Identifying Tension in Holocaust Survivors' Interview: Code-switching/Code-mixing as Cues

Xinyuan Xia, Lu Xiao, Kun Yang, Yueyue Wang

Department of Languages, Literatures, and Linguistics, School of Information Studies, College of Arts and Science,
Syracuse University
{xxia16, lxiao04, kyang23, ywang153}@syr.edu

Abstract

In this study, we thrive on finding out how code-switching and code-mixing (CS/CM) as a linguistic phenomenon could be a sign of tension in Holocaust survivors' interviews. We first created an interview corpus (a total of 39 interviews) that contains manually annotated CS/CM codes (a total of 802 quotations). We then compared our annotations with the tension places in the corpus. The tensions are identified by a computational tool. We found that most of our annotations were captured in the tension places, and it showed a relatively outstanding performance. The finding implies that CS/CM can be appropriate cues for detecting tension in this communication context. Our CS/CM annotated interview corpus is openly accessible. Aside from annotating and examining CS/CM occurrences, we annotated silence situations in this open corpus, as it is shown to be an indicator of tension in interpersonal communications in previous research. Making this corpus openly accessible, we call for more research endeavors on tension detection.

Keywords: Code-switching, Code-mixing, Silence, Tension, Holocaust Interviews

1. Introduction

"Tension" is defined as "inner striving, unrest, or imbalance often with the physiological indicators of emotion," or "a state of latent hostility or opposition between individuals or groups" on Merriam-Webster dictionary. Therefore, when it comes to tension, it always revolves around feelings and emotions. The state of tension could be affected by psychological, phonological, and morphological aspects. In interpersonal communications, people sometimes feel uncomfortable responding to a question or discussing a topic brought up by their conversation partner, who, on the other hand, is eager to explore. They may tolerate the situation, and their responses do not reveal their discomfort. Or they may use various strategies in their responses consciously or subconsciously, such as responding in silence or switching the topic. In the second case, internally, this discomfort experienced by the people may be identified based on how they communicate and what they communicate. We call such situations in communication the tension states. Depending on the communication context, the flow can be negatively affected when tension occurs, and the communication goal may not be achieved (Sins & Karlgren, 2009). With more and more communication data available in electronic form, researchers explore computational approaches that automatically identify tensions from such data (Islam, Mercer, Xiao, & High, 2019), and tension analysis in the data is getting research attention (e.g., Burnap et al., 2015).

This study is particularly interested in identifying and analyzing tensions in survivor or witness interviews about mass violence. It is expected that tensions will occur in these Holocaust survivors' interviews. Specifically, these personal stories and testimonies about mass violence are often explored through life story interviews in oral history (Bornat, 2003). During the interview, the interviewer and interviewee co-construct a conversational space that offers the context for interviewees' stories and feelings. Researchers build rapport to gain trust to open up this conversational space (Owens, 2006), e.g., by starting an interview with general questions (Stevens, Lord, Proctor, Nagy, O'Riordan, 2010). As the interview continues, the interviewer intends to get closer to the interviewee's inner thinking and feelings, whereas the interviewee (intentionally or unintentionally) pushes other people away from that part of the memory or feeling on the violent event. This intrinsic conflict adds momentum to the conversation dynamics in the interview process. Other factors exist as well, making the interviewer and the interviewee go in different directions in this conversation space. For example, suppose the interviewer lacks similar experiences. In that case, it can be challenging for the interviewer to incorrectly interpret the interviewee's feelings, leading to the improper or incorrect assumptions at the beginning of or throughout the interview and making wrong moves. Tensions can easily occur in these scenarios during the interview. Identifying and analyzing them is especially important to understand the interview dynamics better. In a tension state, interviewees may prefer not to discuss a given topic or even challenge the validity of a question being posed (Donovan-Kicken et al., 2011). Or they may reframe the question or re-direct the conversation in another direction (Greenspan, 2010). These points of tensions are not problems that need to be 'fixed,' though it is helpful for interviewers to be able to read these situations. The ability to identify and analyze tensions in these interviews is essential in better understanding the agency of the interviewee and the underlying interview dynamic itself (Tripp, 1983; Koro-Ljungberg, 2008; Tanggaard, 2009).

The interview corpus in our study is obtained from United States Holocaust Memorial Museum in Washington D.C. Interviewees are survivors of the Holocaust, and the interviews are about that history. In this context, we consider interviewees' tensions that are adverse emotion fluctuations or reluctant attitudes caused by specific interview questions or their memories. Also, in this context, interviewees sometimes switched between English and their mother language or second language in the conversation, which is usually a sign of the linguistic phenomenon of code-switching (CS) or code-mixing (CM); and sometimes were silent in their responses. We explored whether the linguistic phenomenon of CS/CM and silence signify tensions in the interviewee's experience at that moment.

The rest of the paper is organized as follows. We first review the linguistic cues explored in tension detection literature as well as the concept of CS/CM and related studies. We then introduce our annotated interview corpus, including the description of the interviewees, the annotation process, and the reliability check of the annotations. We next compare the annotation results with the tension places identified in the data, which are identified by a tool developed by Islam, Mercer, and Xiao (2019) that detects tensions in survivor and witness interviews. We then discuss the implications of our study and suggest directions for future research.

2. Literature Review

2.1 Linguistic cues of tension in interpersonal communication

While discussing identities, emotions, and linguistic decision making, Pavlenko (2005) had noted that "Emotions often fly high in multilingual contexts, where language disputes constitute major sources of tension and instability." A tension state can have many aspects, such as psychological, phonological, and morphological sides. Various cues that could signal the presence of tension. In analyzing spikes in tension in social media content, Burnap et al. (2015) illustrated how lexicons of abusive or expletive terms could identify high levels of tension separated from low levels. Their proposed tension detection engine relies on the lexicons and membership categorization analysis (MCA) (Sacks et al., 1995). As a negative affective condition that people experience when they feel upset, distress is closely related to tension. McCubbin & Sussman (2014) discussed how stressor events produce tension and how stress becomes distressed when it is subjectively defined as unpleasant. Buechel et al. (2018) used a CNN system for detecting distress with Fast-Text word embeddings as inputs.

To our best knowledge, the most recent tension detection framework and the most relevant to our study is by Islam, Mercer, Xiao, and High (2019). Focusing on identifying tensions in interviews about genocide, the researchers first used a multi-channel CNN algorithm (Islam, Mercer, & Xiao, 2019) to recognize emotions in the content of an interviewee's response. Basic negative emotions such as sadness and anger could be the cues for tension (Cherry, 2021). Researchers then examined additional tension cues in a negative emotion place. These cues include hedge words (e.g., *probability, likely*), booster words (e.g., *definitely, never*), silence, laughter, the length of an answer, and a list of discourse markers identified by the researchers[1]. The researchers' examination of these cues was based on the relevant literature. For instance, silence is a critical discursive indicator of various phenomena in oral history interviews, such as reticent responses or, in turn, unnatural pauses caused by personal trauma. (Layman, 2009; Matei, 2013). The use of discourse markers such as *not really*, *not that I remember* or *well, anyway* in responses shows how reticence in an interview might be influential. It is also a common strategy embraced by the interviewees to avoid either complete refusal to reply or full disclosure, which can be revealed by the unusual short or long answers given the question.

2.2 Code-Switching and Code-Mixing

Even though the primary language of our interview corpus is English, there are many different languages interspersed between English in the interviews. This phenomenon is called Code-Switching/Code-Mixing (CS/CM), which means that mixed languages are an outcome of switching or mixing between two or more languages (codes); or alternate between at least two languages or languages varieties (Van Herk, 2012).

There are a few related studies about the phenomenon of CS/CM in Holocaust interviews. Firstly, code-switching and language choice can influence speakers' expressions of positive and negative feelings. Bilinguals or multilinguals tend to use different languages to express different emotions, and the feelings are mixed (Pavlenko, 2005). For instance, Pavlenko's case study (2005) in first language rejection of German Jews and the Holocaust shows that the survivors' longing for their first language (L1) is mixed with hatred. The feelings for using their L1 are intimated childhood remembrances mixed with painful memories of atrocities and losses. Furthermore, in Müller (2014)'s article discussing translations of David Boder's 1946 Interviews with Holocaust Survivors, he also pointed out that interviewees' sudden switch of language at some crucial moment, such as talking about the losses of family members, could be viewed as a sign of trauma or expressing strong feelings.

Building on this literature, we speculate that CS/CM can be a cue for detecting tension in Holocaust survivors' interviews. We also note that the boundaries between CS and CM may be blurred or fixed depending on different analysis purposes. In this study, we view CS/CM as a whole.

3. The Development of the CS/CM Annotated Interview Corpus

Our Holocaust interview corpus is obtained from the United States Holocaust Memorial Museum in Washington, D.C. It contains 39 interview data, including the video recordings of the interviews and the transcripts of the interview content. Of the 39 interviews, 11 were within one hour, 2 were between one and two hours, and 12 were over two hours long (within three hours). The demographics information and interviews date were also provided in our Corpus, as shown in Table 1 below (interviewees' age was calculated according to their DOB information and the interview date).

Document	Gender	Age	Interview Date	Video length
RG-50.030.0001	Male	66	1994 July 28	2h58min

Table 1: Example of the provided information

An initial examination of the corpus shows that the transcripts do not have timestamps for questions and answers, which makes it challenging to associate a video

---

[1] Please refer to the appendices of this Master's thesis for a list of cues used in the researchers' work: https://ir.lib.uwo.ca/etd/5878/

snippet with its corresponding transcript. It is, however, important that we can leverage the video recordings while annotating CS/CM in the transcripts. Hence, the first step in our data processing was to add timestamps for each conversational turn in the transcripts. We next describe how we annotated CS/CM and conducted the inter-coder reliability check in more detail.

3.1 Code-Switching and Code-Mixing (CS/CM) Annotation

In our annotation, we adopted the definition of CS/CM by the Handbook of Second Language Acquisition such that CS/CM is the mixing of various linguistics units such as "words, phrases or sentences" from two different language systems, either across or within sentences (Bhatia & Ritchie, 1996). We also excluded any proper noun from CS/CM because for instance, names of people and places always are pronounced and spelled as their originals and usually viewed as loanwords or borrowings in languages. In the annotation process, we took videos as auxiliary instruments and annotated the sentence with CS/CM code whenever a CS/CM appeared within the sentence. If the sentence had clauses and was too long, we annotated the clause that contained the CS/CM (as the second example in Table 2). As shown in Table 2, we have two codes for each annotated content: one is CS/CM, which indicates the existence of code-switching/code-mixing phenomenon in the annotated content; and the other is the language that notes the mixed or switched language. If we could not identify the mixed language, we have the code *not sure* for this situation. Besides these two codes, we also provided a basic translation for the CS/CM parts. We utilized Google translate and the Holocaust Museum information archive to help determine the language source and meaning of each CS/CM.

Document	Quotation Content	Comment	Codes
RG-50.030.0028	My father lost his business, he had from before, as a legionnaire, he had a liquor store business, so-called "monopol koncesja" in Poland, uh...as a veteran from the legionnaires.	monopol koncesja means Monopoly concession	CS/CM Polish
RG-50.030.0029	The Straf company is the...everybody knows what a straf company...company was,	Straf means punitive	CS/CM German

Table 2: Example of the Annotation Contents

The annotation process involved two research assistants – the first and last authors. The first author holds a master's degree in Linguistics and the last author is a senior undergraduate in Information Science and Technology. The first author annotated CS/CM in the interview data and trained the last author on the CS/CM annotation. There was only enough time for the last author to independently annotated CS/CM in 16 interviews. The Krippendorff's c-α-binary value is 0.884 indicating the reliability between two annotator's coding results. Then, the first author re-annotated the rest 23 interviews also for the sake of double

checking, and the intra-coder reliability is measured and supported by Krippendorff's c-α-binary value as well (0.889). The two annotators discussed all identified discrepancies together and reached agreements afterwards. The updated annotation results are used for later tension analysis.

The annotation results show that there are 802 annotated quotations in 39 transcripts. A total of 22 languages has been detected for CS/CM, and the most frequently occurring languages are German, Polish, Hebrew, and Russian. Combining the annotation results with the demographic information provided, we could see in Table 3 that more than half of the interviewees were male, and among all most interviewees were between the age of 61 to 70 (the age total for the female is 14 because one of the interviewees didn't provide DOB information).

	Gender		Age			
	Female (15)	Male (24)	50-60 (7)	61-70 (23)	71-80 (6)	81+ (2)
CS/CM	221	581	64	477	222	24

Table 3: CS/CM annotations' distribution in gender and age range

4. Code-Switching & Code-Mixing (CS/CM) as a Tension Cue

We are interested in whether and how the use of CS/CM by an interviewee was related to the tension she/he was experiencing. It helps us understand the potential of CS/CM as a tension cue by comparing whether and how much tension places and CS/CM occurrences overlap. In this comparison, we leveraged the tension detection tool by Islam, Mercer, Xiao, and High (2019) to identify the tensions given the annotated interview transcripts. As described in the literature review section, this tool is trained on the interview transcripts about mass violence, and it detects tensions in the interviewees' responses based on various information such as the emotions recognized in the text, multiple language cues (e.g., booster words, hedging words, and the length of the interview response), and discourse markers (e.g., laughter, silence, etc.).

Our comparison was conducted as follows. An interviewee's response is labeled as *tension* or *not tension* by the tool. In total, there are 1095 *tensions* and 3362 *not tensions* in the interview corpus. Based on our annotation results, if the response includes one or more CS/CM occurrences, we marked it as *CS/CM,* otherwise *no CS/CM*. There are 275 *CS/CM* and 4182 *no CS/CM* responses. Table 4 shows how these two categories overlap in the corpus. The chi-square test shows that the existence of tension and that of CS/CM are indeed correlated (the chi-square value is 232.963, p <0.001).

	Tension responses	Not-Tension responses
CS/CM responses	172	103
No-CS/CM responses	923	3309

Table 4: Contingency table for CS/CM and Tension responses

We also explored whether and how the occurrences of CS/CM are captured by the tension detection tool. For each transcript, we noted the total number of CS/CM responses and the percentage of them captured as tensions.
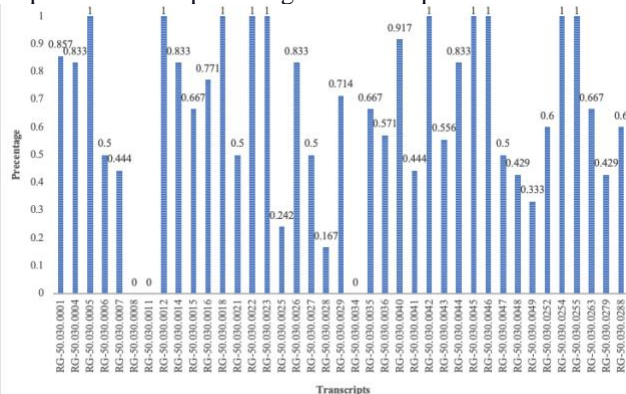

Figure 1: Percentage of the CS/CM responses in tensions

Figure 1 shows that amid 39 transcripts, 11 had all their CS/CM occurrences captured by the model, and 19 had over 50 percent coverage. In interview transcript RG-50.030.0008, only one CS/CM annotation and response did not appear in the identified tension place. The tension detection tool did not identify any tension in RG-50.030.0011, but there was 1 CS/CM response (2 annotations) in the transcript. Hence the percentage coverage was also 0 in this case. As for RG-50.030.0034, there were 3 CS/CM responses (40 annotations) in the transcript, but none overlapped with the only tension response identified in the transcript. Overall, the comparison suggests that CS/CM could be a cue for detecting tension in Holocaust survivors' interviews.

4.2 Code-Switching & Code-Mixing (CS/CM) in No Tension Places

There are 342 CS/CM annotations in the CS/CM responses that have no tension identified by the tension detection tool. We categorized them into three main cases according to our interpretation of their occurrences (see table 5 below).

Translation	**186**
Glossary	**117**
Scene Representation	**22**
Glossary & Translation	6
Scene Representation & Translation	11

Table 5: Summary of each category for CS/CM annotations appeared in the interview responses that have no tension identified

The Translation category refers to three situations where CS/CM occurred: 1). the interviewees came up words or phrases that they could not find the equivalents in English at that moment; e.g., "He figured that probably that I am not his in Polish they call it *wychowane dzieci* (means pupil or foster-child /Polish). He did not raise me, you understand?" (RG-50.030.0001) 2). The foreign words came up naturally with the flow of their storytelling as the background information; e.g., "It was the First of Mai (means May/German) '45. (RG-50.030.0029) 3). The interviewees first said the foreign words or sentences, and then gave their interpretations in English; e.g., "We used to have what do you call *kukiełki* (means puppets/Polish), all kind of marionettes, you know." (RG-50.030.0001)

The Glossary category refers to the case in which CS/CM appeared without providing the translations or equivalents by the interviewees. These often happened in one of the three situations. One is that the words or phrases were commonly used in concentration camps or during that historical period hence were familiarized by people who have adequate knowledge of this history, e.g., *Appellplatz*: the German word for roll call square where prisoners were forced to assemble (USHMM, 2021). The second case is when the words or phrases were specifically about the camps, German officers, barracks, etc. Examples of this case are the words for German officers' rankings and camp work departments. The third situation is that the words or sentences were derived from religions such as "*Shema Yisrael*" means "Hear, O Israel" is a Jewish prayer that serves as a central piece for the Jewish prayer services.

The Translation and Glossary categories account for the majority of the CS/CM annotations in the *no tension* interview responses. This situation is expected because the reasons for switching/mixing the languages in these cases were the interviewees' familiarity (or lack of) with the linguistic terms and foreign language as opposed to their affective or emotional state (e.g., stress). Individual factors like language proficiency and dominance are all essential elements of language choice (Hudson, 1980) in all kinds of communication contexts and, in our case, in the process of reviewing and narrating the past life or memories (Pavlenko, 2005).

The third category is Scene Representation that refers to the situations when the interviewees were reliving the scenes happened in the past; they restored the scenes in their original settings, including having the past conversations represented in the original languages. Here is one example - "I witnessed one scene where a uh a tailor by the name of Singer whom I had known in the ghetto uh was on his knee before a German and said, h*err Gestapo, gib es mir zurück bei Frauen beim Kind*. (means Lord/Mr. Gestapo, give it back to me with women with children/give back my women and children/German)" (RG-50.030.0021). Prior study suggests that when reliving painful, frightening, or disturbing situations people may have a sudden switch of language at that crucial moment, which could be viewed as a sign of trauma (Müller, 2014). We anticipate that the interviewees may feel tension when reliving specific memories.

As the current tension detection tool that utilizes the language cues and emotions failed to identify the tensions in these places, we suggest the exploration of detecting a story structure as an approach for tension detection in these types of interviews about violence in which people may relive the past scenes and describe them to the interviewers.

5. Silence Annotation

Besides annotating code-switch and code-mixing (CS/CM) in the interview data, we also annotated silences. Silence is shown to be an indicator of tension in sensitive interviews (Laymen, 2009; Matei, 2013). Our examination of the interview data shows that silence could appear as in-turn long or unnatural pauses. It usually lasts at least about 4 seconds and has no maximum limit. To annotate silence in the data, we listened to the video recordings measuring the duration of the silence, identified the boundary for each silence code, and included two words before and after the

silence interval. Our annotations also noted the length of silence. Here is an example in table 6.

Document	Quotation Content	Comment	Codes
RG-50.030.0001	Ghetto. They	4s	silence

Table 6: Example of the silence annotation contents

The annotation process and annotated documents for Silence were the same as for CS/CM, involving two annotators, the first author, and the last author. The reliability between the two annotators' coding results is 0.859, indicated by Krippendorff's c-α-binary value. And the intra-coder reliability for the first author is measured and supported by Krippendorff's c-α-binary value as well (0.91). The few disputes between the two annotators were also discussed and solved since the annotation for Silence was pretty straightforward. The final results showed that among 39 transcripts, 21 had silence appeared, and the total annotations were 103.

Moreover, we did not go further in making the comparative analysis of the silence cue was due to the small amount of it we found in our corpus. The comparison results will be more reliable and credible when all interviews in the corpus contain silence.

6. Discussion

The most time-consuming tasks in our study were to find CS/CM and to add the timestamps to the interview transcripts so that the video recordings and transcripts could be synched in the analysis. In addition, our CS/CM annotation was conducted manually mainly because of the complexity of the language context. Specifically, there are over 20 languages involved, and the translation part requires proper background linguistics knowledge. Moving forward, a combination of computational and manual data processing and annotations is expected to speed up the process. Nevertheless, our study shows the potential of using CS/CM as cues in identifying tensions in Holocaust interviews. Furthermore, our analysis of the CS/CM annotations not captured in the tension places identified by the tension detection tool suggests that CS/CM alone is not sufficient to signify tensions, and the detection of story structure in Holocaust interviews may contribute to the detection of tensions in these communication record. Also, it would be essential to investigate the potential false positives in the dataset to improve and perfect the tool's performance and result.

And, our annotations included the silence in the interview data – an important tension detection cue. We make the annotated interview corpus open accessibility to the research community,[2] calling for further investigation on this topic.

Our study only explored the potential of CS/CM occurrences for tension detection. Previous studies imply that other language features may be important to investigate as well, and some have not been examined in the tension detection tool by Islam, Mercer, Xiao, and High (2019). For instance, while acknowledging that in the interviews about mass violence tension tends to occur

when the interviewee experiences negative emotions, their tool only recognizes the negative emotions in the interviewee's response content. Other important emotion cues in video data are not leveraged such as, crying, sobbing, sighing, changing the tone, etc. A previous study analyzing the audio recordings of interview data suggests that prosodic features can be indicative of tensions in interviews (Zhang & Xiao, 2020). In our data context, we speculate that the following cues may contribute to the detection of tension: a sudden 'rise' in intonation; a sudden 'drop' in intonation; a sudden change of tone to emphasize something important, surprising, or to express certain strong emotion. Consistent with Islam, Mercer, Xiao, and High (2019)'s study, we also noticed the use of laughter by the interviewees which may express a feeling of nervousness and uneasiness in this interview context.

The orality of discourse in conversations has many reformulations and self-interruptions such as glottal stops, direct speech, fillers, and many other discourse markers that usually appear more than usual between an apparent strong personal and emotional involvement and stance-taking, to deliberately or seemingly detached themselves from their narration (Busch & McNamara, 2020; Matei, 2013). Our reading of the interview data suggests that some of these cues may be explored for their potential for detection tensions. For example, we noticed that interviewees might use repeated fillers/words/phrases in the conversation and sometimes it was too much than expected. Also, the interviewees' words and sentences sometimes tangled up, implying a sense of fluster and anxiety about the mentioned topics.

7. Conclusion

We annotated 39 Holocaust interview data for analyzing and discussing the potential of the linguistic phenomenon of code-switching/code-mixing (CS/CM) as a cue for detecting tension in survivor or witness interviews. Our data are obtained from United States Holocaust Memorial Museum in Washington D.C.. We first added timestamps for each conversational turn. We then annotated CS/CM in the interviewees' responses. We compared our annotation results with those tension places identified from the tension detection tool (Islam, Mercer, Xiao, & High, 2019). The comparison suggests that CS/CM and tension have a strong correlation, and this linguistic phenomenon could be a cue for detecting tension in Holocaust survivors' interviews. We observed that interviewees may use mixed languages in non-tension situations such as for the purpose of providing a glossary. Our observation also suggests that the current tension detection tool may fail to detect tension places even when there are CS/CM cases, and this likely happens when the use of CS/CM is situated in a storytelling structure.

To encourage further research investigations on the detection of tensions in survivor interviews, we make the annotated corpus openly accessible. Besides the CS/CM annotations, this corpus also has silence annotated as it has been shown that silence is a good indicator of tension in such interview context.

8. Reference

---

[2]https://drive.google.com/drive/folders/1GE2DD5Y3KS8lfmPwSsWmYq6-gfBy1dbg?usp=sharing

Bornat, J. (2003). A Second Take: Revisiting Interviews with a Different Purpose, Oral History, 31(1):47-53.

Buechel, S., Buffone, A., Slaff, B., Ungar, L. & Sedoc, J. (2018), 'Modeling empathy and distress in reaction to news stories', arXiv preprint arXiv:1808.10399.

Burnap, P., Rana, O. F., Avis, N., Williams,M., Housley, W., Edwards, A., Morgan, J. & Sloan, L. (2015), 'Detecting tensionin online communities with computational twitteranalysis', Tech- nological Forecasting and Social Change 95, pp. 96–108.

Busch, Brigitta. McNamara, Tim. (2020). Language and Trauma: An Introduction, Applied Linguistics, 41(3), 323–333, https://doi.org/10.1093/applin/amaa002.

Cherry, Kendra. (2021). The 6 Types of Basic Emotions and Their Effect on Human Behavior. https://www.verywellmind.com/an-overview-of-the-types-of-emotions-4163976 Accessed on June, 2021.

Donovan-Kicken, E., Guinn, T. D. , Romo, L. K., & Ciceraro, L. D. (2011). Thanks for asking, but let's talk about something else: Reactions to topic-avoidance messages that feature different interaction goals, Communication Research, 40(3): 308-336.

Eubanks, C. F., Muran, J. C., & Safran, J. D. (2014). Rupture resolution rating system (3RS): Manual.

Greenspan, H. (2010). On Listening to Holocaust Survivors: Beyond Testimony, 2nd ed., St. Paul: Paragon House.

Islam, J., Xiao, L., Mercer, R., & High, S. (2019). Tension Analysis in Survivor Interviews: A Computational Approach, 2019 Digital Humanities (Jul 9 – 12, Utrecht, the Netherlands)

Islam, J., Mercer, R. E. & Xiao, L. (2019), Multi-channel convolutional neural network for twitter emotion and sentiment recognition, in 'Proceedings of the 2019 Conference of the North Amer- ican Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1', pp. 1355–1365.

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed).

Conversation Analysis: Studies from the First Generation. 13-31. Amsterdam: John Benjamins.

Koro-Ljungberg, M. (2008). A social constructionist framing of the research interview. Handbook of constructionist research, 429-443.

Layman, Lenore. (2009). Reticence in oral history interviews. The Oral History Review. 36(2), 207–230.

Matei, Mădălina-Georgiana. (2013). The Linguistic Mechanisms of Trauma Discourse.

Procedia – Social and Behavioral Sciences. 92, 517-522. https://doi.org/10.1016/j.sbspro.2013.08.711. (http://www.sciencedirect.com/science/article/pii/S1877042813028425).

Mauriello, Matthew Louis; Lincoln, Thierry; Hon, Grace; Simon, Dorien; Jurafsky, Dan and Paredes, Pablo. (2021). SAD: A Stress Annotated Dataset for Recognizing Everyday Stressors in SMS-like Conversational Systems. In Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21). Association for Computing Machinery, New York, NY, USA, Article 399, 1–7. DOI:https://doi.org/10.1145/3411763.3451799.

Mc Cubbin, H. I. & Sussman, M. B. (2014), Social Stress and the Family: Advances and Developments in Family Stress Therapy and Research, Vol. 6, Routledge.

Müller, Beate. (2014). Translating Trauma: David Boder's 1946 Interviews with Holocaust Survivors. Translation and Literature. 23(2), 257-271. Available Online May 2014.

Noone, Peter. (2017).The Holmes–Rahe Stress Inventory. Occupational Medicine. 67(7), 581–582.

Owens, E. (2006). Conversational space and participant shame in interviewing. Qualitative Inquiry, 12(6), 1160-1179.

Pavlenko, Aneta. (2005). *Emotions and bilingualism*. Cambridge, UK: Cambridge University Press.

Ritchie, William C., and Tej K. Bhatia. (1996). Handbook of Second Language Acquisition. 637-667. Academic Press, San Diego.

Rodero, Emma. (2011). Intonation and Emotion: Influence of Pitch Levels and Contour Type on Creating Emotions. Journal of Voice. 25 (1), 25-34. https://doi.org/10.1016/j.jvoice.2010.02.002.(https://www.sciencedirect.com/science/article/pii/S0892199710000378).

Sacks, H., Jefferson, G. & Schegloff, E. A. (1995), Lectures on Conversation, Wiley-Blackwell.

Sins, P.H.M., Karlgren, K.: Identifying and overcoming tension in interdisciplinary teamwork in professional development: two cases and a tool for support (2009).

Stevens, M. M., Lord, B. A., Proctor, M. T., Nagy, S., & O'Riordan, E. (2010). Research with vulnerable families caring for children with life-limiting conditions. Qualitative health research, 20(4), 496-505.

Tanggaard, L. (2009). The research interview as a dialogical context for the production of social life and personal narratives. Qualitative Inquiry, 15(9), 1498-1515.

Tripp, D. H. (1983). Co-authorship and negotiation: The interview as act of creation. Interchange, 14(3), 32-45.

Van Herk, Gerard. (2012). What Is Sociolinguistics? 1st Edition. 127-146. Blackwell Publishing Ltd. https://www.merriam-webster.com/dictionary/tension Accessed on June 18th, 2021.

USHMM. Glossary. Holocaust Encyclopedia. https://encyclopedia.ushmm.org/content/en/article/glossary. Accessed on Oct, 18, 2021.