Eberhard Karls Universität Tübingen

Seminar für Sprachwissenschaft

**Contrastive Dialogues Involving Ellipsis (?)**

Thesis submitted for the degree of Master of Arts

Supervised by

Prof. Dr. Michael Franke and

Jun.-Prof. Dr. James Griffiths

Summer Term 2023

Miriam Schiele

miriam.schiele@student.uni-tuebingen.de

Matriculation Number: 6098290

Käthe-Kollwitz-Straße 15

72074 Tübingen

Date of Submission: September 30, 2023

Abstract

This research co

Contents

1. Introduction 5

2. Literary review 5

3. Data and method 5

4. Findings 6

5. Discussion 6

6. Conclusions 6

7. References 7

8. Appendix 8

8.1 Notational Conventions 8

Table of figures

# 1. Introduction

In morphology,

Some basic, methodologically-oriented work is needed for the foundations of the project. This concerns determining the best medium for reliably obtaining judgments about dialogues that involve contrastive focus, (e.g. (1))  
(1) A: I heard that John is travelling from Africa tomorrow.  
    B: No, to.  
In (1), "from" and "to" contrast. (1B) feels like an odd answer to (1A), but I think that's because the reader isn't primed for the contrast. The following seems better:  
(2) A: I heard that John is travelling \*FROM\* Africa tomorrow.  
    B: No, \*TO\*.  
I want to know if orthographic marking of contrast is sufficient for priming. To find this out, I need to run an experiment that compares differs media: I would compare (1), (2), and their verbal equivalents (i.e., the verbal stimuli would be someone's recorded voice). Likert scale responses; I'm unsure yet whether between- or within-subject design would be best.  
Your student could conduct this study using German stimuli, she could record her own voice (or a friend's) for the verbal stimuli. We would z-score the 1-7 Likert scale responses and fit LMMs, using R. There's a bit of prior work on conducting experiments on contrast, and some prior work on contrast in dialogues like (1) and (2), but not much!

# 2. Literary review

Natural Language

Overview of research on L2 Subordination

# 3. Data and method

The data used in

Data base: EF Cam Dat

Tool: CTAP

Measures: MLC, clauses per sentence (fewer measures are better)

Description of CTAP (see PhD example with Meurers):

“The Common Text Analysis Platform (CTAP) (Chen and Meurers, 2016b, see Chapter 2) developed by the author extends the functionalities of previous complexity analysis systems by adding modules for corpus management, feature selection, and results visualization. The system was implemented as a Web application with a friendly user interface that makes it easy to be used by linguists and researchers who are not familiar with computer programming or NLP technologies. The first release of the system provides over 170 lexical and syntactic measures. The integration of the full set of over 570 complexity measures listed in Appendix A is underway.“

# 4. Findings

Using t-SNE, the semantic

# 5. Discussion

Since fastText allows for an analysis that

# 6. Conclusions

This paper gives an overview of the

# 7. References

Baayen, R. H. 2008. *Analyzing linguistic data*: *A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

# 8. Appendix

## 8.1 Abbreviations, symbols and other notational conventions

? questionable/marginal acceptability

\* unacceptable

**bold** pitch accent

FocP focused position

Declaration of Authorship

I hereby confirm that this paper and the work presented in it is entirely my own. Where I have consulted the work of others this is always clearly stated. All statements taken literally from other writings or referred to by analogy are marked and the source is always given. This paper has not yet been submitted to another examination office, either in the same or similar form.

Tübingen, December 6th, 2022



Miriam Schiele