Motivation

The aim of the project is to investigate how people process language with respect to personal and demonstrative pronouns (ER vs. DER pronouns). In particular, given two perfectly grammatical sentences:

1. Der Oberarzt untersucht den Patienten.
2. (a) Er ist nicht Herzspezialist.
   (b) Er ist nämlich Herzspezialist.

why is there a preference for one pronoun form over the other?

Research Question

The formal task set out was to examine how the DER / ER pronouns are used in natural language and to produce a theoretical framework that accounts for speakers’ preference by characterising the syntactic, semantic, pragmatic and processing differences in the use of the DER and ER pronouns in everyday German discourse.

Sentence Internal Features

Two corpora were made available to the project group: the Frankfurter Rundschau Corpus (~ 4 million tokens) and the NEGRA Treebank (~ 0.35 million tokens). Sentences containing DER & ER pronoun forms were extracted using regular expressions. The pie chart below is a comparison of the number of occurrences of each pronoun form in the FR Corpus.

The pronoun forms were subsequently classed under two types of divisions: the Subject/Object division and the Vorfeld/Mittelfeld division.

Figure 1: Total no. of tokens of ER & DER-forms in FR Corpus

Sentence-internal features chosen for further investigation

In isolating these sentence-internal features, it was hoped that distinct patterns of distribution for each pronoun type would emerge with respect to these features and as the chart below shows, this turned out to be the case.

Figure 2: Sentence-internal features chosen for further investigation

The statistics above suggest that the DER pronoun prefers a fronted position in the sentence in which it occurs. We wanted to test whether native speakers were indeed sensitive to this preference and designed our first psycholinguistic experiment around this goal.

Experiment One

The idea behind this experiment was to take a closer look at those sentence internal features described above (Subject/Object & Vorfeld/Mittelfeld) and to test how varying these features would affect the preferences underlying the choice of the DER and ER pronouns.

Thirty participants were given a self-paced reading task with the sentences in which the position of the pronoun (either ER- or DER-pronoun) varied with respect to case (nominative/non-nominative) & position in the sentence (Vorfeld/Mittelfeld). Subsequently they were given a questionnaire containing sentence pairs (both ER and DER forms) and asked to choose which of the two they felt to be the most appropriate.

Figure 3: Distribution of DER & ER pronouns in FR Corpus with respect to sentence-internal features chosen for investigation

The reading time for sentence (3) was measured. We found a significant difference between the DER and ER version of sentence type (3a). In other words, when world knowledge supported reference to a nominative antecedent, but there was no actual antecedent, participants were asked to complete a questionnaire.

Figure 4: Reading time course for sample text sentence

In the experiment we wanted to investigate the influence of context on the usage of the DER & ER pronouns, as well as finding evidence to support our hypothesis.

Figure 5: Results of questionnaire comparing appropriateness of pronoun form for a given sentence

Two corpora were made available to the project group: the Frankfurter Rundschau Corpus (~ 4 million tokens) and the NEGRA Treebank (~ 0.35 million tokens). Sentences containing DER & ER pronoun forms were extracted using regular expressions. The pie chart below is a comparison of the number of occurrences of each pronoun form in the FR Corpus.

The statistics above suggest that the DER pronoun prefers non-topical referents. Further work would include operationalising the term “topic-hood” using appropriate features and testing these features against native-speaker preferences.

Figure 6: Diagrammatic representation of antecedent tagging procedure

Background

The formal task set out was to examine how the DER / ER pronouns are used in natural language and to produce a theoretical framework that accounts for speakers’ preference by characterising the syntactic, semantic, pragmatic and processing differences in the use of the DER and ER pronouns in everyday German discourse.

Figure 7: Screenshot of the tagging user-interface

The formal task set out was to examine how the DER / ER pronouns are used in natural language and to produce a theoretical framework that accounts for speakers’ preference by characterising the syntactic, semantic, pragmatic and processing differences in the use of the DER and ER pronouns in everyday German discourse.

Figure 8: Distribution of antecedents in previous sentence

The formal task set out was to examine how the DER / ER pronouns are used in natural language and to produce a theoretical framework that accounts for speakers’ preference by characterising the syntactic, semantic, pragmatic and processing differences in the use of the DER and ER pronouns in everyday German discourse.

Figure 9: Sample text sequence for Experiment Two

The statistics above suggest that the DER pronoun prefers a fronted position in the sentence in which it occurs. We wanted to test whether native speakers were indeed sensitive to this preference and designed our first psycholinguistic experiment around this goal.

Experiment One

The idea behind this experiment was to take a closer look at those sentence internal features described above (Subject/Object & Vorfeld/Mittelfeld) and to test how varying these features would affect the preferences underlying the choice of the DER and ER pronouns.

Thirty participants were given a self-paced reading task with the sentences in which the position of the pronoun (either ER- or DER-pronoun) varied with respect to case (nominative/non-nominative) & position in the sentence (Vorfeld/Mittelfeld). Subsequently they were given a questionnaire containing sentence pairs (both ER and DER forms) and asked to choose which of the two they felt to be the most appropriate.

Figure 3: Distribution of DER & ER pronouns in FR Corpus with respect to sentence-internal features chosen for investigation

The statistics above suggest that the DER pronoun prefers a fronted position in the sentence in which it occurs. We wanted to test whether native speakers were indeed sensitive to this preference and designed our first psycholinguistic experiment around this goal.

Experiment One

The idea behind this experiment was to take a closer look at those sentence internal features described above (Subject/Object & Vorfeld/Mittelfeld) and to test how varying these features would affect the preferences underlying the choice of the DER and ER pronouns.

Thirty participants were given a self-paced reading task with the sentences in which the position of the pronoun (either ER- or DER-pronoun) varied with respect to case (nominative/non-nominative) & position in the sentence (Vorfeld/Mittelfeld). Subsequently they were given a questionnaire containing sentence pairs (both ER and DER forms) and asked to choose which of the two they felt to be the most appropriate.

Figure 3: Distribution of DER & ER pronouns in FR Corpus with respect to sentence-internal features chosen for investigation

The statistics above suggest that the DER pronoun prefers a fronted position in the sentence in which it occurs. We wanted to test whether native speakers were indeed sensitive to this preference and designed our first psycholinguistic experiment around this goal.