



HAL
open science

Semi-Structured Interview Analysis: A French NLP Approach for Social Sciences

Robin Quillivic, Charles Payet

► **To cite this version:**

Robin Quillivic, Charles Payet. Semi-Structured Interview Analysis: A French NLP Approach for Social Sciences. JADT - 2024, LasLa, Jun 2024, BRUXELLES, Belgium. hal-04627806

HAL Id: hal-04627806

<https://hal.science/hal-04627806>

Submitted on 27 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Semi-Structured Interview Analysis: A French NLP Approach for Social Sciences

Robin Quillivic¹, Charles Payet²

¹EPHE, CESSP (UMR 8209), ISC-PIF (UAR 3611) – robin.quillivic@ephe.sorbonne.fr

²ENS Paris-Saclay, ISP (UMR 7220) – charles.payet@ens-paris-saclay.fr

Abstract

Semi-structured interviews (SSIs) are widely employed in social sciences, facilitating in-depth data collection while maintaining study focus. Typically involving asymmetric roles (an interviewer and interviewee), SSIs follow predefined guides. However, current analyses predominantly scrutinize participant responses, overlooking the interviewer's language, context, and interactions. Even within guided interactions, discourse is a co-construction, yet the investigator is often rendered invisible. Drawing on conversation analysis principles, we present a French toolbox for analyzing SSIs in social sciences. Integrating multiple Natural Language Processing (NLP) approaches—semantic, thematic, and grammatical similarity—the toolbox enhances research methodology by offering precise indicators of interaction influence in interviews. The toolkit comprises three main components. The first provides a comprehensive dataset overview, including interview structure (number and length of speaking turns), and user-provided metadata. The second module analyzes the interview theme and categorizes the dialogic act using few-shot learners. The third module focuses on discussion alignment, building features using cosine similarity, and sentence embeddings. To demonstrate the toolkit's utility, we present two examples using interview transcripts produced as part of a research program on the memory of the November 2015 Paris attacks. The first explores interaction variability based on the socio-demographic profiles of the interviewer and respondent, while the second focuses on psycholinguistics, measuring how interview structure and theme are influenced by avoidance symptoms and PTSD.

Keywords: semi-structured interviews; interviewer-respondent interaction; conversation analysis; NLP; toolkit; computational social sciences

1. Introduction

The interview is one of the flagship research techniques in the social sciences. The semi-structured interview, the most common type, is an investigation method in which an interviewer asks questions and engages with the interviewee, often following a predetermined interview guide while allowing the interviewee to express themselves freely. The goal is to collect rich data from the interviewee's speech about a subject, whether it be experiences or past behaviors, opinions, or attitudes. Its advantage is that it captures the interviewee's perspective in their own words while guiding them toward pre-defined topics (see Brinkmann et Kvale, 2018, for a general introduction to interview-based research). Researchers apply various analytical approaches, including content analysis, narrative analysis, hermeneutics, and discourse analysis.

Most traditional approaches treat interview content as factual "reports" from participants, revealing their memories and thoughts (Talmy, 2010; Koven, 2014). Additionally, the data produced through interviews are perceived as reflecting the interviewees' reality beyond the interview setting (Rapley, 2001). This perspective, considering interview data as a window into the thoughts and life of the interviewee, tends to direct the researcher's attention toward the content of the interviewee's talk. This ultimately aligns with the notion of the interview as a "research instrument" (Talmy, 2010). However, treating the interview as a research instrument

often leads researchers to treat its content as separable from the interview context, thereby overshadowing its interactional aspects, particularly the role of the interviewer. In contrast, our approach aligns with a divergent (but not less complementary) conception of the interview as a "social practice" (Talmy, 2010).

Rapley (2001) identifies three points of scrutiny to regard the interview as a social practice: considering the interviewee's talk as potential "accounts" or "versions"; these must be interpreted in direct relation to the interview context; and finally, paying special attention to interviewers' talk and their interaction with the interviewee. Here, we emphasize this third aspect as our main concern in interview transcription analysis.

This approach has allowed for the demonstration of how certain conversational resources and rhetorical devices operate within the interview conversation but has also been beneficial to the specific subjects/objects of the researchers (Roulston, 2019). This has mostly been done through conversation analysis and qualitative methods.

Computational or quantitative text analysis studies that focus on interview transcriptions generally consider the material as originating from a research instrument and very rarely as a social practice. Consequently, they often treat it as decontextualized content with the interviewer's intervention being deemed non-existent. By adopting an interactional perspective of the interview and utilizing computational methods, we have developed a methodology and proposed an implementation in Python for analyzing interview transcriptions.

Following a brief review of related work, we demonstrate the utility of the methodological approach and toolkit on a case study involving a corpus of 926 semi-structured interviews about the November 13, 2015, attacks in Paris. This analysis considers not only the characteristics of the interviewees but also those of the interviewers and the dynamics between the two. Using conversation analysis and natural language processing (NLP), we measure interview structure in terms of utterances, pace, and other factors, characterizing the questions posed by interviewers. Our goal is to develop a methodology for constructing features for interview configuration, leaving the choice of statistical analysis to researchers. As an illustration, we conduct a statistical analysis to unveil (1) gender configuration differences in interviews (interviewee - interviewer) and (2) explore the association between interview structure and PTSD (Post-Traumatic Stress Disorder). Finally, we discuss potential applications and areas for improvement.

2. Related Work

Textometry quickly embraced semi-directive interview transcriptions, employing its classic tools, notably Reinert's method (Scelles, 1997). While these studies focused solely on the content of interviewees, some research explored the interventions of interviewers. For instance, Bonneau et Dister (2010) demonstrated the influence of interviewer-interviewee configurations on the use of pronouns and raised the question of the possibility of interview conduct categories. Dalud-Vincent (2011) also applied Reinert's method to corpora with and without interviewer interventions. The results showed no changes in the different categories. However, a notable gap in existing research lies in the treatment of interviewee speech and interviewer interventions not merely as distinct entities based on metadata but also in neglecting the essence of interaction—specifically, the sequences of utterances unfolding in a meaningful order. A central concept in many of these studies is linguistic coordination or linguistic alignment (we use both terms interchangeably). This concept refers to the tendency in a conversation for participants to "coordinate" or "align" what and how they communicate with others. This

linguistic alignment is then measured between two turns of speech and repeatedly throughout a conversation, allowing for the assessment of conscious or unconscious coordination among participants. It was initially operationalized by comparing "function words"¹ between two turns of speech. Danescu-Niculescu-Mizil and Lee (2011) used it to investigate whether coordination occurs in scripted movie dialogues and observed a gender difference: characters in movies, regardless of gender, align more with female characters than with male characters. Danescu-Niculescu-Mizil and al. (2012) demonstrated that the dynamics of linguistic coordination among conversation participants could reveal power relationships among them (in the context of discussions among Wikipedians and arguments before the U.S. Supreme Court). This coordination, based solely on linguistic style (utilization of function words), has been questioned (Doyle and Frank, 2016). More recently, Nasir et al. (2019) suggested using distributional representations of words (word embeddings) to primarily measure semantic and syntactic alignment. They, apply it to transcriptions and demonstrate correlations with empathy in therapeutic settings and affective behaviors in Couples Therapy.

In studies where dialogic acts and conversation were the focus of interest, researchers have sought to characterize speech turns. Niven and Kao (2019) studied the connection between dialogic argumentation and linguistic alignment. They demonstrated that measuring alignment allows the distinction of argumentative discourse acts, such as differentiating between questioning assertive or questioning pure. Chang et al. (2020) introduced an Expected Conversational Context Framework as a general method for analyzing utterances in their context. This framework connects utterances to the surrounding turns, including replies and predecessors, enabling the characterization of utterances. The latter can be interpreted as roles of utterances in the interaction. For example, Zhang et al. (2017) applied this method to question-and-answer sessions in the British Parliament, providing insight into the interactions among parliamentarians based on their party, positioning, and career strategies. Using transcripts of crisis counseling, Zhang and Danescu-Niculescu-Mizil (2020) developed a measure to quantify whether an intervention was more backward-oriented (in their case, empathetically addressing the crisis situation) or forward-oriented (advancing the conversation towards a resolution).

While the aforementioned examples focus on situations involving asymmetric interactions (which can be beneficial for our work), to our knowledge, no study has engaged in the analysis of semi-structured interview transcriptions using an interactional perspective.

3. Data

The "Etude 1000" protocol collected the data used in this study (Eustache et Peschanski, 2022) which consists of collecting the narratives of 934 volunteers who have been affected by the attacks in some way. The particular and innovative methodology is based on three questionnaires submitted to the volunteers, most of which were videotaped. Moreover, this longitudinal study will take place four times, in 2016, 2018, 2021, and 2026. The present analysis focuses on the data collected in 2016. Participants were recruited voluntarily, through information given by newspapers, the city hall of Paris, and victims' associations. Active recruitment involved contacting potentially eligible individuals by the program team. Participants were included from April 1st, 2016 to November 10th, 2016, 5 to 11 months after the traumatic event. The film interview, divided into two parts, collects data about emotional,

¹Articles, auxiliary verbs, conjunctions, high-frequency adverbs, impersonal pronouns, negations, personal pronouns, prepositions, and quantifiers, typically from LIWC dictionaries.

physical, and social reactions. In the first part of the interview, the participants were asked to produce a narrative relating to their experience of the attack. A second question asked about their interpretation of the event and analysis of causes and consequences (see Table 1 for the thematic distribution). It was followed by a shorter structured interview, an emotional memory questionnaire, that included 16 closed and some open-ended questions about post-traumatic symptoms and behaviors according to DSM-5. The interviews were transcribed using the Vocapia (Speech to Text API | Vocapia) software and then corrected by a human annotator. The metadata on the participants (gender, occupation, etc.) and on the interviewer (gender, occupation at the time of the interview, field of study, etc.) were also collected using written questionnaires.

Thematic of the Question in the semi-structured interview	Sub-questions thematics
Personal Experience	Receiving News of Attacks. (Nouvelle)
	Reaction to the event (Reaction)
	Perception of Time (Time)
	Relation to Media (Media)
	Capacity to reconstitute the scene (Reconstitution)
	Relation to Others (Connaissance)
	Relation to location (Places)
Causes and consequences	Causes analysis
	Consequences interpretation

Table 1: Themes from Etude 1000 interview's guide

For our use case, we kept 926 interviewees, because 2 were non-native French speakers and seven were conducted with two interviewers at the same time.

4. NLP methodology and toolkit

Our approach comprises two key components: constructing features based on cosine similarity between sentence embeddings and enriching interview data using "few-shot learning" classifiers. In the first instance, we aim to quantify the alignment of discourse between the interviewer and the interviewee. This quantification is performed at two levels: the turn of speech and the sentence (the last of the previous and the first of the next turn of speech). For each interview, we compute the averages and standard deviations of lexical alignments per turn of speech and sentences. We used the sentence embeddings "paraphrase-multilingual-MiniLM-L12-v2" proposed by Reimers and Gurevych (2019). Moving to the second approach, we enhance utterances with their thematic and dialogical categories. This enrichment is achieved through a "few-shot learning" classification using the SetFit Python package (Tunstall et al., 2022). Thematic themes are derived from the interview guide (Table 1) provided by the user, while dialogical categories are drawn from Brinkmann and Kvale, (2018) see Table 2. In our study, approximately 500 turns of speech out of 50,000 were labeled to fine-tune our models on our corpus and ensure robust classification performance. The source code of this project is available². The trained models are available on the Hugging Face (HF) hubs: Thematic model³ and dialog act categories⁴. The training information is available in Table 2.

² <https://github.com/psycholinguistics2125/ssi-toolbox-fr>

³ <https://huggingface.co/binbin83/setfit-MiniLM-dialog-themes-13-nov>

⁴ <https://huggingface.co/binbin83/setfit-MiniLM-dialog-act-13nov>

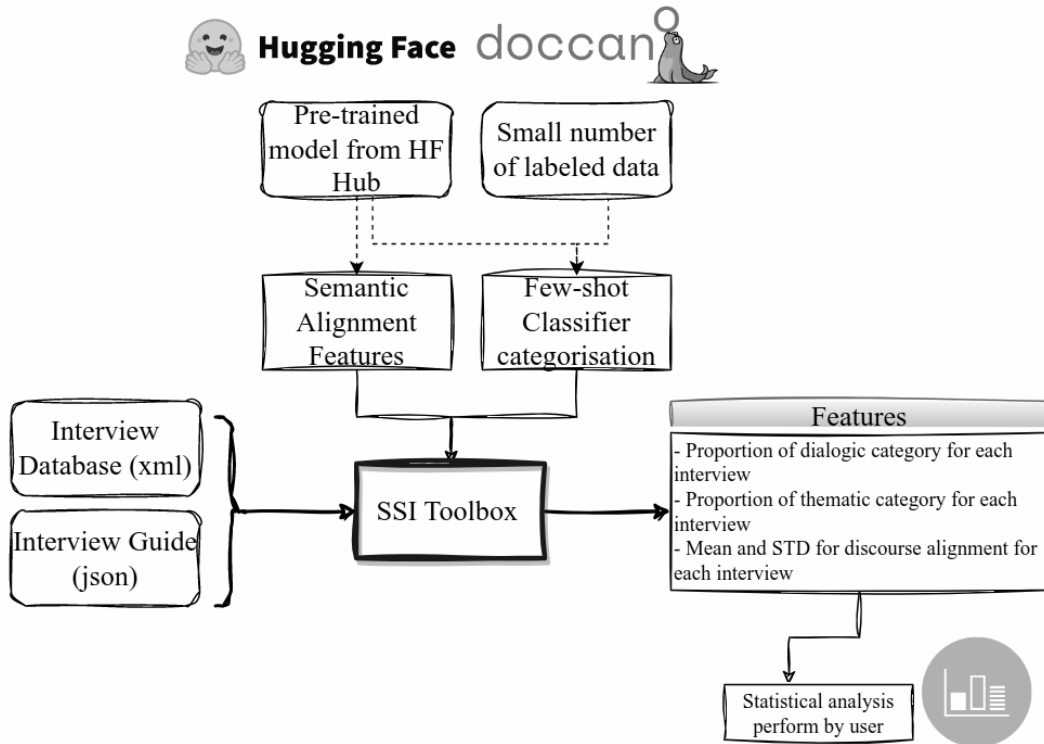


Figure 1: Illustration of the pipeline proposed to analyze Semi-Structured Interview

5. Results

In our analysis, we focus on the interview as the primary unit, emphasizing interaction dynamics. Our methodology constructs features capturing nuances in both interviewer-interviewee interaction and interview structure from the interviewer's perspective. In this section, we propose to use these features in two research scenarios, one on gender configuration and the other on the relationship between PTSD diagnosis and interview structure. We conduct a short inductive and exploratory statistical analysis using simple methods such as Principal Components Analysis (PCA) and t-tests, to illustrate the possibilities of our methodology.

Dialog Act Name	Example	Number in labeled data
Introductory	Can you tell me about ... (something specific)?	12
Follow-up	mmh, you can continue, ok.	134
Probing	Can you say a little more about X?	146
Specifying	Can you give me an example of X?,	135
Structuring	Thank you for that. I'd like to move to another topic..	27
Direct	When you mention X, are you thinking like Y or Z?	125
Interpreting	So, what I have gathered is that...	44

Ending	I have asked all the questions I had, but I wanted to check whether there is something else about your experience/understanding we haven't covered?	12
--------	---	----

Table 2: Dialog Act category from (Brinkmann and Kvale 2018) and number of examples in training dataset for few-shot learning

5.1. Does gender configuration influence the interview structure?

A principal component analysis is performed on three sets of variables. The first group comprises six variables concerning the structure of exchanges: average words spoken per utterance by the interviewer, the ratio of voluntary words to the interviewer, the ratio of short utterances (< 3 words) by both interviewer and interviewee, and the average linguistic coordination with the preceding and following sentences. The second group consists of dialog acts (see Table 2), and the third group comprises dialog themes (see Table 1). The variable of interest is the interview gender configuration (WxW, WxM, MxW, MxM), which is projected into the new space as an additional variable. The analysis is restricted to interviews conducted by the team in Paris, totaling 630 interviews. We show the factorial plane comprising axes 1 and 2 (Figure 2). The variables that most strongly contribute to axis 1 are (positively correlated) follow-up utterances, the ratio of short utterances by the interviewer, average linguistic coordination with the previous sentence, and (negatively correlated) direct questioning, along with media-related, place-related, and cause and consequence dialog themes. The variables that contribute most to axis 2 (positively correlated) are structuring, average words spoken per utterance by the interviewer, and (negatively correlated) interpreting, specifying, and questions regarding personal experience.

Figure 2 displays the barycentric positions of the four gender configuration categories. Hence, the interviewer's gender seems to shape the variations along axis 1. Female interviewers tend to retrieve information by asking more direct questions on predefined guide themes, while male interviewers appear more passive, with shorter follow-up interventions that linguistically align with the interviewee's preceding sentence (especially when both are men). Further investigation is needed to interpret these results, for instance, exploring whether the interviewer's age or status (student vs experienced researcher) is associated.

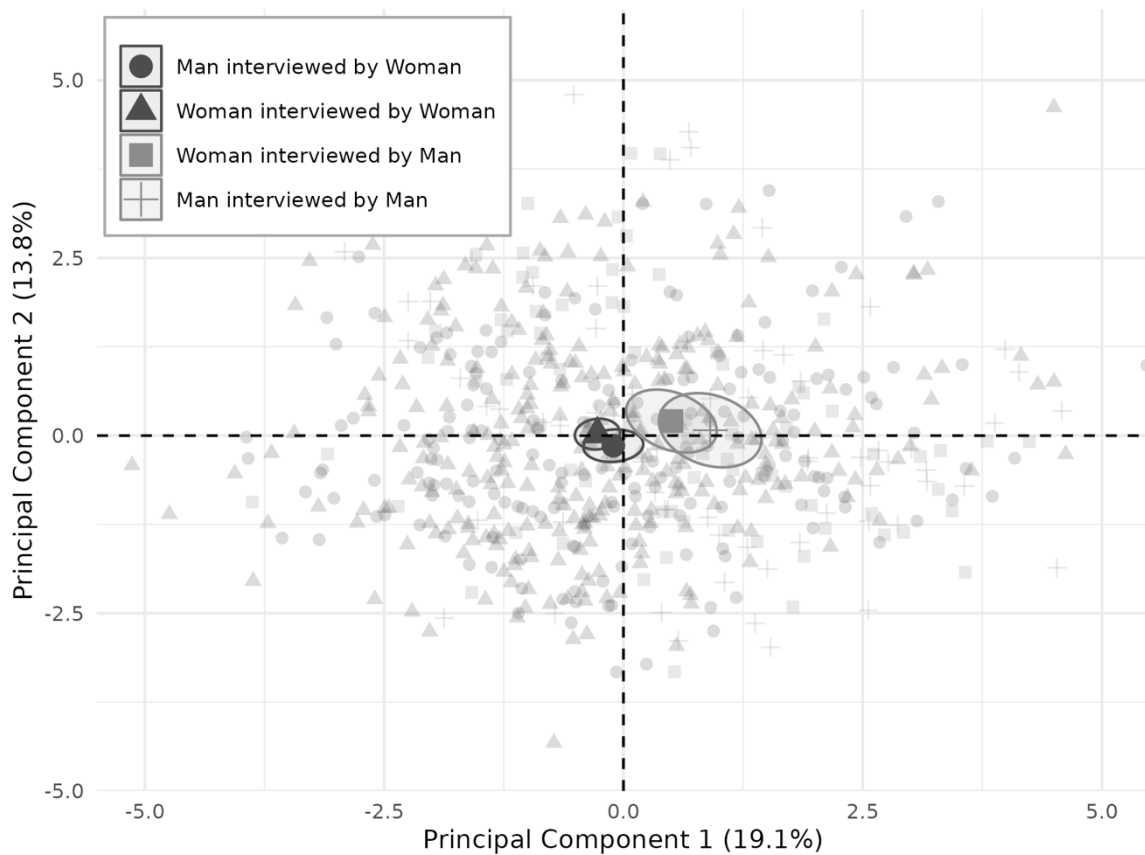


Figure 2: Axis 1 and 2 of the PCA on interview features, with interview gender configuration projected as an additional variable

5.2. Does PTSD diagnosis influence the interview structure?

Among the 926 participants, we selected the exposed population, following the methodology proposed by Quillivic et al. (2023). Among this population (N=148, see Table 3) we found that the interviewer asked significantly more questions related to time to participants without PTSD (p -value = 0.02, cohen = -0.32). Concerning the dialogic act categories, we found that the proportion of “probing” questions was inversely associated with the avoidance symptoms (p -value = 0.01, cohen = -0.32). One must recall that the interviewer does not know the diagnosis, at the moment of the interview. When using multivariate analysis, namely PCA, using only thematic and dialogic act features. Figure 3, shows that the diagnosis did not influence the structure of the interview, the ellipses do not show strong differences. Figure 4, confirms that avoidance symptoms are inversely associated with probing.

Variable	Cohort 5.2
N	148
% Female	58.0
% PTSD	53.0
% full or Partial PTSD	75.0
% Avoidance Symptoms	58.9

Table 3: Description of the population used in the 5.2 used case.

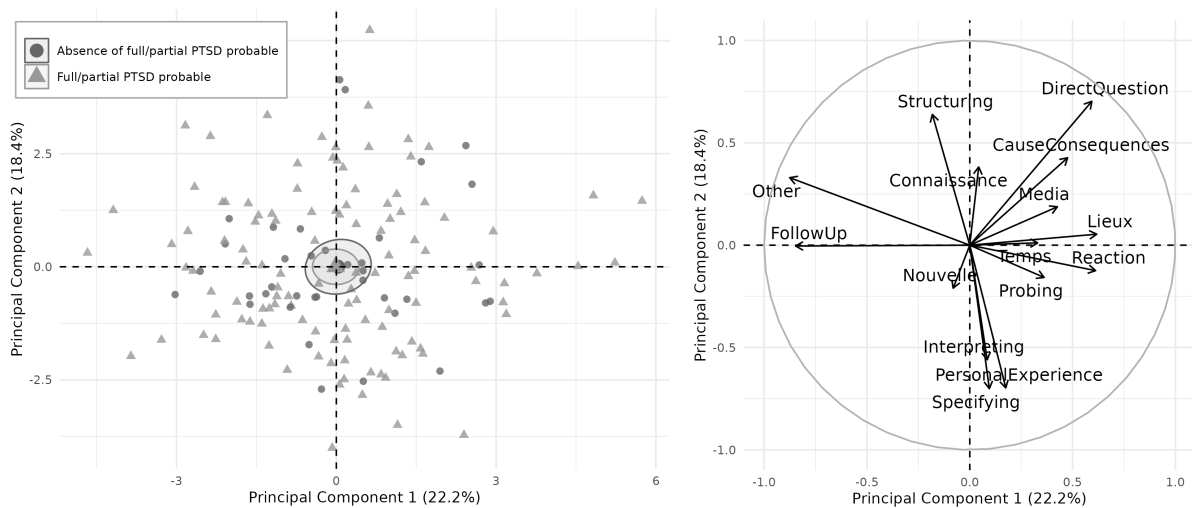


Figure 3: Axes 1 and 2 of the PCA, with on the left the visualization of the diagnosis of full or partial PTSD and on the right the features contribution to the axes.

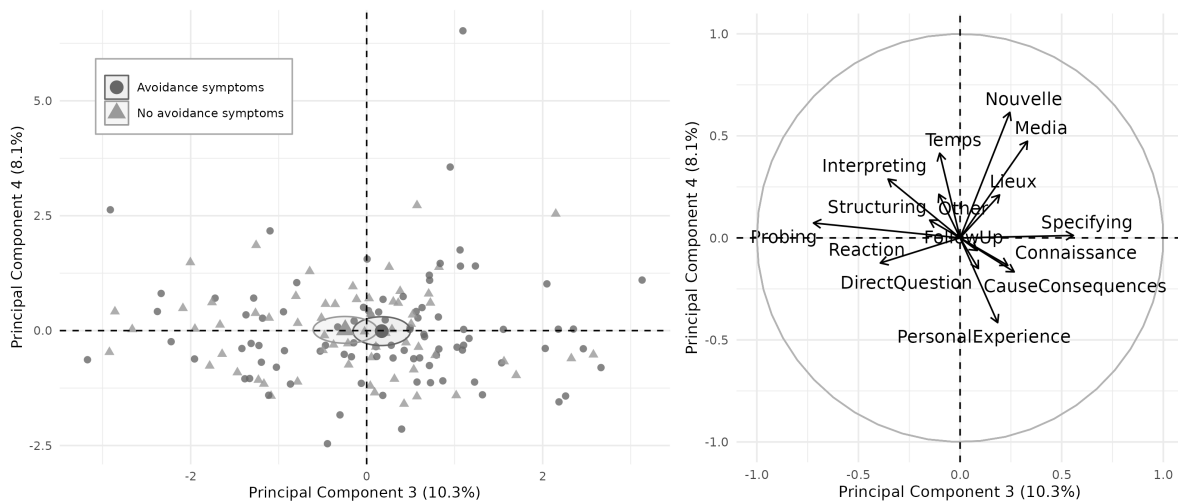


Figure 4: Axes 3 and 4 of the PCA, with on the left the visualization of the avoidance symptoms and on the right the features contribution to the axes.

6. Discussions

In this study, our primary objective was to develop a methodological framework for analyzing data from semi-structured interviews (SSIs), with a specific focus on understanding interaction dynamics. While our work is in its preliminary stages, we demonstrated the potential of recent advancements in Natural Language Processing (NLP) models through two specific examples. Specifically, we leveraged Sentence Embeddings to compute discourse alignment and utilized Multilingual Language Models in combination with the SetFit Framework to train few-shot learners. On the corpus of “Etude 1000”, we estimate how gender configuration influences the interview structure and quantified how PTSD symptoms impact the structure.

Given the vast amount of material generated by semi-structured interviews in the social sciences, our approach offers significant benefits for researchers seeking to compute quantitative measures on their datasets. Moreover, it provides valuable indicators of interaction to effectively navigate the corpus and facilitate qualitative analysis, aligning with the principles of the ADT (Textual Data Analysis).

To continue this work, future research could explore more sophisticated features that take into account the context of each utterance, as proposed by Zhang (2021) with the expected conversational framework. Additionally, our approach extends beyond SSIs and holds potential for application in other types of asymmetric linguistic interactions, such as job interviews. By embracing the concept of interactional dynamics, computational analysis of interview transcriptions presents exciting opportunities for uncovering insights and comprehensively understanding complex human interactions.

Funding

The study was conducted within the framework of the “Programme 13-Novembre”, funded by the French General Secretariat for Investment (SGPI) through the National Research Agency (ANR) and the “Programme d’investissement pour l’Avenir” (PIA ANR-10-EQPX-0021-01). It also receives the support of Region île-de-France through deliberation n°CP2021-297.

Bibliography

- Bonneau J. and Dister A. (2010). Logométrie et modélisation des interactions discursives. L’exemple des entretiens semi-directifs. *10^e Journées internationales d’Analyse statistique des Données Textuelles (JADT)*, Rome, 253-264.
- Brinkmann S. and Kvale S. (2018). *Doing Interviews*. London: SAGE Publications Ltd.
- Chang J. P., Chiam C., Fu L., Wang A., Zhang J. and Danescu-Niculescu-Mizil C. (2020). ConvoKit: A Toolkit for the Analysis of Conversations. In *Proceedings of the 21th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, The Association for Computational Linguistics, 57–60. <https://aclanthology.org/2020.sigdial-1.pdf>
- Dalud-Vincent M. (2010). Les << choix >> du sociologue avec Alceste—De la forme du corpus aux résultats obtenus. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 105 (1), 25–52.
- Danescu-Niculescu-Mizil C. and Lee L. (2011). Chameleons in imagined conversations: A new approach to understanding coordination of linguistic style in dialogs. In *Proceedings of the ACL workshop on Cognitive Modeling and Computational Linguistics*, 76-87. *arXiv*. <http://arxiv.org/abs/1106.3077>
- Danescu-Niculescu-Mizil C., Lee L., Pang B. and Kleinberg J. (2012). Echoes of power: Language effects and power differences in social interaction. *arXiv*. <http://arxiv.org/abs/1112.3670>
- Doyle G. and Frank M. C. (2016). Investigating the Sources of Linguistic Alignment in Conversation. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Berlin, 526–536.
- Eustache F. and Peschanski D. (2022). Toward new memory sciences: The Programme 13-Novembre. *Progress in Brain Research*, 274 (1), 177–201.
- Koven M. (2014). Interviewing: Practice, Ideology, Genre, and Intertextuality. *Annual Review of Anthropology*, 43 (1), 499–520.
- Nasir M., Chakravarthula S. N., Baucom B., Atkins D. C., Georgiou P. and Narayanan S. (2019). Modeling Interpersonal Linguistic Coordination in Conversations using Word Mover’s Distance. *Interspeech*, 2019, 1423–1427.
- Niven T. and Kao H.-Y. (2019). Detecting Argumentative Discourse Acts with Linguistic Alignment. In *Proceedings of the 6th Workshop on Argument Mining*, Florence, 104–112.
- Rapley T. J. (2001). The art(fulness) of open-ended interviewing: Some considerations on analysing interviews. *Qualitative Research*, 1 (3), 303–323.
- Reimers N. and Gurevych I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks, Hong Kong, 3982-3992. *arXiv*. <https://doi.org/10.48550/arXiv.1908.10084>

- Roulston K. (2019). *Interactional studies of qualitative research interviews*. John Benjamins Publishing Company.
- Scelles R. (1997). Apports de logiciels d'analyse de données textuelles (ADT) dans les procédures d'analyse de contenu d'entretiens semi-directifs de recherche: Alceste et hyperbase. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 57 (1), 25–48.
- Talmy S. (2010). Qualitative Interviews in Applied Linguistics: From Research Instrument to Social Practice. *Annual Review of Applied Linguistics*, 30, 128–148.
- Tunstall L., Reimers N., Jo U. E. S., Bates L., Korat D., Wasserblat M. and Pereg O. (2022). Efficient Few-Shot Learning Without Prompts. *arXiv*. <https://doi.org/10.48550/arXiv.2209.11055>
- Zhang J. (2021). *Towards Actionable Understandings of Conversations: A Computational Approach* [PhD Thesis, Cornell University].
- Zhang J. and Danescu-Niculescu-Mizil C. (2020). Balancing Objectives in Counseling Conversations: Advancing Forwards or Looking Backwards. *arXiv*. <https://doi.org/10.48550/arXiv.2005.04245>
- Zhang J., Spirling A. and Danescu-Niculescu-Mizil C. (2017). Asking too much? The rhetorical role of questions in political discourse. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, Copenhagen, 1558–1572.