End-of-Turn Detection

Sean Leishman

MInf Project (Part 1) Report
Master of Informatics
School of Informatics
University of Edinburgh

2023

Abstract

This skeleton demonstrates how to use the `infthesis` style for undergraduate dissertations in the School of Informatics. It also emphasises the page limit, and that you must not deviate from the required style. The file `skeleton.tex` generates this document and should be used as a starting point for your thesis. Replace this abstract text with a concise summary of your report.

Research Ethics Approval

Instructions: *Agree with your supervisor which statement you need to include. Then delete the statement that you are not using, and the instructions in italics.*
Either complete and include this statement:
This project obtained approval from the Informatics Research Ethics committee.
Ethics application number: ???
Date when approval was obtained: YYYY-MM-DD
[If the project required human participants, edit as appropriate, otherwise delete:]
The participants' information sheet and a consent form are included in the appendix.
Or include this statement:
This project was planned in accordance with the Informatics Research Ethics policy. It did not involve any aspects that required approval from the Informatics Research Ethics committee.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Sean Leishman*)

Acknowledgements

Any acknowledgements go here.

Table of Contents

Chapter 1

Introduction

The preliminary material of your report should contain:

- The title page.
- An abstract page.
- Declaration of ethics and own work.
- Optionally an acknowledgements page.
- The table of contents.

As in this example `skeleton.tex`, the above material should be included between:

```
\begin{preliminary}
    ...
\end{preliminary}
```

This style file uses roman numeral page numbers for the preliminary material.

The main content of the dissertation, starting with the first chapter, starts with page 1. ***The main content must not go beyond page 40.***

The report then contains a bibliography and any appendices, which may go beyond page 40. The appendices are only for any supporting material that's important to go on record. However, you cannot assume markers of dissertations will read them.

You may not change the dissertation format (e.g., reduce the font size, change the margins, or reduce the line spacing from the default single spacing). Be careful if you copy-paste packages into your document preamble from elsewhere. Some LaTeX packages, such as `fullpage` or `savetrees`, change the margins of your document. Do not include them!

Over-length or incorrectly-formatted dissertations will not be accepted and you would have to modify your dissertation and resubmit. You cannot assume we will check your submission before the final deadline and if it requires resubmission after the deadline to conform to the page and style requirements you will be subject to the usual late penalties based on your final submission time.

1.1 Turn-taking Prediction

1.1.1 From the Conversation Analysis Perspective

The most widely accepted model of turn-taking originates from Sacks et al. [1974]. They created their model based off a clear pattern in behaviour during social interaction. With this in mind they determine that the behaviour exhibits a joint and coordinated effortto determine turns while attempting to have a minimal number of turns. Although there is this effort, turn-taking organisation is not preplanned and it is coordinated in a manner that changes along the course of a conversation. During a dialogue Sacks et al. [1974] note that convservation is dominated by a single speaker at a time although multiple speakers at one time do occur commonly but they are brief. When the speaker switches, the transitions are characterised with no gap or overlap, most commonly, but also with a small overlap or gap. From these observations Sacks et al. [1974] suggested a series of definitions for these units of speech and rules to describe their interaction. Turn-taking is separated into units of speech called *Turn-Constructional-Units (TCU)* where predominately one speaker is speaking and after each TCU there is a *Transition-Relevant-Phase (TRP)* where the current speaker switches (turn-shift). Sacks et al. [1974] noted that a turn-shift can but does not have to occur. The same could be said in reverse, that a turn-shift can occur when there is no TRP. At each TRP Sacks et al. [1974] noted the turn-allocation techniques used to determine the next speaker. These techniques can be grouped into two categories: 'current selects next' and 'self-selection'. The current speaker selection could employ techniques such as gaze but for self selection it is simply the case that the first to start has the turn if the current speaker exhibits a non-selection. The current speaker is also able to select themselves if no other particpant self-selects.

The development of speech technology has allowed for the automatic analysis of turn-taking by employing statistical analyis of a large corpora. Levinson and Torreira [2015] provides a systemic overview of the properties first noted by Sacks et al. [1974] and produced their own model of turn-taking where: turns are mostly short (mean=1680ms, median=1227ms); gaps in a speaker's current utterance are longer than those between speakers suggesting that the next speaker has the right to speak prior to the current speaker and also that overlaps are more common at turn transitions and within turns. They also develop a psycholinguistic model for turn-taking that backs the claim made by Sacks et al. [1974] that turn-taking requires projection. ¡F8¿ Levinson and Torreira [2015] turn transitions typically occur between -100ms and 500ms. This is proven probelmatic, if we employ a reaction-based system, where studies have shown that language production latenices range from 600ms - 1500ms Indefrey and Levelt [2004], Bates et al. [2003] and are at minimum 200ms (for a prepared vowel Fry [1975]). A combination with a silence becoming reognisable (180ms), the reaction to a silence (100ms) and language productions would result in at least a 550ms turn transition, outside the typical range.

As such, the question remains, how, if a non-speaker wishes to self-select as described in Sacks et al. [1974], can they predict a TRP if they are not simply responding to silence as claimed by Heldner and Edlund [2010].

1.1.1.1 Turn-taking Cues

The first systematic studies of turning taking cues came from Duncan [1972, 1974] where they identified a number of turn-taking cues that signal turn-completion across a few different domains. Syntacially complete phrases, phrase-final intonation, termination of hand gesticulation were all identified as turn-taking cues. In literature, it appears that syntactic and prosodic features tend to have received the most focus as turn-taking cues. Sacks et al. [1974] first argued that syntax and semantics played a more influential role in prediction than a prosodic feature such as intonation. This idea was investigated further by Ford and Thompson [1996] who concluded that intonational completion re than previously considered, as compared to linguistic features such as syntactic completeness. Ford and Thompson [1996] defined an utterance as being `syntactically complete` if "in it's discourse context, it could be intepreted as a complete clause, that is, without an overt or directly reoverable predicte without considering intonation of interactional import". They also define another feature of pragmatic completness that is judged as the completion of intonational and conversational action sequencing. (Explain conversational action sequencing?). Ford and Thompson [1996] found that points of both pragmatic and intonational complete are nearly always syntactic complete points however the reverse is not always true. As such, they defined these points of pragmatic, intonational and syntactic completeness as a Complex Transition Relevance Place, where a 71% of speaker changes occur at CRTPs. (CRITIQUE). Ford and Thompson [1996]

Chapter 2

Your next chapter

A dissertation usually contains several chapters.

Chapter 3

Conclusions

3.1 Final Reminder

The body of your dissertation, before the references and any appendices, *must* finish by page 40. The introduction, after preliminary material, should have started on page 1.

You may not change the dissertation format (e.g., reduce the font size, change the margins, or reduce the line spacing from the default single spacing). Be careful if you copy-paste packages into your document preamble from elsewhere. Some LaTeX packages, such as `fullpage` or `savetrees`, change the margins of your document. Do not include them!

Over-length or incorrectly-formatted dissertations will not be accepted and you would have to modify your dissertation and resubmit. You cannot assume we will check your submission before the final deadline and if it requires resubmission after the deadline to conform to the page and style requirements you will be subject to the usual late penalties based on your final submission time.

Bibliography

Elizabeth Bates, Simona D'Amico, Thomas Jacobsen, Anna Székely, Elena Andonova, Antonella Devescovi, Dan Herron, Ching Ching Lu, Thomas Pechmann, Csaba Pléh, et al. Timed picture naming in seven languages. *Psychonomic bulletin & review*, 10: 344–380, 2003.

Chen-Chung Chang and H. Jerome Keisler. *Model Theory*. North-Holland, third edition, 1990.

Starkey Duncan. Some signals and rules for taking speaking turns in conversations. *Journal of personality and social psychology*, 23(2):283, 1972.

Starkey Duncan. On the structure of speaker–auditor interaction during speaking turns1. *Language in society*, 3(2):161–180, 1974.

Cecilia E Ford and Sandra A Thompson. Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns. *Studies in interactional sociolinguistics*, 13:134–184, 1996.

Dennis B Fry. Simple reaction-times to speech and non-speech stimuli. *Cortex*, 11(4): 355–360, 1975.

Mattias Heldner and Jens Edlund. Pauses, gaps and overlaps in conversations. *Journal of Phonetics*, 38(4):555–568, 2010.

Peter Indefrey and Willem JM Levelt. The spatial and temporal signatures of word production components. *Cognition*, 92(1-2):101–144, 2004.

Stephen C Levinson and Francisco Torreira. Timing in turn-taking and its implications for processing models of language. *Frontiers in psychology*, 6:731, 2015.

H. Sacks, E.A. Schegloff, and G. Jefferson. A simplest systematics for the organization of turn-taking for conversation. volume 50, page 696 – 735. 1974.

Appendix A

First appendix

A.1 First section

Any appendices, including any required ethics information, should be included after the references.

Markers do not have to consider appendices. Make sure that your contributions are made clear in the main body of the dissertation (within the page limit).

Appendix B

Participants' information sheet

If you had human participants, include key information that they were given in an appendix, and point to it from the ethics declaration.

Appendix C

Participants' consent form

If you had human participants, include information about how consent was gathered in an appendix, and point to it from the ethics declaration. This information is often a copy of a consent form.