Discourse processing of oral history

Ryan Shaw
ryanshaw@unc.edu
https://aeshin.org/
Contours of the Past

- Project to invent new tools for exploring oral histories using computational discourse processing techniques
- Funded by an Early Career Development grant from the Institute of Museum and Library Services
“… marking up a text with structural descriptions on several levels …”

Topic-based segmentation

Coreference resolution

http://doi.org/10.2200/S00354ED1V01Y201111HLT015
### Object Description

<table>
<thead>
<tr>
<th>Interview no.</th>
<th>U-0017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restrictions</td>
<td>No restrictions. Open to research.</td>
</tr>
<tr>
<td>Project</td>
<td>U.2. Long Civil Rights Movement: School Desegregation in Robeson County, N.C.</td>
</tr>
<tr>
<td>Date</td>
<td>21 October 2003</td>
</tr>
<tr>
<td>Interviewee</td>
<td>Thompson, Angus, 1919-</td>
</tr>
<tr>
<td>Interviewee occupation</td>
<td>Tailors, Farmers, Construction workers</td>
</tr>
<tr>
<td>Interviewee DOB</td>
<td>1919</td>
</tr>
<tr>
<td>Interviewee ethnicity</td>
<td>African American/Black</td>
</tr>
<tr>
<td>Interviewer</td>
<td>Lowery, Malinda Maynor.</td>
</tr>
<tr>
<td>Abstract</td>
<td>Angus Thompson recalls decades of civil rights activism, from pushing school integration to opposing segregated public facilities. Thompson inherited a legacy of activism from his father and became a forceful leader in the African American community in Lumberton, N.C., forging voting alliances with local Indians and opposing other black politicians'; accommodationist impulses. Thompson's story is one of undiluted support for integration, which he sees as the cornerstone of racial progress in the second half of the twentieth century. This interview will prove useful for researchers looking for on-the-ground narratives of civil rights activism and an impassioned defense of the progress of the past fifty years.</td>
</tr>
</tbody>
</table>
Transcripts are great, but…

- Names mentioned only in passing
- Misspeaking
- Oblique references
- Transcription errors
- Replaces listening with reading
Finer-grained access

- segments 5–15 minutes in length, with boundaries at "a natural ending of a topic, a break before a new question, or other natural pause in the flow of speech."

- “unify the scale of the navigable unit”

Lambert & Frisch, “Meaningful access to audio and video passages,” http://ohda.matrix.msu.edu/?p=1817
Manual segmentation
“Oral Histories of the American South”

Selected segments, with brief description, notes and descriptive classification: 11

- [00:10:48 - 00:21:33] NAACP campaigns for desegregation
- [00:21:34 - 00:33:30] Struggling to find successful desegregation plan
- [00:33:31 - 00:39:09] Blacks and Native Americans form political alliance
- [00:39:10 - 00:42:24] Desegregation creates white need for black political support
- [00:42:25 - 00:56:15] Blacks sue to enforce integration order
- [00:56:16 - 01:03:52] African American man describes legal struggle to enforce integration
- [01:03:53 - 01:09:17] Fighting against accommodation in black community
- [01:09:18 - 01:12:46] Black reluctance to resist white hegemony
- [01:12:47 - 01:16:56] Desegregation creates opportunities for African Americans
- [01:16:57 - 01:20:59] School mergers improve resources
- [01:21:00 - 01:32:37] Threat of violence hangs over black activists
Getting more judgments

CHARLES THOMPSON: Well, let's talk about the flood itself.

BILLY RAY HALL: Sure.

CHARLES THOMPSON: What are the effects of the flood as you describe them?

BILLY RAY HALL: I think anytime you start talking about numbers they are only understandable when you put them in relation to something. For example, if you say you had a hundred dollars worth of damage, you're in a developing nation where you house costs 150 dollars and is made out of natural materials, it's a lot different than in America when you have a hundred dollars worth of damage. So the first thing to do is let's put perspective. This was a six billion dollar disaster in North Carolina. The entire agriculture output for the state of North Carolina last year was seven billion dollars. So six billion dollars is huge relative to this state's economy, relative to our resource base. And that is big by any stretch of the imagination—600,000 acres of crops were damaged or destroyed and 7,000 are totally destroyed. So you get really huge numbers to go along with the fifty-one people that lost their lives. What that does is it gives you a feel for the size of the community—we lost 190 million dollars of cotton in the fields that were flooded. In addition, we lost over 300 million dollars of other crops in the field for a total of 538 million dollars worth of crops that were washed away.
Manual segment lengths
Different segmentation granularities

Fig 6. Possible boundaries per boundary placed
Fig. 5. Mean boundary edit distance between annotators' segmentations, by transcript.
Automatic segmentation
Automatic segmentation

- Considerable work on topically segmenting texts
- Transcripts of broadcast news, lectures, and meetings
- MALACH project (Franz et al. 2003) used supervised learning to segment Shoah interviews

Franz et al., “Automated transcription and topic segmentation of large spoken archives”
Unsupervised segmentation

- Assumption: greater similarity in the language within topically coherent segments than across such segments
- How to define language similarity?
- How to postulate boundaries?
Automatic segmentation 👎

Fig. 2: All differences are significant except between TextTiling and Uniform.
Coreference resolution
MALINDA MAYNOR:

Now go ahead for us Julius Chambers.

ANGUS BOAZ THOMPSON SR.:

Julius had just sued UNC so he could go to school there. So I contacted Julius, and he agreed to come and look our situation over and help us. So we were meeting at McMillan Funeral Home. He came down twice to talk to us and look the situation over. The second time he left and went back, he called me. He said, "Angus, I've been in communication with your black city councilman. I've been in communication with your school board superintendent, Dr. Carroll."

MALINDA MAYNOR:

Is it Gale, G-A-L-E?

ANGUS BOAZ THOMPSON SR.:

Carroll, C-A-R-R-O-L-L.

MALINDA MAYNOR:

I'm sorry, okay.

ANGUS BOAZ THOMPSON SR.:

He said, "Neither one of them want to compromise. In fact, the superintendent told me that he wasn't interested in meeting with our group over there. Nothing he can do for them. He said, "The city councilman said there were two different bodies. The school board was independent. The city council was independent of the school board. He said, "Well, they ought to know that I know better than that. All city councils work together with the school board on certain issues."

He said, "In the first place the school board didn't zigzag that line out there like it did to build that school." He said, "The city council had to put it out there. They did it for the school board." He said, "That councilman insulted my intelligence. He said, "He really made me hot, so I'm going to tell you what you do." He said, "You get your group together down there." He said, "When they did that they broke a statute of the state of North Carolina." He told me what that statute was. At that time if you extend the city limit you had to take into consideration all the [unclear] of the residents, and they did not, not one. He said, "We've got to put heat to the seat, sue the school board and the city council." So he came down the third time. He said, "Be ready to show me what you want in the school district." That's the way he was going to approach it and soon.

So he came down the third time. When he was getting ready to go, as I've said we always met at McMillan Funeral Home, our black city councilman was in the rear. I didn't know he was in there, but when he heard Chambers make his plan of what he was going to do, he immediately jumped up and said, "No, don't sue. We can help you. I can help you."
"Jack Shaws, who was my chief investigator"

Jack Shaws [PERSON]

238: [Jack Shaws] I had Jack Shaws, who was my chief investigator.

238: [Jack Shaws, who was my chief investigator] I had Jack Shaws, who was my chief investigator.

239: [He] He had retired as the chief of detectives, assistant chief of police for the Montgomery Police Department.

240: [a young detective] When he was a young detective he had solved the bombings in Montgomery of Dr. King's house and his church, and made arrests.

240: [he] When he was a young detective he had solved the bombings in Montgomery of Dr. King's house and his church, and made arrests.

240: [his] When he was a young detective he had solved the bombings in Montgomery of Dr. King's house and his church, and made arrests.

241: [Jack] Jack had brought with him to our office one of his other retired detectives that we'd hired named Tom Ward who was a terrific detective.

241: [him] Jack had brought with him to our office one of his other retired detectives that we'd hired named Tom Ward who was a terrific detective.

276: [Jack Nelson] I was in Washington, and I had a good friend that was the LA Times bureau chief named Jack Nelson.

277: [Jack] Jack originally was from Alabama.

278: [Jack] When I would go to Washington I would try to get with Jack, and we'd go out to dinner or something.

279: [his] I'd go over to his house.

280: [He] He had been active in covering some of the civil rights bombings in Mississippi on behalf of the Times, LA Times.

281: [he] So we went out one night, and he wanted to know if I was still working on the bombing case, and I said, ``Yeah.''

282: [him] I said, ``I think we've reached an impasse,'' and I told him about the trouble.

283: [He] He knew I'd been working on it for years.

284: [him] I told him about the problem that we had, if we couldn't get the cooperation of the FBI that I did not think we were going to.

285: [He] He said, ``I think I can help you if you want me to.''

286: [He] He said, ``If you do n't think it's meddling, I'll do what I can to help you.

302: [Nelson] It was either Levi or Saxby, one or the other, but Nelson knew them from his work.

302: [his] It was either Levi or Saxby, one or the other, but Nelson knew them from his work.

304: [Nelson] Nelson goes up there to the Attorney General's office and bypasses the FBI.

305: [He] He says he's working on a big story, and it's not going to be just a story it's going to be a whole series of stuff, and that.

305: [he] He says he's working on a big story, and it's not going to be just a story it's going to be a whole series of stuff, and that.

310: [Nelson] About I think it would have been two weeks, they called Nelson back.

315: [Nelson] I think I was with Nelson.

335: [Jack Nelson] So I'll always be grateful and indebted to Bob Eddy and Jack Nelson both.

451: [Jack Shaws] They had more evidence than what I had to convict Chambliss, but, and this is a big but -- I mentioned early on when
These charts show movie character interactions. The horizontal axis is time. The vertical grouping of the lines indicates which characters are together at a given time.

Lord of the Rings

Star Wars (Original Trilogy)

Jurassic Park

12 Angry Men

Primer

https://xkcd.com/657/
Entity Flow

This is a d3 layout for visualizing patterns of entity participation in some sequence. It is inspired by xkcd’s movie narrative charts as well as subsequent efforts to automatically generate such charts, notably Tanahashi & Ma 2012 and Liu et al. 2013. The algorithm implemented here is not nearly as sophisticated or capable as those. It is based heavily on Mike Bostock’s Sankey plugin.
entity
Access via visualization

**Bill Baxley**: convicted Blanton for sure back when we convicted Chambliss and maybe Cherry also, although Cherry would be a little bit more iffy because a lot of stuff that they convicted Cherry on came out of his own mouth after the fact, bragging about what he had done.

**Willoughby Anderson**: Right.

**Bill Baxley**: But Blanton I think for sure we could have convicted back when we did Chambliss.

**Willoughby Anderson**: So you didn’t have access to any of the tapes that they used or didn’t know anything about him?

**Bill Baxley**: No. No. Well, we knew a lot about him. He was our main suspect.

**Willoughby Anderson**: Okay. Okay.

**Bill Baxley**: Oh yeah. We knew it was his car. We knew he was driving the car.

**Willoughby Anderson**: Okay, okay.

**Bill Baxley**: But like the ID we had. We had a positive ID of Chambliss, and a positive ID of Blanton’s car, but not a positive ID of Blanton. Things like that, there was no question. In fact our suspects were Blanton, Cherry, and Chambliss, and then some other people that helped. There was one other one with them. I went to the FBI about this and got it out of my system kind of. And I don’t blame the ones now. They weren’t there. They weren’t the ones doing it even though the ones now, it hurts their feelings to be criticized. But the FBI office of today is not the one of the ’60s and ’70s, and even saying that, the average agent wasn’t at fault. Had it not been for the FBI there wouldn’t have been anything. I shudder to think what would have happened in the South with racial hate crimes back then. That doesn’t mean that I think they’re all bad. They did a lot of good there. We ended up deciding just to indict Chambliss.

**Willoughby Anderson**: Okay.

**Bill Baxley**: For several reasons, and that was my call. Our evidence on Chambliss was stronger than on the others, and Chambliss also was the ring leader. He was the one that really made it happen. All the others were haters, and all of them were violent, bad people but they were followers. If you’d have removed Chambliss from the scene back then and had all the others here, you wouldn’t have had very many bombings. So we felt like not only getting the ring leader, but if we got him, was the one they...
Thanks!

Ryan Shaw
ryanshaw@unc.edu
https://aeshin.org/