



Open Library of Humanities

## Tension Analysis in Survivor Interviews: A Computational Approach

**Jumayel Islam**, The University of Western Ontario, Canada, [jjislam3@uwo.ca](mailto:jjislam3@uwo.ca)

**Robert E. Mercer**, The University of Western Ontario, Canada, [mercer@csd.uwo.ca](mailto:mercer@csd.uwo.ca)

**Lu Xiao**, Syracuse University, USA, [lxiao04@syr.edu](mailto:lxiao04@syr.edu)

**Steven High**, Concordia University, Canada, [steven.high@concordia.ca](mailto:steven.high@concordia.ca)

---

This study aims to develop computational techniques to analyze and identify points of tensions in interviews with survivors of the 1994 Rwandan genocide. Oral history interviews are a dialogical source composed of questions and answers, producing a conversational narrative. Yet survivor testimony is often approached as though the questions did not exist. This article examines a digital tool that helps us visualize and better understand the underlying interview dynamic that is the heart of oral history and qualitative research more generally. Our tension detection tool identifies those moments in the interview when the interviewer and interviewee are trying to pull the conversation in different directions. This is part of the natural give-and-take of the interview. Hedging, deflection, hesitation, and boosting are all critical components of this interviewer-interviewee tension. By making the interview dynamic central to our analysis, we aim to better understand how the interview dynamic shapes what is being said and what is left unsaid. In this study, we address key components of interview tension and propose a natural language processing model that can efficiently incorporate these components in text-based oral history interviews to identify tension points. With experiments on an annotated transcript, we verify the efficacy of our model. This model provides a framework that can be utilized in future research on the dialogic of the interview.

---

Cette étude vise à développer des techniques computationnelles pour analyser et identifier des points de tensions dans des interviews avec des survivants du génocide rwandais en 1994. Les interviews d'histoire orale sont une source dialogique composée de questions et de réponses, ce qui produit une narration conversationnelle. Cependant, le témoignage de survivant est souvent traité comme si les questions n'existaient pas. Cet article examine un outil numérique qui nous aide à visualiser et à mieux comprendre la dynamique d'interview sous-jacente qui est au cœur de l'histoire orale et, plus généralement, au cœur de la recherche qualitative. Notre outil détecteur de tensions identifie ces moments dans l'interview lorsque l'intervieweur et l'interviewé sont en train d'essayer de guider la conversation dans des directions différentes. Cela fait partie de l'interaction bidirectionnelle naturelle



d'une interview. Le non-engagement, le détournement, l'hésitation et l'exagération sont tous des composants essentiels dans la tension intervieweur-interviewé. En mettant la dynamique d'interview au centre de notre analyse, nous aspirons à mieux comprendre comment la dynamique d'interview structure ce qui est dit et ce qui n'est pas dit. Dans cette étude, nous abordons des composants clés de la tension d'interview et proposons un modèle de traitement de langue naturelle qui peut incorporer de façon efficace ces composants dans des interviews d'histoire orale numérisées afin d'identifier des points de tensions. Avec des expériences sur un transcrit annoté, nous vérifions l'efficacité de notre modèle. Ce modèle fournit un cadre à utiliser dans de futures recherches sur la dialogique de l'interview.

---

## 1 Introduction

Survivor testimony is central to our understanding of mass violence and its consequences. More often than not, however, this testimony has been treated by researchers as eyewitness accounts rather than as interviews. As a result, the researcher's questions are usually suppressed in the analysis and rarely included in published excerpts of these first-person accounts. To do so risks undermining their experiential authority. The interview context is effectively hidden or obscured. Yet, an interview is a dialogical process between the interviewer and interviewee, and the resulting question-and-answer structure largely determines what is and is not said. The interview dynamic is therefore central, leading some oral historians to call the recorded interview a "conversational narrative," as it is effectively co-produced (Grele and Terkel 1991, p. 135). How then can researchers analyze the interview dynamic to better understand how it influences survivor testimony?

An interview is a dialogical source consisting of questions and answers. It is essential that we understand better the role played by the interviewer in directing the conversation, but also in understanding the agency of the interviewee and the underlying interview dynamic itself (Tripp 1983; Koro-Ljungberg 2008; Tanggaard 2009). The interview dynamic is influenced by many factors, including the social and political distance between the interviewer and the interviewee and the chemistry between the two. There is no perfect location for the interviewer, but it is essential that we understand what their positionality affords and forecloses. Back in the 1930s, US President Franklin D. Roosevelt's administration undertook an oral history project with Americans who experienced slavery before 1865. One black man was interviewed twice by mistake; one interviewer was a white woman, and the other was a black man: You wouldn't know it was the same interviewee. Knowing this, we realize that these interviews reveal much about race relations in the US South during the 1930s, a period of Jim Crow segregation and widespread lynching of Black men. It is therefore important that we consider the interview context (Davidson and Lytle 2004). A genocide survivor interviewing another survivor will not be the same as an interview conducted by someone who did not experience it firsthand. Similarly, an interview between family members will not be the same as between strangers. This is not to say that one or the other is "better" positioned, but it influences and shapes the resulting conversation in myriad ways. Gender, race, and class all have immediate bearing on the intelligibility of this "mutual encounter" (Portelli 1991). To what degree, and at what point, are the interviewer and interviewee on the same "wavelength"? How is trust built over the course of the interview? At what point do the interviewer and interviewee struggle to connect, be heard, or work at cross-purposes? Put simply, tension is a "strained state or

condition resulting from forces acting in opposition to each other.” 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. 2013). 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 useful for interviewers to be able to read these situations. Our tension tool enables us to visualize the underlying interview dynamic and the ways that this conversation structures the transcribed life story. It also contributes to a more grounded oral history training.

The tension tool helps the researcher map the interview relationship and understand this very important interplay between what is asked and what is answered. Knowledge of where tension arises also offers a new way of investigating interview data. Where is tension most likely to surface in an interview? Are certain types of questions more likely to generate tension? How does an interviewer’s positionality and their social distance from the interviewee influence the interview dynamic and the relative presence of tension points? What do we learn about interviewee agency and the co-creation process in the process? These are just a few research questions that will allow us to better interpret qualitative interviews as a dialogical source and will make a significant original contribution to understanding survivor narratives and improving our training of potential interviewers.

Tension expresses itself through various linguistic cues and conversational strategies, such as reticence in answering or asking questions (Layman 2009; Greenspan 2010), deflection or redirection (Donovan-Kicken et al. 2013), or in explicit disagreement. With natural language processing and machine learning techniques, we built a tension detection tool that automatically identifies places in the interview where these tension moments can be detected. One usage scenario of the tool is that the researchers need to answer the questions in the previous paragraph and identify patterns in a large amount of interview data (e.g., over 100 interviews). The tool emerges out of the Living Archives of the Rwandan Diaspora, a Social Sciences and Humanities Research Council of Canada-funded partnership development project between the Centre for Oral History and Digital Storytelling (COHDS) and PAGE-Rwanda, which represents Rwandan genocide survivors living in Montreal. The project’s goal is to produce an online platform (<https://livingarchivesvivantes.org/>) where researchers, community members, and students can listen to, and work with, the testimony of thirty survivors of the 1994 genocide that killed hundreds of thousands of Rwandan Tutsi. To facilitate this listening, the project has developed a suite of tools that enable us to search, map, and listen to survivor testimony in new and diverse ways (Caquard and Dimitrovias 2017). The tension tool, developed by one of the authors of this study as his master’s thesis in Computer Science, came as a result.

The life story interviews, which vary in duration from ninety minutes to twelve hours, were recorded between 2007 and 2012 by the Montreal Life Stories project, another COHDS-based partnership project that recorded 500 life stories of Montrealers displaced by war, genocide, and other human rights violations. These interviews were then integrated into live theatre performances, radio programs, online digital stories, audio walks, art installations, pedagogical units, and a museum exhibition, and 500 Montreal metro cars were equipped with audio portraits that allowed citizens to listen to these stories. A large number of books and articles (High, Little, and Duong 2014; High 2014; High 2015; Miller, Little, and High 2017) have been written about this earlier project, including some preliminary tool development (Xiao, Luo, and High 2013; Jessee, Zembrzycki, and High 2011; High and Sworn 2009). The Living Archives of the Rwandan Diaspora is one of many initiatives that have built on this research foundation since 2012. In this study, our long-term objective is to identify the tensions in Rwandan genocide victims' transcribed and translated interview transcripts. To achieve this objective, we explore computational methods to automatically identify the tension moments in the transcripts. In this paper, we report our tension detection tool. The rest of the paper is structured as follows: Firstly, we review the related literature on detecting tension or similar phenomena in interview transcripts. We also discuss earlier works on detecting hedges and emotions from text, as these are crucial components of our architecture for tension detection. Then we discuss in detail our tension analysis framework and our experimental results on a survivor interview and give a thorough analysis of the system. Lastly, we give a summary of the research work that has been done in this study. We also give direction for further work that can be done in this field.

## 2 Related work

Though interview dynamics have been studied to some degree in the past (Misztal 2003; Layman 2009; Bornat 2010; Thompson 2017; Ponterotto 2018), there is very little work that has been done to automate the process of detecting tension in interviews with computational approaches. Burnap and colleagues performed conversational analysis and used different text mining rules to identify spikes in tension in social media (Burnap et al. 2015). They illustrated how lexicons of abusive or expletive terms can identify high levels of tension separated from low levels. Their proposed tension detection engine relies solely on the lexicons and membership categorization analysis (MCA) (Sacks 1995). They demonstrated that their model has consistently outperformed several machine learning approaches and sentiment analysis tools.

Distress is a negative affective condition that people experience when they feel upset. Distress is closely related to tension. McCubbin and colleagues discussed how stressor

events produce tension and how stress becomes distress when it is subjectively defined as unpleasant (McCubbin, Sussman, and Patterson 2014). Buechel and colleagues considered the problem of distress and empathy prediction as a regression problem (Buechel et al. 2018). They used a Feed-Forward Neural Network with Fast-Text word embeddings as their inputs and a CNN system with one convolutional layer with three different filter sizes. They claim that CNN models can capture semantic effects from the word order and found that such models are especially successful in detecting distress when compared with detecting empathy from text. The researchers provided the first publicly available dataset for text-based distress and empathy prediction.

While these early studies illustrate the possibility of detecting tensions in interviews using machine learning and natural language processing techniques, they failed to fully leverage the indicators of tensions that are identified from the literature. For instance, tension can be shown as reticence in the interview. Layman discussed how reticence can cause the interviewees to shift the conversation, thus restricting the interviewees' responses (Layman 2009). It is a common strategy embraced by the interviewees in order to avoid either complete refusal to reply or full disclosure. Layman also discussed how necessary it is to be conscious of these circumstances so that the interviewer can better judge whether the interviewee should be questioned (Layman 2009). For example, 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. This phenomenon reveals tension points in an interview and gives an idea to interviewees that the conversational stream has been interrupted somewhat. Layman also showed how certain topics can lead interviewees to use such strategies to avoid answers to certain questions (Layman 2009). Most commonly, these answers are reticent and short or dismissive. Subjects that address individual trauma, whether tormenting or frightful or humiliating, will probably trigger hesitant narrator-induced reactions. This leads to the judgement of the interviewers whether the interviewee is to be pressed if it is clear that they are unwilling to speak on certain issues.

Conceptually speaking, tension moments are where the interviewer and interviewee are working at cross-purposes or are not quite on the same page. Usually, this involves moments when the interviewer wants the conversation to go in one direction, but the survivor either doesn't want to go "there" (deflection) or wants to go in another direction (booster). It also includes moments of outright, though often subtle, disagreement (Ahn 2010). Hesitation is also significant in an interview, particularly when the subject being explored is mass violence. From the language use perspective, these moments are expected to have hedging or booster words/phrases. Hedging refers to the technique used to add fuzziness to a speaker's propositional content. According to De Figueiredo-Silva,

hedging can be viewed as a speaker's reserved attitude towards a claim and towards their audience (De Figueiredo-Silva 2001). It can be as simple as saying "maybe," "almost," or "somewhat" in ordinary discourse. It is a common strategy of hesitation embraced by narrators in interviews with oral history. It gives narrators an opportunity to think and organize their thoughts in order to plan safe answers when they are asked difficult questions. For example, the usage of "I think ..." or "Well ..." in interviews gives interviewees the authority to shape their stories. For example, the sentence, "I assume he was involved in it," shows how the usage of the hedge word "assume" can weaken the propositional content "he was involved in it." Phrases such as "In other words" or "In my understanding" can also be used to shift a topic either completely or partially. It can be used as a filler or delaying tactic. This is frequent when there is a disjuncture between the interviewer and the narrator. Often interviewees insist on individualizing their narrative, because they either do not feel authorized to speak for the group, or they have a realization that their story is theirs. Often, as a substitute for hedge words, discourse markers are used during oral history interviews. A discourse marker can be an utterance or a word or a phrase (such as *oh*, *like*, *well*, and *you know*) that either directs or redirects the flow of conversation without adding any significant meaning to the discourse (Schiffrin 1987). Ponterotto demonstrated how hedging in talks is used to tackle controversial issues (Ponterotto 2018). On the other hand, boosting, using terms such as "obviously," "clearly," and "absolutely," is a communicative strategy for expressing firm commitment to statements. It limits the negotiating room for the audience. It plays a vital role in creating conversational solidarity (Holmes 1984) and in constructing an authoritative persona in interviews (Weiyun He 1993). Interestingly, if booster words are preceded by negated words (e.g., *not*, *without*), it can act as hedging (e.g., *not sure*).

Besides the detection of reticence and the use of hedging and/or booster words/phrases, the presence of negative emotions can also be indicators of tension moments. Jurek and colleagues discussed how negative emotion can lead to tension (Jurek, Mulvenna, and Bi 2015). Misztal discussed how emotions lead directly to the past and bring the past somatically and vividly into the present (Misztal 2003). In survivor interviews, interviewees may experience different negative emotions (e.g., anger, sadness, fear, etc.) and feel discomfort. If the interviewer notices this and shifts the topics, the interviewee may come back to the calm state. If the interviewer keeps pushing, however, the interviewee may become too uncomfortable and stop cooperating (e.g., refusing to answer questions). Emotion, therefore, can act as a strong signal of tension in the conversation. In the following examples from our research data, the Rwandan survivor interviews demonstrate the strong negative emotions when words fail us that interviewees may carry in our data contexts. (In all transcript excerpts, the questions

by the interviewer will be indicated by “Interviewer,” and the interviewee’s response by “Narrator.” The interview transcripts can be found at <http://livingarchivesvivantes.org/>. Note: The interviewees gave full consent to use the interview transcripts for research purposes.)

1. **Interviewer:** You’ve felt different emotions because of the events in Rwanda, but are there things that have stayed with you even to this day?

**Narrator:** Yes ... I couldn’t understand how one can commit acts like these, how one can hate and carry out such atrocities against another human being.

2. **Narrator:** I’m not going to waste my time praying in these circumstances because it’s completely—it’s hogwash.

**Interviewer:** Tell me—

**Narrator:** But what is even more serious is that there are Canadians, especially Quebecers, who stand behind the factions and are even more extremist than we are!

**Interviewer:** Indeed.

**Narrator:** It’s strange!

### 3 Tension analysis framework

The two core components of our proposed framework for detecting tension in interview transcripts are: the Emotion Recognition Module and the Hedge Detection Module. In this section, we provide a brief overview of these components. We also discuss other important features (booster words, markers, etc.) that we found useful during our study in this section. At the end of the section, we provide pseudo-code incorporating all of these components.

#### 3.1 Emotion recognition

Emotion plays an important role in recognizing conditions of tension during survivor interviews, as we discussed earlier. To analyze whether and how the interviewee’s emotional aspect indicates the tension during the interview, we developed an emotion recognition tool to recognize the interviewee’s emotions from the interview transcript. There is often a misconception about sentiments and emotions as these subjectivity terms have been used interchangeably (Munezero et al. 2014). Munezero and colleagues differentiate these two terms along with other subjectivity terms and provide the computational linguistics community with clear concepts for effective analysis of text (Munezero et al. 2014). While



sentiment classification tasks (Pang and Lee 2008; Cambria et al. 2017) deal with the polarity (positive, negative, or neutral sentiment) of a given text and the intensity of it, emotion mining tasks usually deal with human emotions, which in some end purposes are more desirable (Ren and Quan 2012; Desmet and Hoste 2013; Mohammad et al. 2015). Leveraging the high performance of deep learning compared to other machine learning approaches (Kim 2014; Kalchbrenner, Grefenstette, and Blunsom 2014; Islam, Mercer, and Xiao 2019), we used a multi-channel convolutional neural network (CNN) model to recognize the emotions from the transcript. Kim showed the effectiveness of a simple CNN model that leverages pre-trained word vectors for a sentence classification task (Kim 2014). Kalchbrenner and colleagues proposed a dynamic CNN model that utilizes a dynamic k-max pooling mechanism (Kalchbrenner, Grefenstette, and Blunsom 2014). Their model is able to generate a feature graph, which captures a variety of word relations. They showed the efficacy of their model by achieving high performances on binary and multi-class sentiment classification tasks without any feature engineering. More recently, Islam and colleagues proposed a multi-channel convolutional neural architecture with the incorporation of different lexical features in the neural network model, which significantly improves the performance of emotion and sentiment identification tasks (Islam, Mercer, and Xiao 2019). In this study, in order to identify emotion of an interviewee from interview transcripts, we utilize the model discussed in Islam, Mercer, and Xiao (Islam, Mercer, and Xiao 2019).

### **3.2 Hedge detection**

Hedging is a widely used conversational management strategy to show the lack of commitment of the speaker to what they say, which can signify conflicts among the speakers. People use hedging when they try to avoid criticism or evade questions in conversations (Crystal 1988). Identifying hedges in conversational text is another core component of our tension analysis framework. Martín discussed four common hedging strategies: Indetermination, Camouflage, Subjectivization, and Depersonalization (Martín 2003). We provide brief details about these strategies motivated by the description found in Alonso Alonso and colleagues (Alonso Alonso, Alonso Alonso, and Torrado Mariñas 2012). Strategy of Indetermination includes the usage of various epistemic modalities, for example, epistemic verbs (*assume, suspect, think*), epistemic adverbs (*presumably, probably, possibly*), epistemic adjectives (*apparent, unsure, probably*), modal verbs (*might, could*) and approximators (*usually, generally*). The use of such epistemic modalities in the interviewee's response creates vagueness and ambiguity. Strategy of Camouflage includes the use of different adverbs (e.g., *generally*

*speaking, actually*). This approach serves as a lexical tool to stop the interviewer from having a negative reaction. Strategy of Subjectivization is activated by the usage of first-person pronouns followed by verbs of cognition, for example, “*I think*” or “*I feel*.” These expressions have been given the term “Shield” in Prince, Frader, and Bosk (Prince, Frader, and Bosk 1982). In certain cases, this approach allows the interviewees to openly express their opinions and hear them. Strategy of Depersonalization includes the use of impersonal pronouns or constructs, for example, “*we*,” “*you*,” or “*people*.” This makes it possible for interviewees to hide behind an unknown subject.

The following two examples from a conversational interview transcript demonstrate the use of hedging for these purposes:

1. **Narrator:** Well, *I think* we have the duty to our children to teach them where they come from.
2. **Narrator:** *I don't know* if I want to talk about my brothers and sisters just yet.

The use of hedge terms “*I think*” and “*I don't know*” demonstrates the instability in their narrative. Besides hedge words, people use discourse markers to hedge in conversations. These can be an utterance or a word or a phrase (such as “*oh*,” “*like*,” “*well*,” and “*you know*”) that either direct or redirect the flow of conversation without adding any significant meaning to the discourse (Schiffrin 1987). For example, “Well, I don't know if there are other things I'd like to share, except that I think that we still have a very, very long journey to go as a nation.”

Our rule-based hedge detection algorithm leverages lexicons we compiled for hedge words, discourse markers, and booster words. We included different epistemic words in our hedge words lexicon that show their hedging act, such as verbs (*suppose, think, presume*), adverbs (*arguably, barely, seemingly*), adjectives (*unlikely, unsure, unclear*), and modal verbs (*might, maybe*). We also included various approximators (such as *generally, usually*) in the lexicon. People also use discourse markers when hedging in conversations. These markers have a variety of functions. For example, when making an unexpected contrast (*even though, despite the fact that*), making a contrast between two separate things, people, ideas, etc. (*anyway, however, rather*), clarifying and re-stating (*in other words, in a sense, I mean*), or to change topic or return to the topic (*well, anyway*). We also compiled a list of such discourse markers. In order to measure the comparability between the discourse markers of our lexicon and the phrases from the input sentences, we used Jaccard distance, complementary to the Jaccard index. We have built a lexicon for boosting words as well. Boosting, using terms such as *absolutely, clearly, and obviously*, is a communicative strategy for expressing a firm commitment

to statements. Interestingly, if booster words are preceded by negation words such as “not” or “without,” they can act as hedges. For example, “I’m still not sure if I would go back; I don’t know what it would be like.” Here, “sure” is a booster word. However, since it is preceded by a negation word “not,” it changes the meaning completely. We handle this kind of situation in our hedge detection algorithm.

Hedging disambiguation is an important part of our algorithm, given that some commonly used hedge terms also have non-hedge senses in conversational interviews. We apply rules to disambiguate these terms based on the syntactic structure of the sentences. Islam and colleagues discussed several hedge disambiguation rules which we used in this study (Islam, Mercer, and Xiao 2020). We used the Stanford CoreNLP (Manning et al. 2014) parser to parse the sentences (<https://stanfordnlp.github.io/CoreNLP/download.html>). One of the main reasons we chose this rule-based approach over a learning-based approach was that there is no large enough benchmark annotated dataset available in this genre that could have been leveraged to build a good classifier. A statistical model can be very useful in discovery of latent relations between features, which is difficult with a rule-based approach. However, with this study, we mark the start of producing an annotated dataset supervised by the experts in this field that can be utilized in future research. The following is a brief review of a subset of the rules used in our research with examples from our interview datasets.

**Hedge Term:** Feel, Suggest, Believe, Consider, Doubt, Guess, Hope

*Rule:* If token  $t$  is (i) a root word, (ii) has the part-of-speech VB\*, and (iii) has an *nsubj* (nominal subject) dependency with the dependent token being a first person pronoun (*i, we*),  $t$  is a hedge, otherwise, it is a non-hedge.

*Hedge:* I **hope** to, someday, but no, I haven’t reached it yet.

*Non-hedge:* A message of **hope** and daring to shed light on everything we see.

**Hedge Term:** Think

*Rule:* If token  $t$  is followed by a token with part-of-speech IN,  $t$  is a non-hedge, otherwise, hedge.

*Hedge:* I **think** it’s a little odd.

*Non-hedge:* I **think** about this all the time.

**Hedge Term:** Assume

*Rule:* If token  $t$  has a *ccomp* (clausal complement) dependent,  $t$  is a hedge, otherwise, non-hedge.

*Hedge:* I **assume** he was involved in it.

*Non-hedge:* He wants to **assume** the role of a counsellor.

### 3.3 Tension detection

In addition to the two modules we discussed above, our proposed tension analysis framework makes use of a few additional features that proved to be important during our research. We provide brief details about these features along with the pseudo-code of our proposed algorithm below.

#### 3.3.1 Markers

It is interesting that markers (e.g., laughter, silence, sigh) are used in these interview transcripts. These have various functions. Sometimes markers like “laughter” indicate invitations to the interviewer to ask the next question. At other times they represent hesitation or nervous deflection (i.e., the tension). In this work, we have compiled a list of such markers/cues but acknowledge that further exploration is needed to interpret these cues. The example below shows a use of the marker “laugh”:

**Interviewer:** And what would you like Rwandans, your community, to know about you and that maybe we don’t already know, maybe we ... ? If it were necessary ...

**Narrator:** [laughs] ... I don’t know. It’s a difficult question.... I don’t know since ... I think that all Rwandans, well, every Rwandan has his or her own experience, and I’m not sure that I should be asking them to think about me in a certain way.

#### 3.3.2 Asking questions back

When the interviewed person asks for clarification, posing a question back, then it is also a symbol, a good marker for recognizing tension points. During our research, we have found that asking a question back to the interviewer may possibly be a sign that the interviewee is trying to negotiate. We use this as a possible criterion in our tension detection algorithm. The following example illustrates such a situation:

**Interviewer:** So going back now to the period of 1994, during the genocide—you saw it coming, but how did you live through that time?

**Narrator:** How do you mean?

#### 3.3.3 Outliers

In cases where an interviewee gives unusually long or short answers to a particular question form, shorter or longer than three standard deviations from the average length of responses of that sort (for example, wh questions, yes/no questions, etc.), that is an important indicator of some sort of change in interview dynamics. During our

group discussions, we felt that this type of dynamic could be a sign of tension, so we added this as one of the criteria in our tension detection algorithm. We find the mean (Equation 1) and standard deviation (Equation 2) for each question type. (In this study, we considered wh-question, how, yes-no, and mixed [mix of several question types] as the prime question types.)

$$\mu_{q_t} \leftarrow \frac{1}{N(q_t)} \sum_{i=1}^{N(q_t)} \{w_i(q_t)\} \quad (1)$$

$$\sigma_{q_t} \leftarrow \sqrt{\frac{\sum_{i=1}^{N(q_t)} (w_i(q_t) - \mu(q_t))^2}{N(q_t) - 1}} \quad (2)$$

Here,  $\mu_{q_t}$  indicates the mean for the question type  $q_t$ ,  $\sigma_{q_t}$  indicates the standard deviation for the question type  $q_t$ ,  $w_i(q_t)$  indicates the total number of words in excerpt  $e_i$  belonging to  $q_t$ , and  $N(q_t)$  indicates the total number of excerpts belonging to each  $q_t$ . We consider a response to be an outlier, thus a possible point for tension, if it falls below  $3\sigma_{q_t}$  or is above  $3\sigma_{q_t}$ .

#### 3.3.4 Algorithm

Here, we provide the pseudo-code for our tension detection algorithm. Our algorithm detects tension on excerpt level considering different factors (emotions, hedging, markers etc.) that are present in the sentences of an excerpt. If an excerpt is not labelled as having tension by our algorithm, this indicates that the algorithm did not find any tension-causing phenomena that we discussed earlier in this article, though we acknowledge there can be a few cases where the algorithm fails due to the constraints posed by transcribed texts versus the actual video interview.

#### Algorithm Tension Detection algorithm

```

1: function TensionDetection()
2:   Excerpts(E) // List of narrator's responses
3:   Markers(M) // List of evasions markers and cues
4:   Single Excerpt(e) // List of sentences in each response
5:   qt // Question type (wh-question/how/yes-no/mixed)
6:   wi(qt) // Total number of words in excerpt ei belonging to qt
7:   N(qt) // Total number of excerpts belonging to each qt
8:   Mean,  $\mu_{qt}$  // refer to (1)
9:   Standard deviation,  $\sigma_{qt}$   $\leftarrow$  refer to (2)
10:  for each excerpt e in E do
11:    w // Total number of words in excerpt e

```

```

12:   q // Question asked by the Interviewer
13:   nSentences // First n sentences in e
14:   isNegativeEmotion, eneg = False
15:   isHedgedSentence, hs = False
16:   isBoosting, bs = False
17:   markerPresent, mp = False
18:   isQuestion, qs = False
19:   isOutlier, or = False
20:   for each sentence s in nSentences do
21:     if isNegativeEmotion(s) is True then
22:       eneg = True
23:     end if
24:     if isHedgedSentence(s) is True then
25:       hs = True
26:     end if
27:     if isBoosting(s) is True then
28:       bs = True
29:     end if
30:   end for
31:   for Marker/Cue (A) in M do
32:     if A in e then
33:       mp = True
34:     end if
35:   end for
36:   if nSentences[0] is a Question then
37:     qs = True
38:   end if
39:   if  $w > \mu_{qt} + 3 * \sigma_{qt}$  or  $w < \mu_{qt} - 3 * \sigma_{qt}$  then
40:     or = True
41:   end if
42:   if (eneg and hs) or (hs and bs) or (hs and mp) or qs or or then
43:     mark excerpt as Tension
44:   else
45:     mark excerpt as No Tension
46:   end if
47: end for
48: end function

```

## 4 Evaluation of the tension detection tool

In this section, we provide details about the experiments to examine how well our computational approach performs in comparison to the human performance when it is applied to an annotated interview transcript to identify the tensions in the interviewee's responses. Then, we compared the results with the analysis performed by student researchers. There is considerable messiness in manual annotations as researchers identified varying points in the interview. We then identified those places where the majority of the student researchers identified tension and then compared these to the computational results. We believe that this real-life comparison has considerable merit. It also opened up a space in the oral history classroom to discuss these issues. It is a valuable pedagogical exercise in its own right.

### 4.1 Interview transcripts and annotation process

We used 15 interview transcripts in this evaluation process. They were obtained from the interview collection of the Living Archives of Rwandan exiles and genocide survivors in Canada. This digital repository contains life stories of Rwandan genocide survivors. The 15 interviews lasted from 55 minutes to 184 minutes with the average of 128.7 minutes.

The transcripts were annotated by a group of students taking a public history course with a focus on the Living Archives of the Rwandan Diaspora (<http://livingarchivesvivantes.org/>). These students had been watching interviews each week, working with the transcripts, learning about oral history and mass violence. They had been demonstrated interviewer-interviewee dynamic, which is at the heart of the conversational narrative of the oral history interviews. Then, the students were paired to work together to annotate the whole transcript of one interview following the instructions of the instructor who is a co-author of this paper. Specifically, they annotated four types of incidences in the interviewee's responses: the points of tension (T), the interviewee's hesitation (H), the deflection in the response (D), and the interviewee's boosting (B). Tension is often used as an umbrella term for when the interviewer and interviewee work at cross-purposes, whereas hesitation and deflection have more specific meaning. Deflection can also be followed by boosting. Interviewees tend to use boosting when they try to drag the interviewer somewhere in the interview. We also acknowledge the fact that transcribed and translated interviews are not the same as the recorded ones. It's an "echo of an echo." What sounded abrupt on the transcribed interview might be perfectly normal (and tension free) in the video. Similarly, what sounded normal on the transcribed texts might contain tension in the actual video interview as tension might be present in facial and vocal expressions that might not be captured in text.

There were 15 teams in total, each team having 2 students. Team members discussed with each other first, and the annotated results reflect the team's shared interpretation of the categories and the interviewee responses. In total, there were 116 interviewee responses that have been annotated by these teams that have been used in this study for the purpose of evaluating our algorithm. Of the four categories, the point of tension (T) was annotated the most with 15 responses identified as such by Team #3, 11 by Team #13, 12 by Team #14, and 10 by Team #15. The second most annotated category was Boosting (B) with 14 responses identified as such by Team #13, 7 by Team #7, and 6 by Team #8, Team #11, and Team #14.

While the students have had the same level of training and familiarity with the interview content and the interview context, the interpretation of the four categories is so subjective that the teams had different annotations. For example, the highest level of agreement among the teams is that 6 teams agreed on the same annotation of one response, and there are only two such situations in the annotation. This finding illustrates the challenge of conducting tension analysis in this interview context. **Table 1** shows the total number of responses annotated in each category by the teams.

#### 4.2 The comparison of the annotations by the teams vs. our tool

With this transcript, we first segmented the text according to the turns by the interviewer and the interviewee. We applied the tension detection tool to the segmented data and classified each interviewee response as tension or no tension. In total, our tool identified 55 tension points out of 116 interviewee responses.

We compared the performance of our tool with that of the student teams' annotations through three aspects. First, we examined whether the tool would be able to identify all the possible tension points by a researcher. To do so, we considered an interviewee response to be a human annotated tension point if any team marked it as T. Our tool

	Team														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Tension (T)	1	5	15	9	7	8	5	3	6	7	8	1	11	12	10
Hesitation (H)	3	2	6	3	6	4	4	1	3	1	6	1	5	3	4
Boosting (B)	1	2	5	4	4	2	7	6	5	1	6	0	14	6	5
Deflection (D)	4	5	3	9	4	2	5	3	2	1	4	0	9	4	6
No entry	107	102	87	91	95	100	95	97	100	106	92	114	77	91	91

**Table 1:** Total number of responses annotated in each category by the teams.



was able to identify 37 out of 47 annotations humanly labelled as T. It also incorrectly marked 18 as T.

Acknowledging that our model considers hedging, boosting, and deflection as indicators of tension points, we considered another aspect in the comparison—an interviewee response is marked as a point of tension by human annotators if any team has annotated any category on it. From this aspect, the tool was able to identify 16 of 28 hedging annotations, 21 of 34 boosting annotations, and 7 of 32 deflection annotations. It also incorrectly marked annotations in each of these categories: 32, 25, and 12, respectively.

In the third aspect of the evaluation, we utilized a voting system to determine the final annotation of a response by an interviewee. First, we compiled all four categories used by the teams into one single category representing a tension point (T). Next, we assigned a label for each response of the interviewee when at least 8 teams agreed on the label out of the 15 participated teams. Our tool was able to identify all the 4 annotations classified as T. It also incorrectly classified 51 annotations as T.

## 5 Discussion

As a filtering device that facilitates as opposed to replaces the researcher's qualitative analysis process of interview data, this tool has a promising result. Specifically, our evaluation results show that, overall, the tool is able to identify the majority of the interview places that were annotated as containing tensions or indicators of tensions. However, as shown in the above section, there is room for improvement, mainly to decrease the number of cases where the tool labels as tension points but the human experts do not. Language techniques can be explored to improve this performance. For example, one of the problems that must be tackled by any description of discourse markers is their poly-functionality, which means it is very important to distinguish the usage of different markers. Although we tried to disambiguate a number of hedge terms in this work, we need a clearer understanding of certain discourse markers. One of the problems with our strategy is its failure to discern the discourse functions of the marker “*well*.” “*Well*” has various functions. It has been well investigated by many scholars over the years (Ponterotto 2018; Jucker 1993). It appears in seemingly different contexts. According to Jucker, “*well*” can be used as a marker of insufficiency, as a face-threat mitigator, as a frame, or as a delay device (Jucker 1993).

Tensions can build up over time during conversations, and various factors can contribute to that, such as the interviewer's questions, the topics covered right before this response, etc. Our current framework has only considered the interviewee response. We will explore the potential of these contextual factors in detecting tensions in the conversations.

As mentioned in the introduction section, people can keep their tensions internally without letting the other conversation partners notice them. Our work is aimed at detecting those that have external markers in the communication. The external markers can exist in various communication channels—the conversation content; the body language (such as hand movement and facial expression); and the voice and the sound (such as the tone and the pitch). This study has mainly examined the markers in the conversation content with a few more about the voice and sound (e.g., the laughter and the silence). Prior study has shown that prosodic features can be indicative of tensions in interviews (Zhang and Xiao 2020). One of our next steps is to integrate the audio and video recordings of the interview data into the tension detection model.

The last limitation we recognize is the use of Twitter data for training the emotion recognition tool. The interviews were conducted in a conversational style, which offers similarity to the free form of tweets in that sense. On the other hand, interviewees' responses were often much longer than a tweet, and the interview context being about mass violence is very different from day-to-day tweets. These differences between the training data and testing data also put a constraint on the performance of emotion recognition in our study.

Besides survivor interviews, we anticipate that tension between the interviewer and the interviewee may occur in many interviews about sensitive topics. “Sensitive topics” are topics that require participants to reveal their deep personal feelings and/or experiences that are emotionally difficult or stressful for them (Cowles 1988; Johnson and Clarke 2003; Lee 1993), for example, domestic violence, child maltreatment, and sexual behaviour. Unstructured or semi-structured interviews are common in sensitive topic research. Therefore, our work of analyzing tensions in the survivor interviews is expected to contribute to a larger research community that studies sensitive topics through the interview approach (Miller, Little, and High 2017). Our tool is openly accessible at <https://github.com/jumayelo6/Tension-Analysis>. We encourage other scholars in Digital Humanities to conduct tension analysis in their interview projects and further improve the tool.

## 6 Conclusion

Oral history has a pivotal role to play in educating individuals and communities about the social preconditions, experiences, and long-term repercussions of mass violence. Among other things, life story interviews offer us “unique glimpses into the lived interior” of survivors (Thomson 1999, p. 26). Despite its propensity to archive, oral history still privileges fieldwork over secondary analysis. Researchers have been so focused on the making of the interview that we have spent insufficient time thinking about what to do with

the audio or video recordings and transcripts that result. A central strength of qualitative research is “its capacity to furnish contextual detail and to enhance understanding of the salience of contextual diversity in lived experience” (Irwin and Winterton 2012, p. 4; see also Moore 2007; Mason 2007; Corti, Witzel, and Bishop 2005). New digital tools and techniques are therefore needed. To start, we must go beyond what Savage calls the “juicy quotes syndrome,” to engage with interviews in deeper and more holistic ways (Savage 2005). Our tension tool does that, allowing us to research the interview dynamic that is at the very heart of the interview. We agree with Mayernik, who has argued that: “Digital research data, if curated and made broadly available, promise to enable researchers to ask new kinds of questions and use new kinds of analytical methods in the study of critical scientific and societal issues” (Mayernik et al. 2012).

In this work, we explored interview dynamics and how various factors influence this phenomenon. We also talked about survivor interviews and why analyzing the responses of a narrator to identify situations of tension is important. We provided details about our tension analysis architecture and discussed the components of it. We utilized a multi-channel convolutional neural network model, which was trained on social media data, to identify emotions from our transcribed interview data, which is a core component of our framework. We have observed how emotion fluctuates throughout survivor interviews, and negative emotion appears to be the source of a stress situation most of the time. Next, we presented a discussion about hedging and boosting in speakers’ narratives. These phenomena are crucial in tension detection studies and demonstrate the mood of an interviewee during a conversation. We utilize three manually constructed lexicons of hedge words, discourse markers, and booster words and apply predefined rules to disambiguate hedge terms based on the syntactic structure of the sentences. Our framework also takes length of interviewees’ responses and various markers used in such interviews into consideration. We discussed our process of integrating all the discussed components and features by providing an algorithm for detecting tension in oral history interviews.

Our proposed algorithm gives a very good recall score on the transcript that we worked on, and because of its high recall score, it can be used as a filtering tool, which can be of assistance to researchers in this area. Since very little work has been done in this research field, we hope that in the potential continuation of this study, our research findings can be beneficial. It is crucial to have a good understanding of tension phenomenon in order to better analyze such data. Domain experts at Concordia University’s Centre for Oral History and Digital Storytelling are going to perform further analysis on the interview transcripts and will provide us with more insights about the dynamics of such interviews. They are also in the process of annotating more interview

data, which will help us evaluating our model even better in the future. We also plan to have our interview data annotated with different emotion categories and train our emotion recognition model with data from the same domain, which might potentially improve the performance of the model. Another future direction is to identify and integrate the tension markers from various communication channels in the tension detection framework, which has mostly considered the communication content and ignored other places such as the audio and video recordings of the interviews.

Our work of analyzing tensions in the survivor interviews sheds light on the analysis of interviews and conversations that are expected to have tensions (e.g., interviews about sensitive topics). We make our tension detection tool open source, encouraging scholars to apply the tension analysis in their interview research and/or to further improve the tool.

---

## **Competing interests**

The authors have no competing interests to declare.

## **Contributions**

### ***Authorial***

Authorship is alphabetical after the drafting author and principal technical lead. Author contributions, described using the CASRAI CredIT typology, are as follows:

#### **Author name and initials:**

Jumayel Islam (JI)

Robert E. Mercer (RM)

Lu Xiao (LX)

Steven High (SH)

Authors are listed in descending order by significance of contribution. The corresponding author is JI

Conceptualization: JI, RM, LX, SH

Methodology: JI, RM, LX

Software: JI

Formal Analysis: JI, RM, LX, SH

Data Curation: SH, JI

Writing – Original Draft Preparation: JI, RM, LX, SH

Writing – Review & Editing: JI, RM, LX, SH

### ***Editorial***

#### **Recommending Editor**

Barbara Bordalejo, University of Lethbridge, Canada

#### **Section Editor**

Morgan Pearce, The Journal Incubator, University of Lethbridge, Canada

#### **Copy Editors**

Akm Iftekhar Khalid, The Journal Incubator, University of Lethbridge, Canada

Christa Avram, The Journal Incubator, University of Lethbridge, Canada

#### **Layout Editor**

Christa Avram, The Journal Incubator, University of Lethbridge, Canada

#### **Production Consultant**

Virgil Grandfield, The Journal Incubator, University of Lethbridge, Canada

## References

- Ahn, John J. 2010. *Exile as Forced Migrations: A Sociological, Literary, and Theological Approach on the Displacement and Resettlement of the Southern Kingdom of Judah*. Berlin: Walter de Gruyter Inc. DOI: <https://doi.org/10.1515/9783110240962>.
- Alonso Alonso, Rosa, María Alonso Alonso, and Laura Torrado Mariñas. 2012. "Hedging: An Exploratory Study of Pragmatic Transfer in Nonnative English Readers' Rhetorical Preferences." *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos* 23: 47–64. Accessed August 19, 2022. <http://revistaiberica.org/index.php/iberica/article/view/309>.
- Bornat, Joanna. 2010. "Remembering and Reworking Emotions: The Reanalysis of Emotion in an Interview." *Oral History* 38(2): 43–52. Accessed August 25, 2022. <http://oro.open.ac.uk/39509/>.
- Buechel, Sven, Anneke Buffone, Barry Slaff, Lyle Ungar, and Joao Sedoc. 2018. "Modelling Empathy and Distress in Reaction to News Stories." arXiv:1808.10399. DOI: <https://doi.org/10.48550/arXiv.1808.10399>.
- Burnap, Pete, Omer F. Rana, Nick Avis, Matthew Williams, William Housley, Adam Edwards, Jeffrey Morgan, and Luke Sloan. 2015. "Detecting Tension in Online Communities with Computational Twitter Analysis." *Technological Forecasting and Social Change* 95: 96–108. DOI: <https://doi.org/10.1016/j.techfore.2013.04.013>.
- Cambria, Erik, Dipankar Das, Sivaji Bandyopadhyay, and Antonio Feraco. 2017. "Affective Computing and Sentiment Analysis." In *A Practical Guide to Sentiment Analysis*, edited by Erik Cambria, Dipankar Das, Sivaji Bandyopadhyay, and Antonio Feraco, 1–10. Cham: Springer. DOI: [https://doi.org/10.1007/978-3-319-55394-8\\_1](https://doi.org/10.1007/978-3-319-55394-8_1).
- Caquard, Sébastien, and Stefanie Dimitrovas. 2017. "Story Maps & Co. The State of the Art of Online Narrative Cartography." *Mappemonde. Revue trimestrielle sur l'image géographique et les formes du territoire* 121. DOI: <https://doi.org/10.4000/mappemonde.3386>.
- Corti, Louise, Andreas Witzel, and Libby Bishop. 2005. "On the Potentials and Problems of Secondary Analysis: An Introduction to the FQS Special Issue on Secondary Analysis of Qualitative Data." *Forum Qualitative Sozialforschung* 6(1). DOI: <https://doi.org/10.17169/fqs-6.1.498>.
- Cowles, Kathleen V. 1988. "Issues in Qualitative Research on Sensitive Topics." *Western Journal of Nursing Research* 10(2): 163–179. DOI: <https://doi.org/10.1177/019394598801000205>.
- Crystal, David. 1988. "On Keeping One's Hedges in Order." *English Today* 4(3): 46–47. DOI: <https://doi.org/10.1017/S0266078400003540>.
- Davidson, James West, and Mark H. Lytle. 2004. *After the Fact: The Art of Historical Detection*. Vol. 1. New York: McGraw-Hill.
- De Figueiredo-Silva, Maria Isabel Réfega. 2001. "Teaching Academic Reading: Some Initial Findings from a Session on Hedging." In *Proceedings of the Postgraduate Conference 2001 - Department of Theoretical and Applied Linguistics, The University of Edinburgh*, 1–13. Accessed July 25, 2022. <http://www.lel.ed.ac.uk/~pgc/archive/2001/Isabel-Figueiredo-Silva01.pdf>.
- Desmet, Bart, and Véronique Hoste. 2013. "Emotion Detection in Suicide Notes." *Expert Systems with Applications* 40(16): 6351–6358. DOI: <https://doi.org/10.1016/j.eswa.2013.05.050>.
- Donovan-Kicken, Erin, Trey D. Guinn, Lynsey Kluever Romo, and Lea D. L. Ciceraro. 2013. "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. DOI: <https://doi.org/10.1177/0093650211422537>.
- Greenspan, Henry. 2010. *On Listening to Holocaust Survivors: Beyond Testimony*. St. Paul: Paragon House.
- Grele, Ronald J., and Studs Terkel. 1991. *Envelopes of Sound: The Art of Oral History*. New York: Greenwood Publishing Group.
- High, Steven. 2014. *Oral History at the Crossroads: Sharing Life Stories of Survival and Displacement*. Vancouver: UBC Press.
- . 2015. *Beyond Testimony and Trauma: Oral History in the Aftermath of Mass Violence*. Vancouver: UBC Press.
- High, Steven, and David Sworn. 2009. "After the Interview: The Interpretive Challenges of Oral History Video Indexing." *Digital Studies/Le Champ Numérique* 1(2). DOI: <https://doi.org/10.16995/dscn.110>.
- High, Steven, Edward Little, and Thi Ry Duong. 2014. *Remembering Mass Violence: Oral History, New Media and Performance*. Toronto: University of Toronto Press. DOI: <https://doi.org/10.3138/9781442666580>.
- Holmes, Janet. 1984. "Modifying Illocutionary Force." *Journal of Pragmatics* 8(3): 345–365. DOI: [https://doi.org/10.1016/0378-2166\(84\)90028-6](https://doi.org/10.1016/0378-2166(84)90028-6).
- Irwin, Sarah, and Mandy Winterton. 2012. "Qualitative Secondary Analysis and Social Explanation." *Sociological Research Online* 17(2): 1–12. DOI: <https://doi.org/10.5153/sro.2626>.
- Islam, Jumayel, Robert E. Mercer, and Lu Xiao. 2019. "Multi-Channel Convolutional Neural Network for Twitter Emotion and Sentiment Recognition." In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Vol. 1, edited by Jill Burstein, Christy Doran, and Tamar Solorio, 1355–1365. Minneapolis: Association for Computational Linguistics. DOI: <https://doi.org/10.18653/v1/N19-1137>.
- . 2020. "A Lexicon-Based Approach for Detecting Hedges in Informal Text." In *Proceedings of the 12th Language Resources and Evaluation Conference*, edited by Nicoletta Calzolari, Frédéric Béchet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, et al., 3109–3113. Marseille: European Language Resources Association. Accessed August 11, 2022. <https://aclanthology.org/2020.lrec-1.380/>.
- Jessee, Erin, Stacey Zembrzycki, and Steven High. 2011. "Stories Matter: Conceptual Challenges in the Development of Oral History Database Building Software." *Qualitative Social Research* 12(1). DOI: <https://doi.org/10.17169/fqs-12.1.1465>.
- Johnson, Barbara, and Jill Macleod Clarke. 2003. "Collecting Sensitive Data: The Impact on Researchers." *Qualitative Health Research* 13(3): 421–434. DOI: <https://doi.org/10.1177/1049732302250340>.
- Jucker, Andreas H. 1993. "The Discourse Marker Well: A Relevance-Theoretical Account." *Journal of Pragmatics* 19(5): 435–452. DOI: [https://doi.org/10.1016/0378-2166\(93\)90004-9](https://doi.org/10.1016/0378-2166(93)90004-9).
- Jurek, Anna, Maurice D. Mulvenna, and Yaxin Bi. 2015. "Improved Lexicon-Based Sentiment Analysis for Social Media Analytics." *Security Informatics* 4(1): 9. DOI: <https://doi.org/10.1186/s13388-015-0024-x>.

- Kalchbrenner, Nal, Edward Grefenstette, and Phil Blunsom. 2014. "A Convolutional Neural Network for Modelling Sentences." In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, Vol. 1, edited by Kristina Toutanova and Hua Wu, 655–665. DOI: <https://doi.org/10.48550/arXiv.1404.2188>.
- Kim, Yoon. 2014. "Convolutional Neural Networks for Sentence Classification." In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, edited by Alessandro Moschitti, Bo Pang, and Walter Daelemans, 1746–1751. DOI: <https://doi.org/10.48550/arXiv.1408.5882>.
- Koro-Ljungberg, Mirka. 2008. "A Social Constructionist Framing of the Research Interview." *Handbook of Constructionist Research*, edited by James A. Holstein and Jaber F. Gubrium, 429–444. New York: The Guilford Press.
- Layman, Lenore. 2009. "Reticence in Oral History Interviews." *The Oral History Review* 36(2): 207–230. DOI: <https://doi.org/10.1093/ohr/ohp076>.
- Lee, Raymond M. 1993. *Doing Research on Sensitive Topics*. Thousand Oaks, CA: SAGE Publications.
- Manning, Christopher D., Mihai Surdeanu, John Bauer, Jenny Rose Finkel, Steven Bethard, and David McClosky. 2014. "The Stanford CoreNLP Natural Language Processing Toolkit." In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, edited by Kalina Bontcheva and Jingbo Zhu, 55–60. Baltimore: Association for Computational Linguistics. DOI: <https://doi.org/10.3115/v1/P14-5010>.
- Martín, Pedro. 2003. "The Pragmatic Rhetorical Strategy of Hedging in Academic Writing." *Vigo International Journal of Applied Linguistics*: 57–72. DOI: <https://doi.org/10.35869/vial.v0i0.3867>.
- Mason, Jennifer. 2007. "Re-using Qualitative Data: On the Merits of an Investigative Epistemology." *Sociological Research Online* 12(3): 1–4. DOI: <https://doi.org/10.5153/sro.1507>.
- Mayernik, Matthew S., G. Sayeed Choudhury, Tim DiLauro, Elliot Metsger, Barbara Pralle, Mike Rippin, and Ruth Duerr. 2012. "The Data Conservancy Instance: Infrastructure and Organizational Services for Research Data Curation." *D-Lib Magazine* 18(9/10). DOI: <https://doi.org/10.1045/september2012-mayernik>.
- McCubbin, Hamilton I., Marvin B. Sussman, and Joan M. Patterson, eds. 2014. *Social Stress and the Family: Advances and Developments in Family Stress Therapy and Research*. New York: Routledge. DOI: <https://doi.org/10.4324/9781315804439>.
- Miller, Elizabeth, Edward Little, and Steven High. 2017. *Going Public: The Art of Participatory Practice*. Vancouver: UBC Press.
- Misztal, Barbara A. 2003. *Theories of Social Remembering*. Berkshire: Open University Press.
- Mohammad, Saif M., Xiaodan Zhu, Svetlana Kiritchenko, and Joel Martin. 2015. "Sentiment, Emotion, Purpose, and Style in Electoral Tweets." *Information Processing & Management* 51(4): 480–499. DOI: <https://doi.org/10.1016/j.ipm.2014.09.003>.
- Moore, Niamh. 2007. "(Re)Using Qualitative Data?" *Sociological Research Online* 12(3): 1–13. DOI: <https://doi.org/10.5153/sro.1496>.
- Munezero, Myriam, Calkin Suero Montero, Erkki Sutinen, and John Pajunen. 2014. "Are They Different? Affect, Feeling, Emotion, Sentiment, and Opinion Detection in Text." *IEEE Transactions on Affective Computing* 5(2): 101–111. Piscataway, NJ: IEEE. DOI: <https://doi.org/10.1109/TAFFC.2014.2317187>.



- Pang, Bo, and Lillian Lee. 2008. "Opinion Mining and Sentiment Analysis." *Foundations and Trends in Information Retrieval* 2(1-2): 1-135. DOI: <https://doi.org/10.1561/1500000011>.
- Ponterotto, Diane. 2018. "Hedging in Political Interviewing: When Obama Meets the Press." *Pragmatics and Society* 9(2): 175-207. DOI: <https://doi.org/10.1075/ps.15030.pon>.
- Portelli, Alessandro. 1991. *The Death of Luigi Trastulli and Other Stories: Form and Meaning in Oral History*. New York: SUNY Press.
- Prince, Ellen F., Joel Frader, and Charles Bosk. 1982. "On Hedging in Physician-Physician Discourse." In *Linguistics and the Professions: Proceedings of the Second Annual Delaware Symposium on Language Studies*, edited by Robert J. Di Pietro, 83-97. Norwood: ABLEX Publishing Corporation.
- Ren, Fuji, and Changqin Quan. 2012. "Linguistic-Based Emotion Analysis and Recognition for Measuring Consumer Satisfaction: An Application of Affective Computing." *Information Technology and Management* 13(4): 321-332. DOI: <https://doi.org/10.1007/s10799-012-0138-5>.
- Sacks, Harvey. 1995. *Lectures on Conversation*. Edited by Gail Jefferson and Emanuel A. Schegloff. Hoboken, NJ: Wiley-Blackwell.
- Savage, Mike. 2005. "Revisiting Classic Qualitative Studies." *Historical Social Research* 6(1): 118-139. <https://doi.org/10.17169/fqs-6.1.502>.
- Schiffrin, Deborah. 1987. *Discourse Markers*. New York: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511611841>.
- Tanggaard, Lene. 2009. "The Research Interview as a Dialogical Context for the Production of Social Life and Personal Narratives." *Qualitative Inquiry* 15(9): 1498-1515. DOI: <https://doi.org/10.1177/1077800409343063>.
- Thompson, Paul. 2017. *The Voice of the Past: Oral History*. Oxford: Oxford University Press.
- Thomson, Alistair. 1999. "Moving Stories: Oral History and Migration Studies." *Oral History* 27(1): 24-37. Accessed August 25, 2022. <https://www.jstor.org/stable/40179591>.
- Tripp, David H. 1983. "Co-authorship and Negotiation: The Interview as Act of Creation." *Interchange* 14(3): 32-45. DOI: <https://doi.org/10.1007/BF01810469>.
- Weiyun He, Agnes. 1993. "Exploring Modality in Institutional Interactions: Cases from Academic Counselling Encounters." *Text-Interdisciplinary Journal for the Study of Discourse* 13(4): 503-528. DOI: <https://doi.org/10.1515/text.1.1993.13.4.503>.
- Xiao, Lu, Yan Luo, and Steven High. 2013. "CKM: A Shared Visual Analytical Tool for Large-Scale Analysis of Audio-Video Interviews." In *2013 IEEE International Conference on Big Data* edited by Xiaohua Hu, Tsau Young Lin, Vijay Raghavan, Benjamin Wah, Ricardo Baeza-Yates, Geoffrey Fox, Cyrus Shahabi, et al., 104-112. Piscataway, NJ: IEEE. DOI: <https://doi.org/10.1109/BigData.2013.6691677>.
- Zhang, Bo, and Lu Xiao. 2020. "Augmented Tension Detection in Communication: Insights from Prosodic and Content Features." In *Human-Computer Interaction: Multimodal and Natural Interaction*, edited by Masaaki Kurosu, 290-301. Cham: Springer. DOI: [https://doi.org/10.1007/978-3-030-49062-1\\_20](https://doi.org/10.1007/978-3-030-49062-1_20).

