NLP assignment 3 report

Damiano Buzzo`<damiano25@ru.is>`

T-725-MALV Natural Language Processing

RU Computer Science

October 25, 2025

Contents

1   Natural discourse

1.1   Transcription Task

This is the manual transcription made from this youtube video https://www.youtube.com/watch?v=IVVAYccdKX8, from CBS sport trasmission(from the minute 8:22 to 10:29) All the people in this discouse will be indicated as follow: Alessandro Del Piero(ADP), Micah Richards(MR), Jamie Carragher(JC), Henry(H), Mc Kennie(MK).

1. **Segment A**

2. ADL: My heart is here, my my house is here

3. MR: Big, big house!

4. (Everyone laughs)(MR and ADP are shoving each other)

5. MR: come si dice GRANDE, GRANDE (how can I say BIG, BIG)

6. (Everyone laughs) (MR and ADP are shoving each other in a stronger way)

7. ADL: so you speak Italian when you want

8. MR: dipende, dipende(it depends, it depends)

9. **Segment B**

10. JC : Alessandro, should I ask to Henry the same question?

11. H: oh you can ask me!

12. (Everyone howling)

13. JC : is this the same for you? I mean..

14. H: listen to me, I will go to Arsenal as a Kitman

15. MR: I would ask the same question to you Jamie, but Liverpool don't want you

16. (everyone laugh)

17. **Segment C**

18. JC: oh big M is back! you are back, you know you are back big man!!

19. (MR takes his sunglasses and wear them)

20. H: look at him!! look at him!!

21. JC: the stage is yours

22. H: oh I love that!! what's the hat?

23. M: it was a gift, just I'm just a little proud! but anyway being with Alessandro is like being with a king

24. (ADP appreciates, making a bow)

25. M: what a superstar you are! and you are so humble with it. I wish I would be like you

26. (everyone laugh)

27. **Segment D**

28. (they are showing a video)

29. MR: look at this! we went for a little ride together. Remiscent, NOSTALGIA! and then DEL PIERO look at this! look at this we are like movie stuff!

30. (everyone laugh)

31. ADP: we had a good time, we had a good time!

32. **Segment E**

33. (Mc Kennie arrives)

34. MR: this is your town now, this is your town now?

35. MK: yeah, yeah

36. H: I have one for you! Micah always want to know trim and whatever can you take his hat off

37. (MK take the micah hat)

38. (everyone scream)

39. MK: it's a very big, it's a very big one...

1.2   Discussion

This analysis reflects on transcribing a segment of the CBS Sports Golazo talk show, highlighting the challenges of capturing face-to-face discourse in a dynamic, informal setting.

A sufficient record for understand what took place requires capturing the interactive and affective state of the participants to explain why and how things are said. This necessitates including physical actions (e.g. (MR and ADP are shoving each other)) and affective cues (e.g. (Everyone laughs)). These non-verbal elements are crucial for understanding the segment's function, which is primarily lighthearted banter and cameradie, not just serious commentary.

While I captured the comedic peaks and transitions, like MR wearing his sunglasses to signal his exaggerated "Big Man", it's impossible to include everything. An exhaustive record would require precise timing of pauses and overlaps, and accurate descriptions of intonation and subtle body language. These minute details were omitted to maintain readability and avoid visual clutter.

I intentionally did not clean spontaneous speech features like repetitions ("dipende, dipende") or code-switching ("come si dice..."). These are not errors; they are conversational signals. The repetition serves as emphasis or a filler, while code-switching highlights the cross-cultural humor. To "clean" the transcript would be to remove the very texture of the naturally occurring discourse. So I think that I left the right amount of details.

The most difficult parts to transcribe were the non-spoken actions requiring a judgment of intensity and intention. For example, describing the shoving as simply "in a stronger way" is a vague interpretation (summarized) of a physical escalation. Similarly, the note (ADP appreciates, making a bow) misses the theatricality of the gesture, which the subsequent dialogue suggests. The transcript can use words to map the actions, but it can never fully replicate a live performance.

1.3   Reflection on computation

The greatest challenge for a computer participating in this discourse, even assuming perfect speech and word recognition, would be pragmatic interpretation and understanding what is meant rather than what is literally said. The core difficulty lies in mapping non verbal and physical actions to conversational intent. For instance,

the computer would struggle to classify the act of (MR and ADP are shoving each other); without human social context and affective knowledge, it might flag this physical action as genuine aggression, entirely missing the social signal of playful banter and familiarity that defines their relationship and sets the tone for the exchange. Furthermore, the computer must decode the metaphorical role shifts that occur. When MR puts on his sunglasses and JC says "the stage is yours," the AI must understand that this sequence(a simple physical act followed by a metaphorical declaration) signifies the intentional adoption of a comedic persona ("big M"), which is crucial for interpreting the subsequent enthusiastic comments. Ultimately, the computer lacks the shared history and emotional context necessary to distinguish genuine aspiration from sarcasm ("I wish I would be like you") or to appreciate the comedic timing inherent in spontaneous, multi-layered human interaction.

2 Discourse Structure

2.1 Segmentation task

The designed segment are:

- Segment A from line 2 to 8

- Segment B from line 10 to 16

- Segment C from line 18 to 26

- Segment D from line 28 to 31

- Segment E from line 33 to 39

2.2 Segment Analysis task

Segment	1. Strongest Evidence, Purpose, or Rhetorical Relation	2. Explicit Start Signal	3. Explicit End Signal	4. Linear or Hierarchical Structure
A	The segment is about **ADL's connection to the city** ("My heart is here, my house is here") and **MR's Italian language ability**. The evidence for separation is the clear topic shift, initiated by ADL's statement, moving away from whatever preceded it. The final exchange ("so you speak Italian..." / "dipende") resolves this particular subtopic.	No	No	**Linear**. It introduces a new topic in the conversation.

B	This segment **shift the focus and attention from ADL to H** (Henry) by asking him the "same question." The strongest evidence is the explicit **shift in addressee and topic** (from ADL/MR to H/JC), signaled by **JC's direct question** and the fact that he call his name. The final joke by MR serves as a closing to this interjection.	No. It begins with a direct question and name-call ("Alessandro, should I ask to Henry the same question?"), which acts as a **functional transition** but is not a discourse connective.	No	**Linear**. It's just next turn in the conversation's flow, with a change in the person being addressed.
C	This is a **segment of appreciation and spectacle**, focusing on **M's entrance and attire** (sunglasses and hat) and his subsequent praise for Alessandro Del Piero. The evidence is the sharp change in focus to M's appearance, signaled by JC's enthusiastic address. The core purpose is M's **rhetorical act of homage** ("being with Alessandro is like being with a king").	It begins with "oh" ("oh big M is back!"), which functions as an attention-getter.	No	**Linear**. It follows the previous segment, allowing M to take the "stage" after the attention shifted to H in Segment B.
D	The segment is clearly separated by the **introduction of a new medium** (a video) and a **shift in topic to a shared memory** between MR and ADP. MR's exclamations ("look at this!") and the shared theme of "Nostalgia" and "a good time" define its boundaries and purpose.	No. It begins with the stage direction indicating the video start, followed by MR's immediate reaction.	No	**Linear**. It follows the previous segment, a natural shift in the conversation's flow as a video is introduced as a new point of focus.
E	The segment's purpose is to **incorporate a new person (MK, McKennie) into the discourse** and create a new comedic event. The evidence is the stage direction indicating **MK's arrival** and the subsequent, distinct **interaction around MK's local knowledge and the removal of M's hat**. The entire event centers on this new action.	No. It starts with the stage direction of MK's arrival, followed by MR's greeting/question.	No	**Linear**. It's the next event in the sequence, triggered by the entrance of a new participant.

2.3   Discussion

In my opinion it's highly unlikely that another person would place the segment boundaries in the exact same locations, as discourse segmentation is often subjective. The segments (A-E) were defined by shifts in topic, addressee, or external action (for example a video starting or a person arriving). While the initiation of a new activity (for example JC addressing Henry in B, or McKennie arriving in E) provides a strong anchor for a boundary, the internal boundaries are highly fluid.

The primary topic of disagreement would be the treatment of jokes, quick retorts, and laughter as distinct segments. If someone wants to prioritize turn-by-turn conversational mechanics can create numerous small segments for each joke or reaction within the playful atmosphere of A and B. Conversely, if you want to focus on rhetorical structure you can combine Segments B and C, arguing that the entire sequence—shifting focus from Alessandro to Henry and then to Micah's attire—serves the single rhetorical purpose of playful banter and celebrity focus, overriding the minor shifts in topic.

A purely rule-based computer algorithm could perform a similar segmentation, but it would be brittle and require complex programming to handle non-linguistic cues. The success of the segmentation would depend heavily on recognizing the structure of the input, such as:

1. Explicit changes in speaker turn

2. Direct address/name-calling

3. The explicit stage directions like (Mc Kennie arrives) or (they are showing a video), which function as unambiguous boundary markers in this transcript.

The main difficulty for such an algorithm lies in interpreting the function of the discourse rather than just the words. For instance, distinguishing when a sequence of jokes constitutes a single, cohesive segment (like A or B) versus when a boundary truly marks a shift in the central meaning or activity requires a high level of semantic and pragmatic understanding.

3   Discourse Function vs. Device

3.1   Classification task

#	Phenomenon	Classification
1	Requesting payment	Discourse Function
2	Acting surprised	Discourse Function
3	Requesting the speaking turn	Discourse Function
4	Saying "I do not have any money!"	Discourse Device
5	Crossing the arms	Discourse Device
6	Showing sympathy	Discourse Function
7	Smiling and shaking the head	Discourse Device
8	Uttering "Don't worry"	Discourse Device
9	Showing the signs of relief	Discourse Function
10	Making the sound "Ah" (Or "/æ/ /h/")	Discourse Device

3.2   Spotting examples

#	Discourse Device	Segment/Line	Discourse Function
1	Exclamation/Repetition	MR: Big, big house!	Exaggerating
2	Code-Switching	MR: come si dice GRANDE, GRANDE(how can I say BIG, BIG)	Joking about his Italian language knowledge
3	Non-Verbal Action (Shoving)	(MR and ADP are shoving each other)	Establishing playfulness
4	Attention Getter (Direct Address)	JC : Alessandro, should I ask to Henry the same question?	Initiating topic shift
5	Metaphor/Hyperbole	M: ...being with Alessandro is like being with a king	Showing appreciation