Information theoretic historical morphology: A case study of High German adjectives

Rory Turnbull, Rachel Steindel Burdin, David M. Howcroft, Cynthia A. Johnson

The Ohio State University

{turnbull, burdin, howcroft, caj}@ling.osu.edu

Goals

To use the tools from information theory to:
- Assess claims about New High German adjectival system
- Provide a methodology to quantify simplification and complexity in language change

Background on High German adjectives

Middle High German (MHG)
Two paradigms:
- Weak, used with definite articles
- Strong, used with indefinite articles, or no article

New High German (NHG)
Three paradigms:
- Weak, used with definite articles
- Mixed, used with indefinite articles
- Strong, used when no article present

- NHG, Durrell (2002:118): Trade off between “fuller” strong endings and clarity of determiners

Information theory

- Branch of mathematics concerned with representation and transmission of signals (Shannon 1948); applied in studies of linguistic complexity.
- Entropy is a measure of uncertainty based on the frequency of signals and is measured in bits.

Key for H(A/B)

- Higher entropy → Less informative
- Lower entropy → More informative

<table>
<thead>
<tr>
<th>Paradigm</th>
<th>MHG</th>
<th>NHG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entropy (bits)</td>
<td>0.694</td>
<td>0.966</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Paradigm</th>
<th>MHG</th>
<th>NHG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Entropy</td>
<td>0.00</td>
<td>0.580</td>
</tr>
</tbody>
</table>

Table 1: Informativeness of B w.r.t. A

Methodology

Calculations
- Measure the conditional entropy of the adjective form given the article inflection (i.e., \(H(\text{Adj}|\text{Art})\)) in both MHG and NHG.
- Measure the conditional entropy of the article inflection given the adjective form (i.e., \(H(\text{Art}|\text{Adj})\)) in both MHG and NHG.

Synchronic Predictions

- Strong: \(H(\text{Adj}|\text{Art}) > H(\text{Art}|\text{Adj})\)
- Weak: \(H(\text{Adj}|\text{Art}) < H(\text{Art}|\text{Adj})\)

<table>
<thead>
<tr>
<th>Paradigm</th>
<th>MHG</th>
<th>NHG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entropy (bits)</td>
<td>0.756</td>
<td>0.978</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Paradigm</th>
<th>MHG</th>
<th>NHG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Entropy</td>
<td>0.240</td>
<td>0.978</td>
</tr>
</tbody>
</table>

Table 2: Relations between the entropy calculations suggested by Durrell’s (2002) claims

Simulations
- Generate alternative mixed paradigms to compare the attested systems to an alternative tripartite system
- Generate possible single paradigm systems to compare the attested systems to an alternative system where the paradigms collapsed

Table 3: Entropy of MHG and NHG adjective paradigms

<table>
<thead>
<tr>
<th>Paradigm</th>
<th>MHG</th>
<th>NHG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entropy (bits)</td>
<td>0.694</td>
<td>0.966</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Paradigm</th>
<th>MHG</th>
<th>NHG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Entropy</td>
<td>0.00</td>
<td>0.580</td>
</tr>
</tbody>
</table>

Table 4: Entropy of MHG and NHG adjective paradigms

Q1: Does the adjective compensate for a lack of information on the article?
- In MHG, the article is more informative about the adjective than vice versa.
- In NHG, strong adjective given the article is more informative than article given the strong adjective, and vice versa for weak adjective.
- This result supports Durrell’s (2002) claims.

Q2: How does the MHG system compare to the NHG system?
- Overall relative entropy for: MHG 0.658, NHG 0.537
- In MHG, the adjective is more informative about the article than the article is about the adjective.
- In NHG, the article is more informative about the adjective than the adjective is about the article.
- This result supports Durrell’s (2002) claims.

Q3: How does NHG compare to systems with alternative mixed paradigms?
- Intuitive claims are correct that NHG strong forms compensate for a lack of information on the article.
- NHG system is ‘simpler’, but unclear what this means in the language as a whole.
- NHG system has relatively low entropy compared to other possible tripartite systems.
- Attested tripartite system has entropy comparable to a collapsed system with only one paradigm for adjectives.
- Relative entropy decreased from MHG to NHG, suggesting that the NHG system is easier for listeners to process.

Figure 1: Relative conditional entropy for distinct paradigms in MHG and NHG

Figure 2: Varying the mixed paradigm in a tripartite system

NHG system at lower, but not lowest, end of the possible entropy space.

Discussion & Conclusions

- Language change can move in different directions.
- Other languages descended from Middle High German have a single paradigm (cf. Yiddish, Luxembourgish).
- Intuitively, one might assume a single paradigm is “simpler”.

Simulating a single paradigm

- Attested NHG system has relative entropy comparable to a system with a single paradigm.
- Some possible single paradigm systems have higher entropy than the NHG system.

Acknowledgements

We are grateful to Brian Joseph, Andrea Sims, and the MMPR group at OSU for helpful discussion and comments.

Works Cited
