1. The German Recipient Passive

- Formed by: {bekommen, kriegen, erhalten} + past participle
- Alternatives: active, regular passive

The Classifier

- Maximum Entropy classification (MegaM)
- Build a loglinear model for voice prediction
- Learns to predict the voice of a sentence given a feature representation of it
- Features in terms of semantic roles: Argument Type, Grammatical Features, Syntactic Complexity, Surface Order
- Discourse givenness/newness encoded by morpho-syntactic properties

Results

Accuracy: 98.05% (Baseline: 94.6%)

2. Building the Data Set

- We extracted the most frequent verbs using the German Recipient Passive by means of flat PoS patterns
- We parsed all sentences containing these verbs using a hand-crafted, broad-coverage Lexical Functional Grammar

3. Classification Experiments

The Classifier

- Accuracy: 98.05% (Baseline: 94.6%)

Data

- We mapped F-Structures to labelled feature sets
- Non-agent triviality: Divide test data into 2 subsets

4. Data Analysis

- Each argument is most often pronounriminalized in the voice where it is realised as the subject
- The agent is more likely to be a non-third person entity in the active than in the passive voices
- Definiteness does not match theoretical hypotheses: the recipient is more often indefinite in recipient passives than in active.
- The definiteness features of the agent agrees in the hypotheses
- The least complex argument in a voice is that which is realised as the subject
- The agent occurs twice as often in the recipient passive as in the regular passive
- For each voice, the subject precedes the objects most often

Examining the overt traces, the object theme in recipient passives often occurs more frequent than in the regular passive.