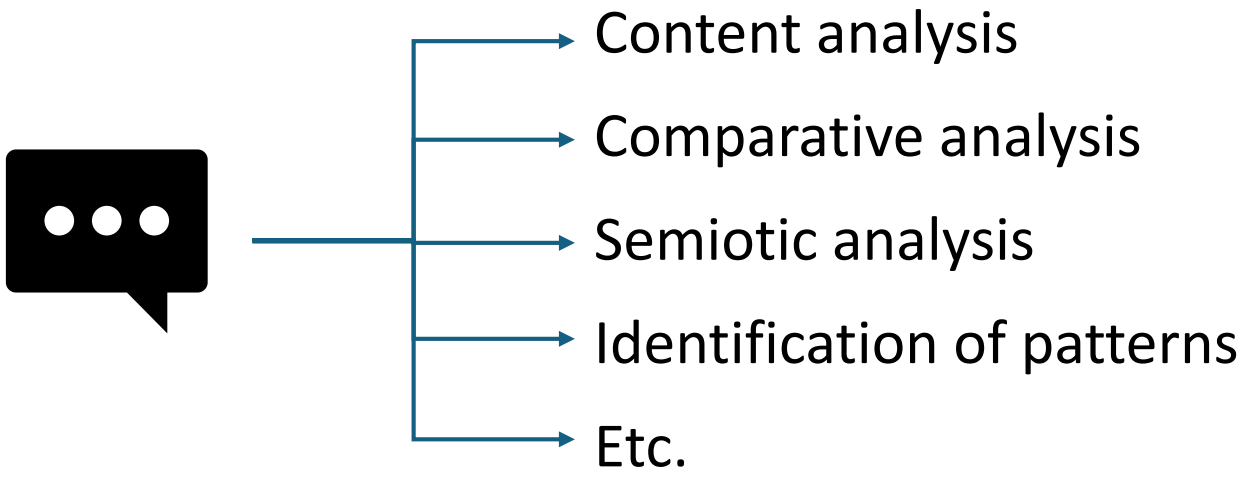
Applicability and other possibilities with AI

Applicability of the procedure to other variables

Analysis of open-ended questions with AI

Qualitative analysis

The first way we will have to analyse the data obtained is by carrying out a qualitative analysis of the information. Of course, this will always depend on the objective of our research.

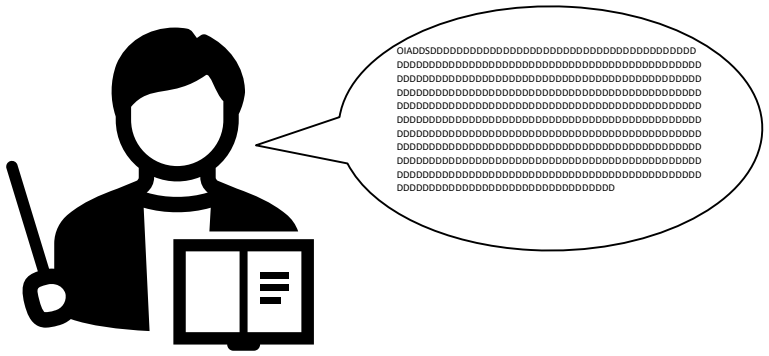


Development of new theories or modification of existing ones based on empirical evidence

Quantitative analysis time! But before…

These methodologies allow for the analysis of large samples and large amounts of data, so another possibility is to carry out quantitative descriptive analyses or statistical modelling to make inferences.

Teacher A



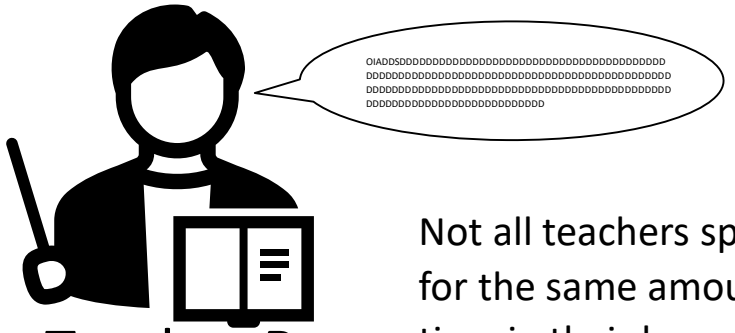
5 messages

30 000 words

$$\frac{5}{20\,000} = 0.00017$$

1.7 messages per 10 000 words

$$\frac{Number\ of\ messages}{Number\ of\ words}\ x\ 10\,000$$



Teacher B

5 messages

20 000 words

$$\frac{5}{10\,000} = 0.00025$$

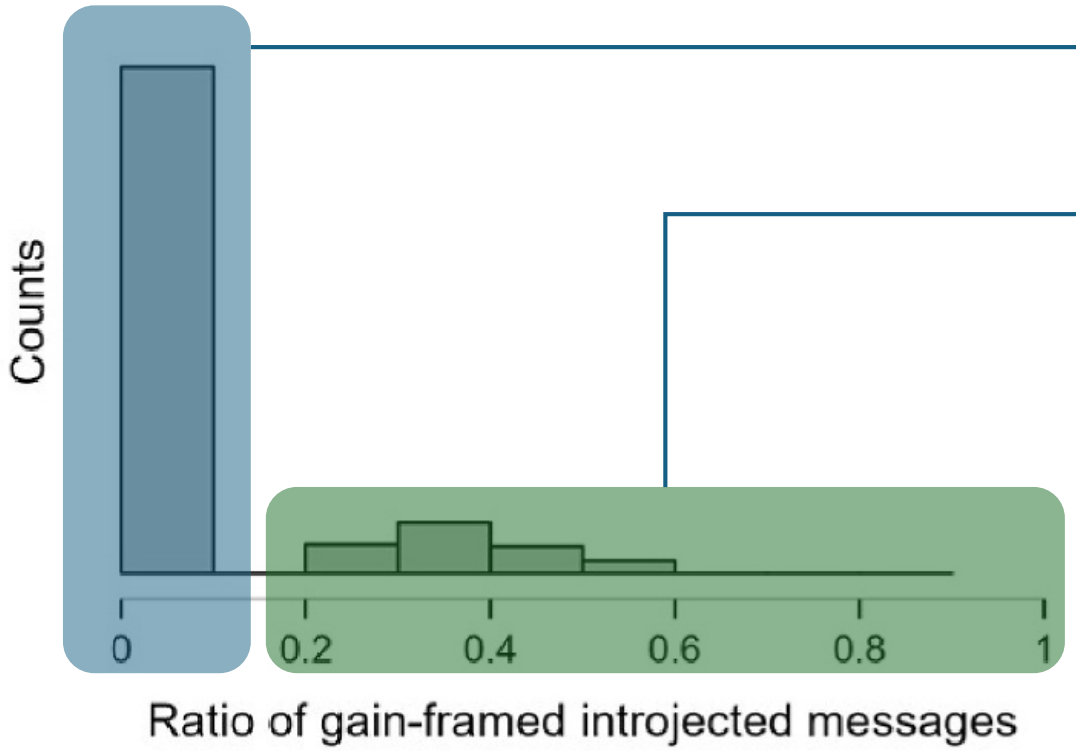
2.5 messages per 10 000 words

Not all teachers speak for the same amount of time in their lessons.

Now we can compare the teachers

Quantitative analysis

Now that we have standardised the data, we can carry out different types of analysis. However, if we plot the data and look at its distribution, we will see something peculiar.



Zero-inflated distribution

Many teachers don't use this type of messages with their students

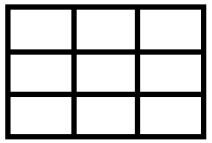
Some teachers use them to a different extend

Similar situations:

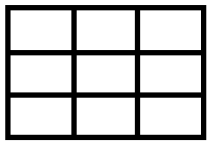
- Many people don't think about suicide in the last twelve months, but people who do think about it will do so a different number of months per year.
- It rarely rains in the desert, but when it does it can rain in different amounts.

How to analyse zero-inflated data

One way (but not the only one) to analyse these data is through hurdle models (for count data) or two-part models (for continuous data)



Original dataset with zero-inflated data

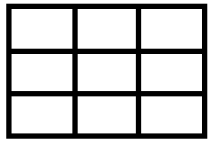


Dichotomous dataset: absence or presence of the phenomenon

Logistic regression

How the teacher's enthusiasm influences his **likelihood** of using engaging messages

If we were interested in analysing the influence of enthusiasm on the use of messages



Dataset with only observations of the phenomenon

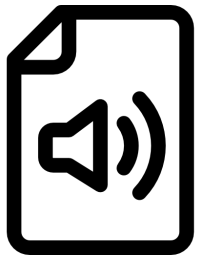
Linear regression

Among teachers who do use messages, how enthusiasm influences the number of messages used

Analysing audio data

Remember the audio clips that we created after identifying the engaging messages?

There are many tools and models, both commercial and open-source, that allow the extraction of different types of information from audio clips. For example, they can extract paralinguistic information, such as pitch, rhythm, pauses and voice inflections, which are important for analysing the emotions and intentions behind spoken words.



- Emotional Temperature model by Alonso et al. (2015) → Open-source model that allows to extract a numerical value of the emotional intensity present in the message, a value related to arousal.
- HumeAI service based on the work of Schuller et al. (2023) and Brooks et al. (2023) → Commercial model that allows to analyse an audio and give a numerical value to 12 universally present emotions.

This field is really broad and there are many, many more tools than these.

Any questions?

Qualitative analysis of the messages

Quantitative analysis: Standardization

Quantitative analysis: Models

Analysis of audio clips

Context

Some background on the framework

Data collection

Devices

Naturalistic procedure

Aspects to consider

Data processing

Transcription

Text filtering with lexical rules (traditional approach)

Text filtering with AI models (new approach)

Data analysis

Statistical analysis: Two-part models and how to handle zero-inflated data

Audio analysis: Paralinguistic features of the messages

Applicability and other possibilities with AI

Applicability of the procedure to other variables

Analysis of open-ended questions with AI

Applicability of the methods

There are many verbal behaviours that can be analysed through audio-recorded lessons using this methodology:

Teacher motivational messages: Messages that teachers use before exams to push students to study (Putwain et al., 2021)

Praise: Messages that teachers use to commend the worth of or to express approval or admiration (Jenkins et al., 2015)

Many other verbal behaviour: Empathic messages, feedback, autonomy supportive messages, etc.

Caution: Make sure that the phenomenon that you want to study can be operationalised

Bonus track: Automatic coding using LLMs

Large Language Models like ChatGPT can be used for many purposes

One of them is to perform automatic classification of text, which can be really useful for easily coding huge amounts of answers to open-ended questions

This can save time and resources, and allow you to analyse huge amounts of data for performing advanced statistical analyses and make inferences

Procedure for automatic classification using LLMs

1º Identify a theoretical framework : It should include clear and concise definitions of the categories into which you want to classify the responses.

2. Write the open-ended question and test them if you can: Avoid questions that can be answered with a Yes/No, whose answers are very unstructured or long, etc.

3º Collect the data

4º Create a prompt for GPT to classify the open-ended responses: It usually has an instruction about the task, the definitions of the categories, instruction about the response format and some examples. There is a limit to the length of prompts, so it is advisable not to have too many categories or extremely long definitions.

5º Analyse a random percentage of the sample independently to calculate the agreement between humans and GPT. In this way we ensure that GPT is correctly classifying the answers to the open-ended question.

6º If the results are satisfactory, do the rest of the classification automatically only with GPT and then perform the analysis we want with the data obtained.

Example of classification using LLMs

Álvarez-Álvarez, C., & Falcon, S. (2023). Students' preferences with university teaching practices: analysis of testimonials with artificial intelligence. Educational Technology Research and Development, 71, 1709–1724. https://doi.org/10.1007/s11423-023-10239-8

We were interested in examine students' preferences on university teacher pedagogical practices

We found a framework containing 9 definitions of different pedagogical practices and we add another category for "no preferred practices"

1089 students answered the open-ended question

The agreement humans-GPT on a random 10% of the sample was very good

We classified the rest of the sample using only GPT and saving a lot of time

Now we are using the same methodology in the LaRAC to examine students' answers to open-ended questions on sources of motivation and stress → It is a very flexible methodology that allows many types of variables to be studied in different contexts.

Thank you for your attention

Any questions?

samuel.falcon@ulpgc.es