From verbal prefixes to aspectual derivations in early Slavic

Dag Haug

University of Oslo

Formal Diachronic Semantics
Konstanz, 13 September 2016
Introduction

- Eckhoff and Haug (2015) analyze the rise of the derivation-based system of aspect in early Old Church Slavonic (OCS)
- The bulk of the paper is devoted to a synchronic analysis of the OCS aspectual system
- We also offered some thoughts on how this system arose in terms of “grammaticalization of telicity”
- The latter part was sketchy and very simplified, which is why I want to revisit it here.
Overview

1. Synchronic OCS

2. Diachrony – previous accounts

3. A new account
Most modern Slavic languages express aspect with **derivational** morphology: prefixes and suffixes

There is evidence of this system in OCS already

However, there is also a **inflectional** distinction between the aorist and the imperfect in the past tense
Old Church Slavonic verbs

<table>
<thead>
<tr>
<th>Stem</th>
<th>Present</th>
<th>Aorist</th>
<th>Imperfect</th>
<th>Infinitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>tvori-(ipfv?)</td>
<td>tvoritъ</td>
<td>tvori</td>
<td>tvorjaašе</td>
<td>tvoriti</td>
</tr>
<tr>
<td>sъ-tvori-(pfv)</td>
<td>sъ-tvoritъ</td>
<td>sъ-tvori</td>
<td>*sъ-tvorjaašе</td>
<td>sъ-tvoriti</td>
</tr>
<tr>
<td>sъ-tvarja-(ipfv)</td>
<td>sъ-tvarjatъ</td>
<td>*sъ-tvarja</td>
<td>sъ-tvarjaašе</td>
<td>*sъ-tvarjati</td>
</tr>
</tbody>
</table>

- Different stem formations (derivations) seem to be associated with aspectual values
- There is an inflectional exponent of ‘aspect’ in the past tenses (and in the participles)
- The present and the infinitive do not express ‘aspect’ in the inflection, but they do have the derivational distinction
Eckhoff and Haug (2015)

The questions we asked:
- What are the meanings of the stem formations?
- What are the meanings of the aorist and imperfect past tenses?

To answer those we used evidence from the Greek-OCS parallel corpora PROIEL and TOROT

These contain the New Testament in the Greek original and in two manuscripts of the OCS translation (Marianus and Zographensis)

The idea was to look at translation correspondences because
- Greek aspect is relatively well understood (a large traditional literature, and formal analysis in Bary 2009, Bary and Egg 2012)
- There are no transparent similarities between Greek and OCS aspectual morphology (despite etymological connections)
### OCS translations of Greek aorists and imperfects

<table>
<thead>
<tr>
<th></th>
<th>Sl. aorist</th>
<th>%</th>
<th>Sl. imperfect</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>aorist</td>
<td>2887</td>
<td>98.6</td>
<td>42</td>
<td>1.4</td>
</tr>
<tr>
<td>imperfect</td>
<td>79</td>
<td>11.1</td>
<td>631</td>
<td>88.9</td>
</tr>
</tbody>
</table>

**Table:** Marianus: n=3639

<table>
<thead>
<tr>
<th></th>
<th>Sl. aorist</th>
<th>%</th>
<th>Sl. imperfect</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>aorist</td>
<td>2604</td>
<td>98.2</td>
<td>47</td>
<td>1.8</td>
</tr>
<tr>
<td>imperfect</td>
<td>73</td>
<td>11.0</td>
<td>592</td>
<td>89.0</td>
</tr>
</tbody>
</table>

**Table:** Zographensis: n=3316
Inflectional and derivational aspect

- The numbers strongly suggest a functional similarity between the aor./imperf. distinctions in Greek and OCS
- This raises the question of what the role of the seemingly aspectual, derivational morphology is
  - What do the prefixes mean?
  - What do the suffixes mean?
  - How does affixation relate to inflectional aspect?
What role for the affixes?

<table>
<thead>
<tr>
<th>affix</th>
<th>aorist</th>
<th>imperfect</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>1143</td>
<td>224</td>
</tr>
<tr>
<td>prefix</td>
<td>1399</td>
<td>6</td>
</tr>
<tr>
<td>suffix</td>
<td>851</td>
<td>1388</td>
</tr>
<tr>
<td>both</td>
<td>1679</td>
<td>499</td>
</tr>
</tbody>
</table>

- The meaning of prefixation (without suffixation) must be incompatible with imperfective aspect, *unlike* in Greek.
- We concluded that prefixation is already grammaticalized in the meaning of perfective aspect, but this may have been too quick.
- Suffixation increases the frequency of the imperfect, but there is no general incompatibility.
<table>
<thead>
<tr>
<th>Suffix</th>
<th>All Verbs</th>
<th>Unprefixed Verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Aorist</td>
<td>Imperfect</td>
</tr>
<tr>
<td>–</td>
<td>4825</td>
<td>450</td>
</tr>
<tr>
<td>ě</td>
<td>563</td>
<td>555</td>
</tr>
<tr>
<td>ja</td>
<td>305</td>
<td>392</td>
</tr>
<tr>
<td>ěj</td>
<td>104</td>
<td>14</td>
</tr>
<tr>
<td>jaj</td>
<td>412</td>
<td>245</td>
</tr>
<tr>
<td>a</td>
<td>1185</td>
<td>1068</td>
</tr>
<tr>
<td>aj</td>
<td>98</td>
<td>251</td>
</tr>
<tr>
<td>a/aj</td>
<td>72</td>
<td>93</td>
</tr>
<tr>
<td>(V)va</td>
<td>2</td>
<td>63</td>
</tr>
<tr>
<td>i</td>
<td>1731</td>
<td>826</td>
</tr>
<tr>
<td>nɔ</td>
<td>197</td>
<td>0</td>
</tr>
<tr>
<td>ova</td>
<td>128</td>
<td>128</td>
</tr>
</tbody>
</table>

- The affixation pattern does not by itself reveal aspectual markedness
- But we can identify pairs by looking at the past tense behaviour
Pairing up verbs

- **Group 1**: a prefixed verb partnered by an unprefixed verb: *na-uč-i-ti*, *uč-i-ti* ‘teach’
- **Group 2**: an unprefixed verb partnered by an unprefixed derived (suffixed) verb: (*pust-i-ti*, *pušt-a-ti* ‘let go’)
- **Group 3**: a prefixed verb partnered by a derived prefixed and suffixed partner (typically with vowel lengthening in the root): *o-stav-i-ti*, *o-stavl-ja-ti* ‘leave’

- In group 1, the prefixed verbs typically takes only the aorist, whereas the simplex one can take both aspects
- In group 2, the simplex verb typically takes both aspects and the suffixed one only the imperfect
- In group 3, both verbs take only a single aspect
Verb classes - group 1

Distribution of verbs and aspect in translations of σπείρω
Verb classes - group 2

Distribution of verbs and aspect in translations of βαπτίζω

<table>
<thead>
<tr>
<th>Freq</th>
<th>Pres. (Gk. impf.)</th>
<th>Pres. (Gk. aor.)</th>
<th>Impf./pres.ptcp (Gk. impf.)</th>
<th>Impf./pres.ptcp. (Gk. aor.)</th>
<th>Other (Gk. impf.)</th>
<th>Other (Gk. aor.)</th>
<th>Aor./past (Gk. impf.)</th>
<th>Aor./past (Gk. aor.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Verb classes - group 3

Distribution of verbs and aspect in translations of ἐπερωτάω

- prašati
- prositi
- glagolati
- prošati
- Pres. (Gk. impf.)
- Pres. (Gk. aor.)
- Impf./pres.ptcp (Gk. impf.)
- Impf./pres.ptcp. (Gk. aor.)
- Other (Gk. impf.)
- Other (Gk. aor.)
- Aor./past (Gk. impf.)
- Aor./past (Gk. aor.)
Synchronic OCS

Diachrony – previous accounts

A new account

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>tvoriti, svtoriti ‘do’</td>
<td>priimati, prijeti ‘take’</td>
<td>dajatai, dati ‘give’</td>
</tr>
<tr>
<td>znati, poznati ‘recognize’</td>
<td>pristopati, pristopiti ‘arrive’</td>
<td>padati, pasti ‘fall’</td>
</tr>
<tr>
<td>učiti, naučiti ‘teach’</td>
<td>obrětati, obrěsti ‘find’</td>
<td>krštati, krštiti ‘baptize’</td>
</tr>
<tr>
<td>jasti, sëněsti ‘eat’</td>
<td>ostavljati, ostaviti ‘leave’</td>
<td>puštati, pustiti ‘divorce’</td>
</tr>
<tr>
<td>písati, napísati ‘write’</td>
<td>výprašati, výprositi ‘ask’</td>
<td>kupovati, kupiti ‘buy’</td>
</tr>
<tr>
<td>rasti, včzdrašti ‘grow’</td>
<td>ubivati, ubiti ‘kill’</td>
<td>saždati, saditi ‘plant’</td>
</tr>
<tr>
<td>klěti, proklěti ‘curse’</td>
<td>razumevati, razuměti ‘understand’</td>
<td>sveštati, sveštiti ‘sanctify’</td>
</tr>
<tr>
<td>alškati, vzalškati ‘fast’</td>
<td>sǐpasati, sǐpasti</td>
<td>plvati, plinžti ‘spit’</td>
</tr>
<tr>
<td>dělati, sđědati ‘work’</td>
<td>icěljati, icěliti ‘cure’</td>
<td></td>
</tr>
<tr>
<td>sě(ja)ti, vřsě(ja)ti ‘sow’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>slyšati, uslyšati ‘hear’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Overview

1. Synchronic OCS

2. Diachrony – previous accounts

3. A new account
Pragmatic strengthening

- ‘Default aspect’:
  - telic predicates → perfective interpretation
  - atelic predicates → imperfective interpretation

- This is a common inference in many languages (event realization, Bohnemeyer and Swift 2004)

- Pragmatic strengthening: → entailment
  - sn-ĕsti ‘eat’, o-stav-i-ti ‘leave’ → perfective
  - jasti ‘eat’ → imperfective (?)
  - o-stavl-ja-ti ‘leave’ → imperfective
Problems

- *jasti* ‘eat’ with a quantized theme should actually come out as telic and hence perfective but it doesn’t (raised by Grønn 2007, p. 4)

- What does it mean to say that *sən-ěsti* is a telic *jasti*, especially since the latter can also occur with a quantized theme?

- Preverbs “expressing telicity” are found throughout IE but only in Slavic are they grammaticalized as perfective markers
  - We partly addressed this arguing that what’s special in Slavic is that the prefixes get superimposed on a pre-existing aspectual system
  - But that’s also the case in e.g. Latin
  - Moreover, Latin also restricts the combination of prefixation and perfective aspect (Haug 2005)
Bidirectional OT (Grønn 2007)

<table>
<thead>
<tr>
<th>$f_1 &gt; f_2$; $m_1 &gt; m_2$</th>
<th>incomplete events ($m_1$)</th>
<th>complete events ($m_2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>simplex ($f_1$)</td>
<td>√ (optimal)</td>
<td>*(blocked)</td>
</tr>
<tr>
<td>prefixed ($f_2$)</td>
<td>*(blocked)</td>
<td>√ (weakly optimal)</td>
</tr>
</tbody>
</table>

1. incomplete events unmarked meaning (across verbal situations)
2. generalized to $f_1$ even with a quantized object
3. $\langle f_1, m_1 \rangle$ strengthened by associative learning
4. $\langle f_2, m_2 \rangle$ emerges as weakly optimal
5. $\langle f_2, m_2 \rangle$ strengthened
6. secondary imperfectivization fills morphological gaps
Problems

- No account of the interaction with inflectional aspect
- Hard to see why $m_1$ would ever be the unmarked meaning for VPs with quantized objects
- Incorrect chronology:
  - assumes $m_1$ becomes central meaning of $f_1$ before $f_2$ associates with $m_2$
- Again, no account of the difference between $s\ddot{a}n\ddot{-}\ddot{e}sti$ and $jasti$ with a quantized object
Strong points

- Grønn is right to start from group 1 (prefixed verbs with unprefixed partners)
- The spread of secondary suffixation takes off in the 16th and 17th centuries, implying that what we see in OCS is the beginning of an S-curve (Andersen 2009)
- So chances are that the origin of aspect lies in group 1 and that secondary suffixation is a response to the grammaticalization of aspect
Overview

1. Synchronic OCS

2. Diachrony – previous accounts

3. A new account
A new perspective

- Previous accounts are vague on the meaning of the source construction (e.g. *sjan-ěsti*) except that it is somehow related to telicity

- In general there is a large philological literature on the use of ‘telicizing prefixes’ in Old IE, but no account of their meaning

- A tempting candidate is Bary and Egg’s operator MAX (≈ AOR from Krifka 1989)

- For Bary and Egg, MAX is a coercion operator triggered by perfective aspect in Greek

- So MAX is not perfective aspect, but a *component* of it, which would seem like a suitable starting point

- Ultimately this does not work, but it fails for interesting reasons
Bary and Egg 2012

- Bary and Egg analyze Greek aspects as relations between event times and topic times
  
  \[
  \text{PFV: } \lambda P \lambda t \exists e. P(e) \land \tau(e) \subseteq t
  \]
  
  \[
  \text{IPFV: } \lambda P \lambda t \exists e. P(e) \land t \subseteq \tau(e)
  \]

- In addition, there are aspectual selection restrictions
  
  \[
  \text{PFV requires a bounded predicate}
  \]
  
  \[
  \text{IPFV requires an unbounded predicate}
  \]

- These restrictions can be satisfied through various coercion operators, including MAX
Consider an unbounded predicate like *bouleuein* ‘be a senator’

- \[ IPFV(bouleuein) = \lambda t. \exists \ e. be.a.senator(e) \land t \subseteq \tau(e) \]
- \[ PFV(boulein) = \lambda t. \exists \ e. be.a.senator(e) \land \tau(e) \subseteq t \]

Because *be.a.senator* has the subinterval property we predict contrary to facts that PFV(bouleuein) could give us a time containing only a part of some maximal eventuality of being a senator.

MAX maps a predicate \( P \) on the set of locally maximal, convex (uninterrupted) eventualities.

- \[ MAX(P)(e) \text{ iff } P(e) \land CONV(e) \land \forall e'. e \sqsubseteq e' \rightarrow \neg P(e') \]
- \[ PFV(MAX(bouleuein))) \text{ now gives us the right readings} \]
prefixation \equiv \text{MAX}?

- What if we took \textit{s\n-\v{e}sti} to denote a set of maximal eating events?
- \[
\begin{align*}
\llbracket s\n-\v{e}sti \rrbracket &= \llbracket \text{MAX} \rrbracket (\llbracket jasti \rrbracket) = \\
\lambda e. \text{eat}(e) \land \text{CONV}(e) \land \forall e'. e \sqsubseteq e' \rightarrow \neg \text{eat}(e')
\end{align*}
\]
- \textit{s\n-\v{e}sti} would denote the set of all maximal, convex eating events
- While a good fit for the Greek aorist, this notion of maximality turns out to be too permissive for OCS prefixation
Prefixed aorist with local maximality

(1) a. pros tên sklêrokardian humôn egrapsen because of your hardness of heart write.3.SG.AOR humin tên entolên tautên you.2PL.DAT this commandment.acc

b. po žestosrđiju vašemu napisa because of your hardness of heart na-write.3.SG.AOR vamъ zapověď sijo you.2PL.DAT this commandment.acc ‘He wrote you this commandment because of your hardness of heart.’ (Mark 10.5)
Unprefixed imperfect without local maximality

(2) a. ho de Iêsous katô kupsas the.NOM but Jesus.NOM down bend.AOR.PTCP tōi daktulôi egraphen eis tên gên the.finger.DAT write.3SG.IPFV on the ground

b. iĉь že nizь poklonь sę the.NOM but Jesus.NOM down bend.PST.PTCP prvstomь pisaašе na zemi finger.DAT 0-write.3SG.IPFV on the ground ‘But Jesus bent down and was writing with his finger on the ground. (When they kept on questioning him, he straightened up and said.) (John 8.6)
Unprefixed aorist with local maximality

(3) a. peri gar emou ekeinos egrapsen
   about for I.GEN he.SG.M.NOM write.3SG.AOR

   b. o mënë bo tь pīsa
   about I.LOC for he.NOM ∅-write.3SG.AOR

   ‘(If you believed Moses, you would believe me.) For he wrote
   about me.’ (John 5.46)

- Here we have a complete, but extendable event
  - The maximal event of Moses writing about Jesus lies within the (past)
    reference time
  - The aorist is sensitive to this maximality, but the prefix requires
    something stronger
- So we typically get unprefixed aorists in OCS (but imperfective in
  MRus.!)
David and the consecrated bread v1

(4) a. kai tous artous tês prothéseôs
and the.ACC breads.ACC the.GEN presentation.GEN
ephagen
eat.3.SG.AOR

b. i xleby prědĕloženíě sĕnĕstĕ
and breads.PL.M.ACC presentation.GEN sĕ-eat.3.SG.AOR
(How David went into the house of God) and ate the showbreads. (Mark 2.26)
David and the consecrated bread v2

(5) a. kai tous artous tês prothéseōs elaben kai and the showbreads.ACC take.3.SG.AOR and ephagen kai edôken kai tois met’ autou eat.3.SG.AOR and give.3.SG.AOR also those with him.DAT

b. i xleby prěděložení přiję i and showbreads.ACC pri-take.3.SG.AOR and estъ i dastъ i ř-eat.3.SG.AOR and ŕ-give.3.SG.AOR also sôštiimь sь nimь

beimg.with.him.DAT

‘(How David entered the house of God), took the showbreads, ate and gave to those with him.’ (Luke 6.4)
Second try

- Prefixation seems to require not only local maximality but global maximality across worlds: the event cannot be extended further.
- Similar observations have been made on verbal prefixes in other languages (incl. Tatevosov on Russian), and about *finish*.
- Piñon: *weak* and *strong* accomplishments, where the latter have two presuppositions:
  - There was a preceding event with the same agent and the same incremental theme such that the theme was partly affected.
  - An event of the type determined by the predicate and the incremental theme QC cannot be continued.
- The first presupposition is optional, as has also been observed for German prefixed verbs (Engelberg 2002, Rossdeutscher 2011).
- For now we focus on the second presupposition.
A standard accomplishment

**a commandment:** \( \lambda R. \lambda e. [y \mid \text{commandment}(y)]; R(e, y) \)

**write:** \( \lambda y. \lambda x. \lambda e. [\mid \text{write}(e, x, y)] \)

**write\(^{TR}\):** \( \lambda Q. \lambda x. \lambda e. Q(e, \lambda y. \lambda e. [\mid \text{write}(e, x, y)] \)

**write\(^{TR}(a \text{ commandment}):** \( \lambda x. \lambda e. \text{commandment}(y) \)
\( \text{write}(e, x, y) \)
Global maximality

- In all worlds, all events of the type defined by the predicate and the incremental theme generalized quantifier have no superparts

- We abbreviate \( \text{MAX}(T, Q) := \)
  
  \[
  \begin{align*}
  & [e'x'] \land \forall y. \lambda e[ |T(e, x', y)| ] \rightarrow \\
  & \neg[e'' | e' \prec e''] \land \forall y. \lambda e[ |T(e, x', y)| ]
  \end{align*}
  \]

\[
\text{MAX(\text{write, a commandment})}
\]

\[
\begin{align*}
& [e'x'y'] \land \text{commandment}(y'), \text{write}(e', x', y') \rightarrow \\
& \neg[e''y'' | e' \prec e'', \text{commandment}(y''), \text{write}(e' \oplus e'', x', y'')]\]
\]
We will take \( \text{MAX}(\text{write}, Q) \) to be a presupposition of \( \text{na-pisati} \).

This amounts to a selectional restriction on the theme \( Q \) (viz. it must be quantized).

Such restrictions are informative in languages like Slavic (and Latin) with no articles.

\[
\lambda Q. \lambda x. \lambda e. Q(e, \lambda y. \lambda e. [\text{write}(e, x, y), \partial(\text{MAX}(\text{write}, Q))] )
\]
Synchronic OCS – previous accounts

A new account

commandment(y), write(e, x, y)

\[
\begin{array}{c}
\text{commandment}(y') \\
\text{write}(e', x', y')
\end{array}
\]

\[\partial
\begin{pmatrix}
e' x' y' \\
\end{pmatrix}
\]
Does this help?

- With a careful analysis of the source semantics, we are better equipped to understand what is going on.
- In particular we look at the combination of a prefixed (and unsuffixed) verb with imperfective aspect, which should be ruled out.
- Since the data show that the Slavic imperfect is close to the Greek one, we adopt Bary (2009)'s semantics (based on Dowty 1979).
Synchronic OCS

Diachrony – previous accounts

A new account

\[ x \ t_{TT} \]

\[ \square \textit{inert} \]

commandment(\(y\)), write(\(e, x, y\)), \(t_{TT} \subseteq \tau(e)\)

\[ \partial \]

commandment(\(y'\))

write(\(e', x', y'\))

\[ e'', y'' \]

\(e' \prec e''\)

commandment(\(y''\))

write(\(e' \oplus e'', x', y''\))
The presupposition is irrelevant for the actual world: the P-event could not be continued in any world, but the imperfective does not assert the existence of a P-event in the first place.

Infelicitous for similar reasons as *The sea was rising ten feet when...* or *John was drinking three cups of tea when...* (Mittwoch 1988)
The infelicity of a strong accomplishment with the imperfective follows from Gricean reasoning.

This is good since we observe the same effect in e.g. Latin, which also has prefixed with no lexical content: *conficio*, *comedo*

This suggests a grammaticalization path from lexical prefix to strong accomplishment marker.

Incompatibility with the imperfect arises already at the strong accomplishment stage.

So Eckhoff and Haug may have been wrong in claiming that prefixation is already grammaticalized as perfective aspect in OCS.
How *did* it change then?

- Since the verbal prefix *qua* marker of strong accomplishment is incompatible with imperfective aspect, speakers are free to reanalyze it *without* visible results
  - \( \text{na-pisa-∅} = \text{AOR}(\text{write}(e, x, Q) \land \partial(\text{MAX}(\text{pisa}, Q))) \)
  - \( \text{∅-pisa-aše} = \text{IPFV}(\text{write}(e, x, Q)) \)

- This makes it possible to express aspect outside the past tense (e.g. in infinitives)

- The crucial event now becomes the *loss* of the imperfect/aorist distinction and its replacement with the *I*-form

---

Dag Haug

Early Slavic aspect

FoDS 13 September 2016
Infinitives

Infinitives and their Greek originals

Freq

0
50
100
150
200

~ipfv -ipfv -ipfv(<4) ~neut ~pfv -pfv -pfv(<4) -unkn.
Consequences of the morphology loss

<table>
<thead>
<tr>
<th>Pre-loss</th>
<th>Post-loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>pisaaše ±QUA</td>
<td>pisal ±QUA</td>
</tr>
<tr>
<td>pisa ±QUA</td>
<td>pisa ±QUA</td>
</tr>
<tr>
<td>napisa +QUA</td>
<td>napisal +QUA</td>
</tr>
</tbody>
</table>

- The loss of the morphology leads to an increase in the scope of imperfective aspect, which makes a number of correct predictions about the later Slavic system.
  - Intimate connection between perfectivity and telicity (Dahl 1985)
  - A very general temporal relation of overlap for the imperfect (Grønn 2003): $\tau(e) \subseteq t \lor \tau(e) \supseteq t \rightarrow \tau(e) \circ t$
  - The rise of the imperfective general-factual, limited to class 1 in Old Russian until the 16th century (Kukuškina & Ševeleva 1991)
  - Incompatibility of the perfective aspect with durational adverbs
  - Some of these effects are partly blurred by the (largely post-OCS) development of specialized prefixes po- and pro- as well as the spread of secondary imperfectivization.
What about the first presupposition?

- Strong accomplishments come with an optional presupposition that the event has already started.
- In fact, it has been proposed that Modern Russian perfectives have exactly this presupposition, to account for their behaviour under negation (Forsyth 1970, Padučeva 1996, contra Grønn 2003:60ff.).
- Whatever the merits of this analysis for Modern Russian, there’s actually no evidence that negation affects OCS aspect in ways that it does not affect Greek.
- It has also been suggested that the use of the imperfective aspect in negated imperatives stems from a desire to prohibit the entire event (Boguslawski 1985, Levinson 2005, see also Partee).
- But again, OCS shows no preference for imperfective aspect in negated imperatives.
Conclusions

- We have identified one grammaticalization path from lexical prefix to strong accomplishment marker.
- This gives us lexical telicity, which is a better source for perfective aspect than VP-level aspect.
- The subsequent change to aspect marker is due to classical reanalysis + subsequent loss that makes the reanalysis visible.
- The large remaining question is how this was generalized to larger classes of verbs with the rise of secondary imperfectivization.
Some observations on the semantic shift of the clause-embedding predicate żałować 'regret' in Polish

Łukasz Jędrzejowski (Universität Potsdam, Germany)
lukasz-jedrzejowski@daad-alumni.de

1. Introduction

In Modern Polish (1900 - ) żałować can be employed in two different ways. On the one hand, it can be used as a factive predicate in the sense claimed by Kiparsky & Kiparsky (1971) and paraphrased as 'regret' (= żałować1):

[1] Żałuję, że tak późno zacząłem strzelać gole
żałować.1SG that so late start.l-PTCP.1SG.M shoot.INF goals
'I regret that I started scoring goals so late'
(NKJP, Express Ilustrowany, 29/4/2002)

On the other hand, it can also mean 'begrudge' (= żałować2):

[2] Żałujesz mi iść na urlop?
żałować.2SG me.DAT go.INF on vacation
'Do you begrudge me a vacation?'

Żałować is an imperfective verb. Its perfective counterpart, pożałować, can also be used in the sense of żałować1 and żałować2:

[3] Pożałujesz, że mnie zatrzymałeś
pożałować.2SG that me.ACC arrest.l-PTCP.M.2SG
'You will regret it that you arrested me'
(NKJP, Gazeta Poznańska, 28/6/2003)

money.GEN NEG pożałować.l-PTCP.SG.M also president Pałasz
'The major Pałasz did not begrudge money either'
(NKJP, Gazeta Poznańska, 12/4/2003)

In this talk, I will focus only on żałować.

1 Following abbreviations are used in this paper: 1/2/3 - 1st/2nd/3rd person, ACC - accusative, AUX - auxiliary verb, CL - clitic, COR - correlate, DAT - dative, GEN - genitive, HAB - habitual, INF - infinitive, LOC - locative, l-PTCP - l-participle (inflected for number and gender), M - masculine, N - neuter, NEG - negation, PL - plural, PPR - present participle, PRF - perfective, REFL - reflexive, SG - singular, VIR - virile.
Leading questions

✓ To what extent does żałować1 differ from żałować2?
✓ What are the c-selectional properties of both predicates?
✓ What are the emergence conditions of żałować2?
✓ Does the development of żałować2 exemplify a grammaticalization process?
✓ What role does the opposition factive vs. non-factive play?

Outline of the talk

Section 1: Introduction
Section 2: Żałować in Modern Polish
✓ Argument structure
✓ Embedded complements
✓ Factivity
Section 3: Żałować in the history of Polish
✓ From Old Polish (-1500) to Modern Polish (1900-)
✓ Semantic change
Section 4: Theoretical perspective: Reanalysis
Section 5: Concluding remarks

2. Żałować in Modern Polish

In this section, I will point out selected differences between żałować1 and żałować2. The main focus will be on: (i) argument structure, (ii) embedded complements and (iii) the opposition factive vs. non-factivity.

2.1. Argument structure

Żałować1 is a two-place transitive predicate. Its internal argument receives Genitive case:

'Sometimes I regret my superfluous words' (NKJP, Dziennik Zachodni, 15/4/2005)

The embedded proposition can also refer anaphorically to the matrix correlate tego:

'I regret very much what I did' (NKJP, Dziennik Zachodni, 25/9/2009)
Żałować₂, in turn, is a three-place ditransitive predicate. Similar to żałować₁, it licenses a direct object checking the Genitive case. Additionally, an indirect object is generated and marked for the Dative case:

\[
[7] \quad [\text{DP pro}] \quad \text{Żałować} \quad [\text{DP mi}]_{\text{DAT}} \quad [\text{DP cukierków}]_{\text{GEN}} \quad [\text{żałować₂}]
\]

'begrudge.1-PTCP.3SG.M me sweets'

(\textit{NKJP, Talki w wielkim mieście, 2002})

If the Dative-DP is realized as an implicit argument, two readings of żałować appear to be appropriate:

\[
[8] \quad [\text{pro}] \quad \text{Żałuję} \quad [\_]_{\text{i}} \quad \text{kupna} \quad \text{tego mieszkania}
\]

'İ regret that I bought this flat' [żałować₁]

'İ begrudge someone to buy this flat' [żałować₂]

However, if the Dative-DP is realized overtly, it disambiguates the reading of the verb:

\[
[8'] \quad [\text{pro}] \quad \text{Żałuję} \quad [\_i]_{\text{you.DAT}} \quad \text{kupna} \quad \text{tego mieszkania}
\]

'*I regret that I bought this flat' [żałować₁]

'I begrudge you to buy this flat' [żałować₂]

2.2. Embedded complements

2.2.1. DPs

Both żałować₁ and żałować₂ can embed DPs marked for the Genitive case:

\[
[9] \quad \text{Nie} \quad \text{żaluje} \quad [\text{DP swojej decyzji}] \quad [\text{żałować₁}]
\]

'He doesn't regret his decision'

(\textit{NKJP, Mazowieckie To i Owo, 7/8/2008})

\[
[10] \quad \text{Nie} \quad \text{żałujemy} \quad [\text{DP urlop-u}] \quad \text{doktor-owi Szczypułu-DAT} \quad [\text{żałować₂}]
\]

'We do not begrudge Doctor Szczypuła a vacation'

(\textit{NKJP, Dziennik Polski, 23/5/2002})
2.2.2. Finite CPs

As far as sentential complements are concerned, żatować1 can embed finite CPs headed by the complementizer że 'that'. There are no restrictions as to the aspect of the embedded verb:

[11] Żaluję, [CP że nie jem owoców] [imperfective]
żałować.1SG that NEG eat.1SG.IMPRF fruits
'I regret that I don't eat fruits'

[12] Żaluję, [CP że sprzedaliśmy dom] [perfective]
żałować.1SG that sell.1PL.l-PTCP.PRF house
'I regret that we sold the house'

[13] Żaluję, [CP że obiecywałem ci nowy dom] [iterative]
żałować.1SG that promise.1SG.l-PTCP.ITER you.DAT new house
'I regret that I used to promise you a new house'

The correlate tego can be used in all three cases too:

[11'] Żaluję, tego_i [CP że nie jem owoców]_i
żałować.1SG COR.GEN that NEG eat.1SG.IMPRF fruits
'I regret it that I don't eat fruits'

[12'] Żaluję, tego_i [CP że sprzedaliśmy dom]_i
żałować.1SG COR.GEN that sell.1PL.l-PTCP.PRF house
'I regret it that we sold the house'

[13'] Żaluję, tego_i [CP że obiecywałem ci nowy dom]_i
żałować.1SG COR.GEN that promise.1SG.ITER.l-PTCP you.DAT new house
'I regret it that I used to promise you a new house'

Żatować2, on the other hand, is more restricted:

[14] Żałujesz mi [CP że idę na urlop?] [imperfective]
begrudge.2SG me.DAT that go.1SG.IMPRF on vacation
'Do you begrudge me a vacation?'

[15] *Żałujesz mi [CP że pójdę na urlop?] [perfective]
begrudge.2SG me.DAT that go.1SG.PRF on vacation
Intended: 'Do you begrudge me a vacation?'

[16] Żałujesz mi [CP że chodzę na urlop?] [iterative]
begrudge.2SG me.DAT that go.1SG.ITER on vacation
Intended: 'Do you begrudge me a vacation?'
However, if the correlate *tego* is used, the examples given in (14)-(16) sound natural:

\[14\] Żałujesz mi tego [\text{CP że idę na urlop?}] 
bebrudge.2SG me.DAT COR.GEN that go.1SG.IMPRF on vacation
'Do you begrudge me a vacation?'

\[15\] Żałujesz mi tego [\text{CP że pójdę na urlop?}] 
bebrudge.2SG me.DAT COR.GEN that go.1SG.PRF on vacation
'Do you begrudge me a vacation?'

\[16\] Żałujesz mi tego [\text{CP że chodzę na urlop?}] 
bebrudge.2SG me.DAT COR.GEN that go.1SG.ITER on vacation
Intended: 'Do you begrudge me a vacation?'

At this moment, I have no explanation for why this is so and where this difference comes from.

\subsection*{2.2.3. Infinitive CPs}

*Żałować*1 as a factive verb cannot embed infinitive clauses (cf. also Słodowicz 2008 for a recent general overview of clause-embedding predicates in Polish disallowing infinitive clauses):

\[17\] *Żałuję, \text{CP nie potrafić wysoko śpiewać} [żałować1] 
żałować.1SG NEG can.INF high sing.INF
Intended: 'I regret to be not able to sing high notes'

Even if one uses the correlate *tego*, the situation does not change:

\[17\] *Żałuję tego, \text{CP nie potrafić wysoko śpiewać} [żałować1] 
żałować.1SG COR.GEN NEG can.INF high sing.INF
Intended: 'I regret to be not able to sing high notes'

On the contrary, *żałować*2 can select for CP-infinitives. It is an object control verb:

\[18\] proi Żałujesz mi [\text{CP PROi/j iść na urlop?}] [żałować2] 
żałować.2SG me.DAT go.INF on vacation
'Do you begrudge me to go on vacation?'
There are no aspectual restrictions with respect to the embedded verb:

[19]  
    Żałujesz mi [cp iść/pójść/chodzić na urlop]?  
    żalować.2SG me.DAT go.INF.IMPRF/PRF/ITER on vacation  
    'Do you begrudge me to go on vacation?'

Notice, however, that the correlate *tego* is disallowed:

[19']  
    *Żałujesz mi tego [cp iść/pójść/chodzić na urlop]?  
    żalować.2SG me.DAT COR.GEN go.INF.IMPRF/PRF/ITER on vacation  
    'Do you begrudge me to go on vacation?'

Remarkably, if *żalować* embeds a non-finite CP complement, the matrix verb and the dependent clause can be modified by two distinct temporal adverbials. I take this to be one of the diagnostics for the CP-hood of embedded complements:

[20]  
    Jeszcze wczoraj żałował-eś mi  
    yet yesterday żałować.1-PTCP.3SG.M-AUX.CL.2SG me.DAT  
    [cp iść jutro na urlop]  
    go.INF tomorrow on vacation  
    ≈'Yesterday, you begrudge me to go on vacation tomorrow'

As *żalować* does not presuppose a factive interpretation of the embedded proposition, infinitive CPs are allowed to occur.

2.2.4. Conditional clauses

As pointed out by Williams (1974), Pullum (1987) and Pesetsky (1991), conditional clauses can be realized as one of the arguments of factive predicates and preference predicates:

[21]  
    You'll regret it [cp if you split up with him]

*Żałować* as a factive predicate allows conditional clauses as well:

[22]  
    Później żałuję, [cp jeśli to zrobić]  
    żałować.3PL if this do.3PL  
    'They will regret it later, if they do this'  
    (NKJP, an internet forum, 2/9/2007)
On the other hand, żałować2 is not compatible with conditional clauses:

[23] *Rodzice żalują mi, [cp jeśli to zrobić] żałować2
parents żalować.3PL me.DAT if this do.1SG

Based on the contrast between [22] and [23], I claim that żałować2 is neither an inherent factive predicate nor a preference predicate.

2.3. Factivity

Żałować1 is a factive predicate:

[24] Żaluje, że tak późno zacząłem strzelać gole [p=1]
żalować.1SG that so late start.1-PTCP.1SG.M shoot.INF goals
'I regret that I started scoring goals so late'
(NKJP, Express Ilustrowany, 29/4/2002)

The presence of a negation marker in the matrix clause has no impact on the truth value of the embedded proposition:

[25] Nie żaluje, że tak późno zacząłem strzelać gole [p=1]
NEG żalować.1SG that so late start.1-PTCP.1SG.M shoot.INF goals
'I don't regret that I started scoring goals so late'

Żałować2, in turn, is a future-oriented predicate, meaning that it remains open whether the embedded proposition is true or false. If a non-finite clause is embedded, the aspect has no impact on the interpretation:

[26] Żałujesz mi [cp iść/pójść/chodzić na urlop]? [p=0 v p=1]
żalować.2SG me.DAT go.INF.IMPRF/PRF/ITER on vacation
'Do you begrudge me to go on vacation?'

If żałować2 selects a finite CP with a verb marked for the present tense, only the iterative aspect gives rise to a factive interpretation:

[27] Żalujesz mi tego, że idę na urlop? [p=0 v p=1]
begrudge.2SG me.DAT COR.GEN that go.1SG.IMPRF on vacation
'Do you begrudge me a vacation?'

[28] Żalujesz mi tego, że pójdę na urlop? [p=0 v p=1]
begrudge.2SG me.DAT COR.GEN that go.1SG.PRF on vacation
'Do you begrudge me a vacation?'
Interestingly, if the embedded verb is marked for the past tense, it triggers a factive interpretation of the embedded proposition. The aspect value plays no role:

- [30] Żałujesz mi tego, że byłem na urlopie? \( [p=1] \)  
  begrudge.2SG me.DAT COR._GEN that be.1SG.IMPRF.PST on vacation  
  'Do you begrudge me a vacation?'

If żałować2 is under the scope of a negation marker, the truth value of \( p \) does not change:

- [33] Nie żałujesz mi tego, że byłem na urlopie? \( [p=1] \)  
  NEG begrudge.2SG me COR._GEN that be.1SG.IMPRF.PST on vacation

2.4. Interim summary

The following table gives an overview of selected differences between żałować1 and żałować2:

<table>
<thead>
<tr>
<th></th>
<th>żałować1</th>
<th>żałować2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Dativ-DPs</td>
<td>-</td>
</tr>
<tr>
<td>2.</td>
<td>Finite że-clauses</td>
<td>+</td>
</tr>
<tr>
<td>3.</td>
<td>Finite że-clauses with the correlate tego</td>
<td>+</td>
</tr>
<tr>
<td>4.</td>
<td>Conditional clauses</td>
<td>+</td>
</tr>
<tr>
<td>5.</td>
<td>CP-infinitives</td>
<td>-</td>
</tr>
<tr>
<td>6.</td>
<td>CP-infinitives with the correlate tego</td>
<td>-</td>
</tr>
<tr>
<td>7.</td>
<td>Factive interpretation of ( p )</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 1: Selected differences between żałować1 and żałować2
3. *Żałować* in the history of Polish

Based on Klemensiewicz (2009) I distinguish the following four language stages in the history of Polish:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Language Stage</th>
<th>Beginning</th>
<th>Ending</th>
</tr>
</thead>
<tbody>
<tr>
<td>OP</td>
<td>Old Polish</td>
<td>-</td>
<td>1535</td>
</tr>
<tr>
<td>MidP</td>
<td>Middle Polish</td>
<td>1535</td>
<td>1780</td>
</tr>
<tr>
<td>NP</td>
<td>New Polish</td>
<td>1780</td>
<td>1900</td>
</tr>
<tr>
<td>MP</td>
<td>Modern Polish</td>
<td>1900</td>
<td>present</td>
</tr>
</tbody>
</table>

Table 2: Language periods in the history of Polish

3.1. Old Polish ( - 1535)

*Żałować* occurs already in Old Polish and it is associated with three different meanings. Its factive use (= *żałować1*) occurs mainly with DPs and finite CPs:

[36] \( [_{DP \text{ swych} \ g<rz>echow}] \) (...) on był żałował  
    his sins.GEN  he be.l-PTCP.3SG.M  żałować.l-PTCP.3SG.M  
    'He regretted his sins'  
(PolDi, *GN*, Kazanie 4; 14r: 26-27)

[37] oni tego byli barzo żałowali  
they this be.l-PTCP.3PL.VIR very żałować.l-PTCP.3PL.VIR  
[CP izć oni swe [s]boże (...) byli ro/z/przedali]  
    that they their crop be.l-PTCP.3PL.VIR sell.l-PTCP.3PL.VIR  
    'They regretted very much that they sold their crop.'  
(Poldi, *KG*)

Remarkably, *żałować* can also mean 'sue', 'proceed against somebody', as shown in [38]:

[38] tedy Jan Piotra pozwał przed  
    then Jan.NOM Piotr.ACC sue.l-PTCP.3SG.M before  
    sąd żałując [CP izć ji uranil]  
    court żałować.PPR that him.ACC hurt.l-PTCP.3SG.M  
    'then Jan sued Piotr saying that he hurt him'  
If żałować occurs with the preposition na 'on', it means 'tell tales about sb. to sb.':

[39]  

\[ \text{<P>iotr żałował} [\text{pp na Jana}] \text{ przed sądem,} \]

Piotr.NOM żałować.l-PTCP.3SG.M on Jan.ACC before court.LOC

\[ \text{[cp kako] pcożyte abo dzienia jego kradmie} \]

how bees or clothes his steal.3SG

'Piotr told tales about Jan that he supposedly steals his bees and clothes'

(PolDi, Dzial, 22:10-11)

What appears to be striking about the latter two uses is that they fell into disuse. In MP żałować can only be employed either as żałować1 or żałować2. In this respect, MP differs from Modern Czech. As Radek Šimík pointed out to me, the Old Polish readings (except for żałować1) are available in Modern Czech.

I was not able to find any Old Polish examples illustrating the use of żałować2.

### 3.2. Middle Polish (1535 - 1780)

First instances of żałować2 occurs in Middle Polish. In all cases a DP is embedded:

[40]  

\[ \text{nie żałować} [\text{dp pieniędzy} \text{ boby to było nierychło} \]

NEG żałować.INF money.GEN because it be l-PTCP.3SG.N not.soon

'Don't begrudge the money; otherwise it (= a party) won't take place soon'

(PolDi, ListyDoMarysi)

Notice that żałować2 does not license any Dative-DP in [40]. Occasionally, we can attest examples in which żałować does generate a DP checking the Dative case. But in these cases żałować is used a speech verb meaning 'complain':

[41]  

\[ \text{Jedna pani przyszła przed wojta i żałowała} \]

one woman come.l-PTCP.3SG.F before voyt and żałować.l-PTCP.3SG.F

\[ \text{[dp mu], iż jej czynszownik zbiewał} \]

him.DAT that her.DAT owner deceive.l-PTCP.3SG.M

'A woman came to the village head and was complaining to him that the owner of the tenement house deceives her.' (PolDi, OrtMac)

### 3.3. New Polish (1780 - 1900)

In late Middle Polish/early New Polish żałować2 starts licensing Dative-DPs:
[42] czy byłby mu kawałka
whether him.DAT piece.GEN
chleba swojego żałował?
bread.GEN his żałować.ĺ-PTCP.3SG.M

'Would he begrudge his own piece of bread to him?'
(Poldi, Krasicki)

As I was not able to find any New Polish examples in which żałować2 would select an infinitive CP, I assume that they start to appear in Modern Polish.

4. Theoretical perspective: Reanalysis

In what follows, I analyze both żałować1 and żałować2 as lexical V-heads, indicating that none of these heads grammaticalized into a functional head associated with a functional projection:

![Figure 1: Base positions of żałować1 and żałować2](chart)

I argue that żałować2 developed out of żałować1 in the 17th century:

![Figure 2: The development of żałować2 and its complements](chart)
Following Larson (1988), I assume the indirect object receiving the Dative case to merge within vP.

The differences between żałować1 and żałować2, in turn, follow from the presence/absence of an [assertion] feature in ForceP of the subordinate clause (cf. Basse 2008). If żałować selects for a CP, the truth-value of $p$ can be either presupposed by the speaker (= żałować1) or asserted by the matrix subject (= żałować2). In the former case CPs are analyzed as defective phases lacking the feature [assertion]. Internally, there is no edge feature on the left periphery in the embedded clause and any kind of movement to the left edge is disallowed (based on Basse 2008):

Evidence for [42] comes from floating auxiliary clitics. In [45], a CP is embedded under żałować1 and the auxiliary clitic cannot move from PtcpP to a higher position within the CP-field. The movement is blocked due to the absence of the [assertion] feature:

If, on the hand, the feature [assertion] is activated, the C-Phase is not defective and the embedded C-head is an accessible goal for an Agree relation, which, in turn, is required both for PRO and secondary predicates in order to check their Case values in the embedded infinitive clause, e.g. the Dative in [46]:

In other words, although żałować underwent a semantic change and although its complement types have changed, the syntactic size of its complements remained the same.
5. Conclusion

In this talk, I examined two different uses of the predicate *żałować* in (the history of) Polish. It has been shown that *żałować* can be used either as an inherent factive predicate ('regret') or as a future-oriented predicate ('begrudge').

If *żałować* is employed as a factive predicate, it cannot license Dative-DPs and select for CP-infinitives. These restrictions do not hold for the non-factive use of *żałować*. If, on the other hand, it occurs as a future-oriented predicate meaning 'begrudge', it cannot occur with conditional clauses and finite CPs headed by the complementizer *że* 'that'.

Both predicates are lexical V-heads. The non-factive use developed out of the factive use of *żałować* in Middle Polish (1535 - 1780). This development does not instantiate any grammaticalization process. What both uses exemplify is a semantic shift of a single predicate entailing syntactic consequences.

Primary sources

NKJP - *Narodowy Korpus Języka Polskiego* ['National Corpus of Polish']
(http://nkjp.pl/)

PolDi - *A Polish Diachronic Online Corpus*
(http://rhssl1.uni-regensburg.de/SlavKo/korpus/poldi)

Secondary references


BARE AND INDEFINITE NOMINAL PREDICATES IN THE HISTORY OF GERMAN

Svetlana Petrova, Wuppertal
petrova@wuppertal.de

1. Subject and Goals

- striking mismatch in the use of the indefinite article with so-called ‘role nouns’ (nominals denoting professions, religious denominations, nationalities etc., see Zamparelli 2008) in copular constructions in modern standard German (1) and its historical prestages (3)-(2), see also Behaghel (1923, pp. 87-91):

(1) Er ist (*ein) Arzt / (*ein) Italiener / (*ein) Katholik
‘He is a physician / an Italian / a catholic’

(2) Dancwart der was marschalc
‘Dankwart, he was Marshal’1 (NibB 11,1)

(3) a. Kain was ein acchermann
‘Cain was a peasant’ (Genesis 6:11)

b. sîn vater was ein Franzois
‘his father was a Frenchman’ (Pz 46, 22)

- Follow-up questions

1. What is the precise distribution of BNPs and INPs with role nouns and what governs the variation in historical German
2. How did the modern-German distribution arise; in view of the data in (3) it cannot be attributed to incomplete grammaticalization of the indefinite article (contra Szczepaniak 2009, p. 85)
3. How can these findings be integrated into current semantic accounts on the drop of the article in nominal predicates denoting social roles

2. Corpus and distribution of bare vs. indefinite role nouns in historical German

2.1. Principles of the distribution in Modern German

- as in a series of languages displaying an indefinite article (e.g. Dutch, see de Swart et al. 2007 or Romance, see Zamparelli 2008) the use of the article in copular constructions in modG depends on the semantic type of the indefinite expression

---

1 Cf. lat. constabularius, one of the high court appointments, in charge of travelling and horses.
Bare vs. indefinite nominal predicates in the history of German

(4) a. **Class A (role nouns) = affiliation to socially well-established social roles → bare nominal (BNP)**


   b. **Class B (class nouns) = classes of individuals or representatives of a class → modified by an article (INP)**

   sub-kinds of humans (*Mann* ‘man’), properties (*Riese* ‘giant’), epitheta (*Held* ‘hero’), swear words (*Idiot* ‘idiot’), etc.

   • what can be derived about the historical pre-stages, apart from the examples cited in Behaghel (1923)

2.2. Historical German

2.2.1. Corpus and search

   • word search for prototypical role nouns (nouns denoting crafts, occupation, nationalities, religious denominations, official functions etc.) in electronic corpora containing texts of the following periods (and dialectal arias)

   o Early New High German, c. 1350 – 1650
   o Middle High German, c. 1050 – 1350
   o Old High German (c. 750 – 1050) is left out because of lack of conclusive data (small number of copular nouns and only one with an article, Petrova 2015)
   o Middle Low German (13th to 16th century) will be tested to assure that the picture gained for MHG/ENHG is not due to high number of texts from the South, which displays a high preference for indefinite articles not only in predicates but with also singular mass and abstract nouns even today (Glaser 1996)

2.2.2. Results

2.2.2.1 Middle Low German

   • *Seelentrost*, only prose MLG text in TITUS (Thesaurus indogermanischer Text- und Sprachmaterialien), 15th century, 130.000 tokens
   • lexical equivalents of role nouns found in this text occur with an article without any exception

(5) crafts / occupations

   a. *do wart he eyn herde*
      ‘then he became a shepherd’ (Seel., 5, 163, 13)
   b. *Kayn de was eyn ackerman*
      ‘Cain, he was a peasant’ (Seel., 5, 163, 14)

---

2 But note that roles nouns are expressed via INPs in some modern dialects, predominantly in the South and South-West, see Henn-Memmesheimer 1986, Weber 1987, Glaser 1996, Kolmer 1999, *inter alia*. For similar variation in the modern standard, see Berman 2009.
c. *de was eyn kopman*
   ‘he was a merchant’ (Seel., 9, 240, 5)

(6) nationality / religious denominations
   a. *Do he horde, dat he eyn yode was*
      ‘When he heard that he was Jewish’ (Seel., 3, 117, 3)
   b. *dat de konningynne ock eyn yodynne was*
      ‘that the queen was also Jewish’ (Seel., 3, 117, 9)
   c. *Desse koningk was eyn heyden*
      ‘This king was a heathen’ (Seel., 0, 3, 17)

 devise preference for INP with role nouns is no strict Southern phenomenon but also typical for other dialectal regions during the period

2.2.2.2 Early New High German

- word search in Bonner ENHG Corpus and TITUS
- nouns denoting crafts and occupations come exclusively with the article through all sub-periods and regional areas

(7) *das er ein Bawr [...] war*
   ‘that he was a peasant’ (Johann Bange: „Chronik, Mühlhausen 1599“; Thuringian; 1550–1600)

(8) *sein Vater war ein Schneider*
   ‘his father was a tailor’ (Hans Michael Moscherosch: „Gesichte, Straßburg 1650“; Asssarine, 1650–1700)

- variation with appointments (teacher, headmaster, messenger) → INP prevail but BNP are possible as well

(9) *Diser ist maister gewesen des manns Pitagore*
   ‘This one was a teacher of Pitagoras’ (Gualtherus Burlaeus: „Vita, Augsburg 1490“; East Swabian; 1450–1500)

(10) *Vnd ein maister Archillej gewesen*
   ‘and was a teacher of Achilles’ (Gualtherus Burlaeus: „Vita, Augsburg 1490“; East Swabian; 1450–1500)

(11) *die stat oder dz land dennen derselb bot ist*
   ‘the city and the land of which he is a messenger’ (Gerold Edlibach: „Chronik, Zürich 1485-1486“; East High Alemannic: 1450–1500)

(12) *ich bin ain bot Priamus*
   ‘I am a messenger of Priamus’ (Hans Mair: „Troja, Nördlingen 1392“; East Swabian; 1350–1400)

- only BNP with nouns denoting positions in administration and legal services
  *(Burgermeyster ‘mayor’, Rentmeister ‘treasurer’, Schreiber ‘clerk’)*
(13) *de Burgermeyster gewest was*
‘who was a mayor’ (Buch Köln 14. Jahrhundert, Köln 1360-1396“; Ripuarian; 1350–1400)

(14) *do he Rentmeister was*
‘when he was a treasurer’ (Buch Köln 14. Jahrhundert, Köln 1360-1396; Ripuarian; 1350–1400)

(15) *Herr Ferdinand Stenglin / ist Stadtschreiber*
‘Mr Ferdinand Stenglin is a clerk‘ (Christoph Schorer: „Chronik Memmingen, Ulm 1660“; Swabian; 1650–1700)

⇒ traces of the modG distribution visible, in that nouns denoting **institutionalized social roles** consequently occurs as BNPs

### 2.2.2.3 Middle High German

- lexical search for word forms of lemma in the online Concordance of the MHG Dictionary (Konkordanz des Mittelhochdeutschen Wörterbuchs)

- **Group 1**: no variation in article use

- role nouns that never take the article (17 hits/100%): those denoting medieval **court appointments** (*trucheze* ‘steward’, *marschalc* ‘marshal/master of the Horse’, *cemerere* ‘treasurer’, *skence* ‘cupbearer’) as well as **positions in the hierarchy of the church** (*abbet* ‘abbot’, *babes* ‘pope’, *bischof* ‘bishop’)

(16) a. *Sindolt der was scenke*
‘Sindhold, he was cupbearer‘ (NibB 11,3)

b. *Hûnolt was kamerære*
‘Hunold was treasurer‘ (NibB 11, 4)

(17) *dô bat der keiser Fôcas/ den herren, der dô bâbest was*
‘then Emperor Focas asked the man who was pope at that time‘ (Eracl 2376)

- **Group 2**: role nouns that allow for variation in article use (69 hits)

- with all remaining role nouns (crafts, occupations: *accherman* ‘peasant’, *hirte* ‘shepherd’, *iegere* ‘hunter’, *vischere* ‘fisher’, *koufman* ‘salesman’, *smit* ‘blacksmith’, *arzat* ‘physician’, *richtaere* ‘judge’, *mûnech* ‘monk’, *priester/pfaffe* ‘priest’, or nationalities), there is variation in the type of NP as indicated by Behaghel (1923), but some factors become visible

⇒ **factor one: coordination / disjunction**

- coordination triggers BNP (as known to be the case cross-linguistically, see Heycock and Zamparelli 2003), see (18)a-b and (19)a-b

(18) a. *daz ich ein ieger bin*
‘that I am a hunter‘ (Herb 18070)
b. *der altere wart jagire und accherman*  
‘the elder one became a hunter and a peasant’ (Gen 1072)

\[(19)\]  
a. *er was kurtoys,/ sîn vater was ein Franzoys*  
‘he was well-mannered, his father was a Frenchman’ (Parz 46, 22)  
b. *ist Kingrün/ Franzoys od Bertûn*  
‘May Kingrun be a Frenchman or a Breton’ (Parz 195, 28)

- coordination/disjunction triggers BNP s in a statistically significant way, see Table 1

<table>
<thead>
<tr>
<th>Coordination/disjunction</th>
<th>Single Nouns</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNP</td>
<td>13</td>
<td>39</td>
</tr>
<tr>
<td>INP</td>
<td>1</td>
<td>33</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>14</strong></td>
<td><strong>72</strong></td>
</tr>
</tbody>
</table>

Fisher exact test statistic value 0.006711, significant at p < .05  
Table 1: The effect of coordination/disjunction on article drop in predicate NPs containing role nouns in MHG

⇒ **Factor two**: referential status of the subject DP

- INPs in clauses in which the subject denotes an individual referent, see (20)a-(21)a  
- BNP s occurs in clauses whose subject expression does not refer to a single individual, but is of the class of non-referring expressions, e.g. a (negative) quantifier, a free relative or a generic noun, see (20)b-(21)b

\[(20)\]  
a. *Asclêpius ein arzât was*  
‘Asclepius was a physician’ (RvEBarl 10083)  
b. *dâ sol nieman arzât wen*  
‘there, no one should be a physician’ (SM:St 3: 2, 7)

\[(21)\]  
a. *Heren Euerhart Hardevuyst […] de zo der zijt ein richter was*  
‘Mr E.H., who was a judge at that time’ (God. Hagen, Reimchr. I, 370-371)  
b. *Wer richter sin muge*  
‘Whoever might be a judge’ (Sachsenspeigel 1, LV)

- the referentiality of the subject is statistically significant, see Table 2

<table>
<thead>
<tr>
<th>Subj [+REF]</th>
<th>Subj [-REF]</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNP</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>INP</td>
<td>32</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>48</strong></td>
<td><strong>7</strong></td>
</tr>
</tbody>
</table>

Fisher exact test statistic value 0.012974, significant at p < .05  
Table 2: The effect of referentiality of the subject expression on article use in predicate NPs containing role nouns in MHG

⇒ denotation of an **institutionalised position or function** is a core and historically constant domain of BNP in German
referentiality of the subject DP as a factor in article use with role nouns denoting non-institutionalized roles is a novel factor in the distribution

3. Theoretical implementation

Interim conclusion

- one of the basic results of the corpus study → BNP is associated with social roles in terms of socially well-established institutionalized kinds of activities (professions, hobbies, professions etc.) or statuses (citizenship, faith etc.), or the concepts of these respectively
  - court appointments are conceptualized as well-established activities into which one may be introduced at some point in his career → expressed by means of BNPs from the beginning of the attestation
  - crafts, occupations, nationalities in the respective period are no social roles proper but affiliations that are obtained by birth → considered as ordinary / class nouns and expressed by INPs
- variation and change in the realization of predicates referring to social roles is due to extra-linguistic factors → inventory of social roles obtained by appointment increases
- modern dialectal situation: overgeneralization of the article in the souths, as with abstract and singular mass nouns as well

follow-up question: why do role nouns occur as bare predicates

- previous accounts: correlation of the semantic deficiency and an impoverished structure of the DP

gender deficiency (Zamparelli 2008)

- role nouns are deficient for gender but receive gender specification via agreement with the subject NP, which is in a close relation with the nominal predicate in the argument structure of the copula
- in English, where we have no gender, no feature is there to be transmitted via Agree → no BNPs among role nouns → subject in an outer position
- criticism on the basis of modern German data in Geist (2014)

(22) Antoinette-FEM war unschuldiges-NEUTR Opfer-NEUTR ... 

problem: how to explain INPs with role nouns in languages which have specification for gender as the languages presented here

number deficiency (de Swart et al. 2007)

- role nouns are deficient for number, and number is located in NumP → lack of NumP in the fine structure of the DP in (23) → no article means no NumP (for discussion, see Zamparelli 2008, Geist 2014)

(23) [DP D [NumP semantic number or REL operator [NP N]]] (de Swart et al. 2007)

Proposal: nouns denoting social roles are number deficient, but this is not because of the lack of functional layers in the DP
Recall that in Longobardi (1994), the common property of bare nouns is number deficiency (singular mass, plural count, existential indefinites etc.) \( \rightarrow [d e] \) is lexically governed;

Could bare nouns denoting social roles be one of these?

Geist (2014) introduces the notion of ‘event kinds’ \( e_k \), similar to definites referring to kinds, to explain the denotation of BNP in modG.

According to Krifka et al. (1995), kinds are proper names for well-established categories, and proper names, according to Longobardi (1994) are subject to N-movement into empty D position

(24) a. \([\text{DP } [d e] [\text{NP } [n \text{ XP }]]] \rightarrow b. [\text{DP } [d \text{ XPt }] [\text{NP } [n \text{ ti }]]]\)

I extend N-movement to D to social roles proper, i.e. to denotations of well-established kinds of activities/states

- \([d e] \) is lexically governed (\( N^o \) is singular but has no number specification)
- \( N \) remains in NP if used not as a social role alias proper name but as an appellative, i.e. if modified by an evaluative adjectival modifier, a relative clause or in figurative meaning \( \rightarrow \) in these cases, an indefinite determiner is required because \( N \) is a regular singular count noun

(25) a. Er ist (*ein) Künstler / (*ein) Athlet / (*ein) Preuße \( \rightarrow \) profession / origin
   ‘He is an artist, an athlete, a Prussian’

b. Er ist *einer bekannter / talentierter / erfolgreicher Künstler
   ‘He is a famous / talented / successful artist

c. Er ist *(ein) Künstler, den jeder kennt
   ‘He is an artist whom many people know

d. Er ist *(ein) Künstler / *(ein) Athlet / *(ein) Preuße
   ‘He is creative / has an athletic appearance /
   \( \rightarrow \) behaviour/appearance
   shares properties of Prussian people’

\( \Rightarrow \) still tentative: role of referentiality of the subject expression in MHG

bare N no role nouns proper, so N-movement to D cannot be assumed
alternative: \( N \) remains in \( N^o \) and licenses empty D under the assumption that the predicate \( N \) agrees with a subject which is number deficient / existentially bound

4. Conclusions

\( \Rightarrow \) bare nominal predicates denoting well-established social roles are interpreted as proper names undergoing N-movement to D (Longobardi 1994)

\( \Rightarrow \) historical development in German

- denotations of social roles is constant over the entire history of the attestation
- changes in social life and structure lead to changes the inventory of social roles and the concepts of them \( \rightarrow \) inventory of nouns undergoing N-movement to D increases, ENHG is the transitional period
- loss of \( [d e] \) in some dialects \( \rightarrow \) not only with nouns denoting social roles but with with singular mass and even plural counts (Glaser 1996)
References

Electronic resources
http://www.mhdwb-online.de/quellenverzeichnis.php
http://titus.uni-frankfurt.de/indexd.htm
https://korpora.zim.uni-duisburg-essen.de/Fnhd/

Secondary literature


de Swart, Henriette, Yoad Winter and Joost Zwarts. 2007. Bare Nominals and Reference to Capacities. NLLT 25(1), 195–222.


1 Introduction

1.1 The Grammaticalization of Indefinite Articles

- The cardinal one is one of the diachronic sources of indefinite articles

- Development of the indefinite article according to grammaticalization theory (see Heine, 1997: 72ff.):

<table>
<thead>
<tr>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Stage 4</th>
<th>Stage 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>numeral</td>
<td>presentative</td>
<td>specific</td>
<td>nonspecific</td>
<td>generalized</td>
</tr>
</tbody>
</table>

- My concern here will be mainly with the first stages of the grammaticalization of the indefinite article (whether these stages mentioned by Heine are real or not)

- Some relevant quotes from Heine:

  An early stage II would obtain in a language where the article is confined to the beginning of a narrative discourse — for instance, where the main participants are introduced with the presentative article at the beginning of the tale. (ibid.: 72; my emphasis)

  [Stage III] The article is no longer confined to presentative uses; it is extended typically to any participant in discourse known to the speaker but presumed to be unknown to the hearer, irrespective of whether or not the particant is expected to be taken up in subsequent discourse. Nevertheless, the use of the article is still associated with subsequent mentions, even if much less so than at stage II. (ibid.: 72f.; my emphases)

- Problems:
  - work used to establish the schema is not diachronic, but typological
  - there is no connection to known meaning components of the cardinal or their uses (which is strange, since grammaticalization theory is supposed to be usage-based…)

(1) Each conference participant ate one Maultasche, and everyone was sick afterwards. The conference, however, was a delightful experience …
* Cardinal can have non-specific uses (see (1): most probable reading of one = narrow scope)
* Cardinal can be used for ‘throw-away references’ *(see (1))
  – Why should these possibilities ever disappear? Or why would speakers ever not use these possibilities?

- My take on the issue: try to determine (discourse) behavior of English one, and use that as a tool for determining why UNUS is used
- Since UNUS is a longer form (less economic/more expressive) with respect to the bare “standard form”, its use needs to be motivated (and ideally, not because the source of translation has ‘one’)

1.2 Brief Notes on the Latin of the Vulgata

- The Vulgata is a Bible translation by St. Jerome (=347–420 CE), commissioned by Pope Damasus in 382, translated from Greek (New Testament) and (probably) from Hebrew (Old Testament, but not 100% certain)
- Features post-classical Christian Latin (and is based on earlier translations, collectively known as Vetus Latina) — basically a language without articles
- Translation of sacred writing might reflect original more closely than contemporary (oral) Latin!
- Downloaded from Bibliotheca Augustana and extracted occurrences of UNUS

<table>
<thead>
<tr>
<th></th>
<th>Old Testament</th>
<th>New Testament</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>766</td>
<td>320</td>
<td>1086</td>
</tr>
</tbody>
</table>

- Compare: English Catholic Public Domain Version of the Bible

<table>
<thead>
<tr>
<th></th>
<th>A(N)</th>
<th>ONE(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNUS</td>
<td>14341</td>
<td>4813</td>
</tr>
</tbody>
</table>

- UNUS is clearly not an indefinite article in the Vulgate; differences to current English ONE also seem huge, but may be partly due to syntactic differences between Latin and English
- The Vulgate contains certain ‘innovative’ uses of UNUS which look suspiciously like modern indefinite articles, see, e.g. Matthew 26:69, classified by Coleman (1991: 390) as one of the ‘unequivocal examples of articular meaning’:

(2)  
[... ] et accessit ad eum una ancilla dicens [...]  
[... ] and came to him UNUS maid saying [...]  
‘and a maid came to him [Petrus] and said’

(2) introduces a new discourse participant, and is generally translated by the indefinite article in English versions of the Bible.

- Most often, elements remain bare, see, e.g., 1 Kings 1:3–4

(3)  
3 quaesierunt igitur φ adulescentulam speciosam in omnibus φ finibus Israel et invenerunt Abisag Sunamitin et adduxerunt eam ad φ regem 4 erat autem φ puella pulchra nimis dormiebatque cum φ regem et ministrabat ei φ rex vero non cognovit eam

'Term suggested by Chiara Gianollo, p.c.'
So they sought a beautiful young woman, in all the coasts of Israel and they found Abisag, a Sunamitess, and brought her to the king. And the damsel was exceedingly beautiful, and she slept with the king, and served him, but the king did not know her.

2 Discourse Properties of English One

- Why English? Clear difference between indefinite article a and unity cardinal one (contrary to Romance, written Standard German, etc.)
- According to standard dynamic semantics (e.g. Kamp and Reyle, 1993), indefinite articles introduce new discourse participants.
- The basic semantic property of one in English seems to be to introduce a new referent, while presupposing the existence of a larger entity that new referent belongs to. Consider (4).

  (4) We saw a lot of cows_{i}. [One cow]_{tel} appeared to be looking at me with an attitude like, “What? You never saw a cow before?” [internet example]

- The entity denoted by one cow in (4) is in a sense discourse new. Yet, it also is a part of a plural discourse referent already introduced (cows).
- With an indefinite article, such a partitive reading is difficult to get:

  (5) We saw a lot of cows_{i}. #[A cow]_{tel} appeared to be looking at me with an attitude.

- One is also often used in contrastive contexts:

  (6) a. One cow was looking at me; another (cow) simply ignored me.
   b. Peter drank one bottle of beer, Fred had three bottles.

- One has determiner and pronominal uses:

  (7) a. John owns no dog, but he wants one.
   b. Over the years, a mythology has developed concerning certain colors of M&M candies. The green ones are supposedly aphrodisiac; if a red one is last to emerge from a bag, make a wish and it will come true; if the last one is yellow, call in sick and stay home. [Example from COCA]

- A partitive projective content comes necessarily with one in these uses. It can be textually given, but it is also involved when the cardinal is used in a contrastive manner — but here, the existence of a collective entity Ns needs to be inferred/accommodated.

- Not that indefinite article as such does not exclude a partitive effect; when contrastively opposed to a definite article, this is precisely what one gets (see (8)).

  (8) That wasn’t A reason I left Pittsburgh, that was THE reason. [Abbott, 2004]

- However, as shown by (5), it cannot refer as a partitive anaphora to a close and explicitly given plural antecedent.
- If the indefinite article has a partitive component (which is somehow part of its lexical meaning), it must be obligatorily accommodated.
Can we simply transfer from English to Latin? Not completely: the alternatives are not the same:

- English has fully grammaticalized indefinite and definite articles
- Latin has bare noun (phrases) and very limited *ille*

3 A Detailed Look at ‘Innovative’ Uses of Adnominal UNUS in the Vulgate

I will have a look at adnominal uses of UNUS, which look like indefinites articles (at least at first sight). I will try to identify elements that may have motivated the use of UNUS.

Most uses of UNUS fall are unproblematic, given an analysis as a cardinal: they are either used in contrastive contexts (see (9a)), or in partitive contexts like (9b-c).

(9) a. *itaque iam non sunt duo sed una caro […]* therefore already NEG two but one flesh […]
   ‘Therefore now they are not two, but one flesh.’
   b. *tunc crucifixi sunt cum eo duo latrones unus a dextris et unus a sinistris* then crucified are with him two thieves one at right and one at left
   c. *[…] propheta unus de antiquis surrexit* prophet one of old rose
   ‘one of the old prophets (lit., one prophet of the old) has resurrected’

In what follows, I will look at uses of UNUS that are adnominal and which I have classified as introducing a new discourse participant.

Often, in looking at the larger context, motivation becomes clear.

3.1 Clear Cases of Cardinal Uses with Unusual Distance

Let us have a second look at (2) in a larger context (Matthew 26:69–71)

(10) 69 Petrus vero sedebat foris in atrio et accessit ad eum *una ancilla* dicens et tu cum Iesu Galilaeo eras 70 at ille negavit coram omnibus dicens nescio quid dicis 71 exeunte autem illo ianuam vidit eum *alia* et ait his qui erant ibi et hic erat cum Iesu Nazareno
69 But Peter sat without in the court. And there came to him a servant maid, saying: Thou also wast with Jesus the Galilean. 70 But he denied before them all, saying: I know not what thou sayest. 71 And as he went out of the gate, another maid saw him; and she saith to them that were there: This man also was with Jesus of Nazareth.

This clearly is an instance of a contrastive use: *una ancilla* vs. *alia*.

Unusual distance (viewed from the perspective of contemporary English) between these two elements, which does not warrant the use of English *one*.

In Latin, the noun *ancilla* is not taken up again, contrary to the English translation.

*una ancilla* in (10) is quite close to a throw-away referent: extremely secondary character; not taken up again with a pronoun (reappears only in *dicis* – you say)
• Similar case: Zechariah 5:7–9

(11) 7 et ecce talentum plumbi portabatur et ecce mulier una sedens in medio amphorae 8 et dixit haec est impietas et proiectit eam in medio amphorae et misit massam plumbem in os eius 9 et levavi oculos meos et vidi et ecce duae mulieres egredientes et spiritus in alis earum et habebant alas quasi alas milvi et levaverunt amphoram inter terram et caelum 7 And behold a talent of lead was carried, and behold a woman sitting in the midst of the vessel. 8 And he said: This is wickedness. And he cast her into the midst of the vessel, and cast the weight of lead upon the mouth thereof. 9 And I lifted up my eyes and looked: and behold there came out two women, and wind was in their wings, and they had wings like the wings of a kite: and they lifted up the vessel between the earth and the heaven.

Verse 7: discourse-new woman, reinforced with ecce.
Verse 9: discourse-new 2 women, reinforced with ecce.

Once again: too great a distance for English one, but still a clear case of numerical contrast
Once again: at best a secondary figure, taken up again twice by eam and eius in verse 8


(12) 51 erat autem turris excelsa in media civitate ad quam conugerant viri simul ac mulieres et omnes principes civilis clausa firmissime ianua et super turris tectum stantes per propugnacula 52 accedentes Abimelech iuxta turrem pugnabant fortiter et adpropinquans ostio ignem subponere nitebatur et ecce una mulier fragmen molae desuper iaciens in lisit capiti Abimelech et confregit cerebrum eius qui vocavit cito armigerum suum et ait ad eum evagina gladium tuum et percute me ne forte dicatur quod a femina interfecit eum 51 And there was in the midst of the city a high tower, to which both the men and the women were fled together, and all the princes of the city, and having shut and strongly barred the gate, they stood upon the battlements of the tower to defend themselves. 52 And Abimelech, coming near the tower, fought stoutly: and, approaching to the gate, endeavoured to set fire to it: 53 And behold, a certain woman casting a piece of a millstone from above, dashed it against the head of Abimelech, and broke his skull. 54 And he called hastily to his armourbearer, and said to him: Draw thy sword, and kill me: lest it should be said that I was slain by a woman. He did as he was commanded, and slew him.

Verse 53: discourse-new woman, reinforced with ecce.
Verse 51: plural discourse referent mulieres, of which una mulier is a part
Standard partitive use of the cardinal, with a non-standard distance.
Throw-away referent: not taken up again (unless one considers a femina in verse 54 to be anaphoric to una mulier)

• Slightly less clear (depending on the interpretation), but in principle unproblematic: Ezekiel 9:2

(13) 51 et ecce sex viri veniebant de via portae superiores quae respicit ad aquilonem et uniuscuiusque vas interitus in manu eius vir quoque unus in medio eorum vestitus lineis et atramentarium scriptoris ad renes eius et ingressi sunt et steterunt iuxta altare aereum 51 And six men came from the way of the upper gate, which looketh to the north: and each one had his weapon of destruction in his hand: and there was one man in the midst of them clothed with linen, with a writer's inkhorn at his reins: and they went in, and stood by the brazen altar.

Is the one man a member of the group of the 6 men? If yes: Partitive; if no: numeral contrast
3.2 Slightly Less Clear Cases

Possibly non-referential use of a (potential) antecedent:

- Chapter 12: sequence of teaching in parables; last teaching: beware of the scribes...who “eat” the houses of the widows, and conveniently, there comes a widow...Mark 12:42

(14) 38 et dicebat eis in doctrina sua cavete a scribis qui volunt in stolis ambulare et salutari in foro 39 et in primis cathedris sedere in synagogis et primos discubitus in cenis 40 qui devourant domos viduarum sub obtentu prolixae orationis hii accipient prolixius iudicium

Not clear whether occurrence of *viduarum* is to be taken referentially — if it is not, that undermines the case for a partitive use (in any case, it is not a very salient discourse referent)

The referent is taken up again below; but once again, this is not an individual who as such has any importance in the story

Contrast, but without the same predicate

- Matthew 8:19–22. It has been a long day of miracle-working and soliciting of all kinds of people. It is evening...

(15) 19 et accedens unus scriba ait illi magister sequar te quocumque ieris 20 et dicit ei Jesus vulpes foveas habent et volucres caeli tabernacula Filius autem hominis non habet ubi caput reclinet 21 alius autem de discipulis eius ait illi Domine permitte me primum ire et sepelire patrem meum 22 Jesus autem ait illi sequere me et dimitte mortuos sepelire mortuos suos

There is no mention of another scribe here, but there is a contrast with *another*, even though they do not share the same predicate

However: By following Jesus, the scribe becomes a disciple of Jesus

Would be perfectly as expected if it was pronominal *unus* with *scriba* as apposition — *some guy, a scribe*

Secondary character
3.3 Cases Without Textual Antecent or Contrastive Element

- What will be looked at: discourse-new referents introduced with *UNUS* which would be translated by an indefinite article and where there is no clear textual antecedent or a contrastive element with the same noun-predicate

### Absolute Beginnings

- 1 Samuel 1:1–3

> (16) 1 [fuit vir unus](1) de Ramathaim-sophim de monte Ephraim et nomen eius Helcana filius Hieroam filii Heliu filii Thau filii Suph Ephratheus 2 et habuit duas uxorres nomen uni Anna et nomen secundae Fenenna fueruntque Fenennae filii Annae autem non erant liberi 3 et ascendebat vir ille de civitate sua statutis diebus ut adoraret et sacrificaret Domino exercituum in Silo erant autem ibi duo filii Heli Ofni et Finees sacerdotes Domini

> 1 There was a man of Ramathaim-sophim, of Mount Ephraim, and his name was Elcana, the son of Jeroham, the son of Eliu, the son of Thohu, the son of Suph, an Ephraimite: 2 And he had two wives, the name of one was Anna, and the name of the other Phenenna. Phenenna had children: but Anna had no children. 3 And this man went up out of his city upon the appointed days, to adore and to offer sacrifice to the Lord of hosts in Silo. And the two sons of Heli, Ophni and Phinees, were there priests of the Lord.

Absolute beginning of discourse: no possible antecedent

No other man mentioned in the text in proximity (apart Ofni and Phinees), but classified as *filii Heli* — sons of Heli

Other elements are numbered: 2 wives; 2 sons of Heli

Elcana is an important character: the father of Samuel

What may have triggered the use of *UNUS* here? Compare Job 1:1–3

> (17) 1 [vir erat](1) in terra Hus nomine Iob et erat vir ille simplex et rectus ac timens Deum et recedens a malo 2 natique sunt ei septem filii et tres filiae 3 et fuit possessio eius septem milia ovium et tria milia camelorum quingenta quoque iuga boum et quingentae asinae ac familia multa nimis eratque vir ille magnus inter omnes Orientales

> 1 There was a man in the land of Hus, whose name was Job, and that man was simple and upright, and fearing God, and avoiding evil. 2 And there were born to him seven sons and three daughters. 3 And his possession was seven thousand sheep, and three thousand camels, and five hundred yoke of oxen, and five hundred she asses, and a family exceedingly great: and this man was great among all the people of the east.

Similarly followed by (even more) numbered entities

Verbal tense is different: (17)– imperfectum; (16)– perfectum

Job is *the* main character of the book; Elcana is secondary — even though the book starts with him

Other than that, there does not seem to be much difference here: it seems extremely unlikely that Job and Elcana were the only men in Hus and Ramathaim-sophim, respectively

### Unclear Contrast Element

- 1 Kings 20:13 & 1 Kings 20:28

- Some context: main character of chapter 19 is the prophet Elias, who is said to have killed all the prophets with a sword (which is strange, as we will see).
Ahab, King of Israel, is being besieged by Benhadad, the king of Syria.

13 et ecce \textit{propheta unus} accedens ad Ahab regem Israel ait haec dicit Dominus certe vidisti omnem multitudinem hanc nimirum ecce ego tradam eam in manu tua hodie ut scias quia ego sum Dominus [. . .] 20 et percussit unusquisque virum qui contra se ven- erat fugeruntque Syri et persecutus est eos Israhel fugit quoque Benadad rex Syriae in equo cum equitibus 21 necnon et egressus rex Israhel percussit equos et currus et percussit Syriam plaga magna 22 accedens \textit{propheta} ad regem Israhel dixit ei vade et confortare et sicut et vide quid facias sequenti anno rex Syriae ascendente contra te [. . .] 26 igitur postquam annus transierat recensuit Benadad Syros et ascendit in Afec ut pugnaret contra Israhel 27 porro filii Israhel recensit sunt et acceptis cibariis de Syris et ascendit in Afec ut pugnaret contra Israhel. [. . .] 35 Then a certain man of the sons of the prophets, said to his companion, in the word of the Lord: Strike me. But he would not strike. [. . .] 37 Then he found another man, and said to him: Strike me. And he struck him and wounded him. So the prophet went, and met the king in the way, and disguised himself by sprinkling dust on his face and his eyes. 39 And as the king passed by, he cried to the king, and said: Thy servant went out to fight hand to hand: and when a certain man was run away, one brought him to me, and said: Keep this man: and if he shall slip away, thy life shall be for his life, or thou shalt pay a talent of silver. 40 And whilst I, in the hurry, turned this way and that, on a sudden he was not to be seen. And the king of Israel said to him: This is thy judgment, which thyself hast decreed. But he forthwith wiped off the dust from his face, and the king of Israel knew him, that he was one of the prophets. And he said to him: Thus saith the Lord. Because thou hast let go out of thy hand a man worthy of death, thy life shall be for his life, and thy people for his people.
Several things going on here

- General schema: X says Y to Ahab; Ahab does Y; Ahab wins
- in verses 13 & 28: propheta unus & unus vir Dei; in verses 22 & 38 \( \varnothing \) propheta; in verse 35 vir quidam
- Why propheta unus in 13? Possibly to indicate that the prophet is not Elias; in verse 22 bare form to indicate that it is the same prophet?
- Why unus vir Dei? Possibly because vir Dei would suggest coreference with prophet?
- It is not that clear whether we should take these elements to co-refer or not...
- Notice the vir quidam in verse 35 opposed to alterum virum
- Why vir unus in verse 39? vir has appeared before, maybe to avoid coreference as well?
- Note: Hebrew has `exad in verses 13 & 35, but the bare noun in verses 28 & 39.\(^5\)

• How could avoidance of coreference with unus N work (in a language without real articles)?
  - Assume an entailment scale: Elias \( \subseteq \) prophet \( \subseteq \) man of god. Such scales are classical elements for identifying discourse referents:
    (18) a. Garfield, was expelled from his home. [The cat], had eaten …
    b. Yesterday, we found [a cat], in the street. [The poor animal], …
  - By assumption, unus N introduces a projective content that there is another entity satisfying N in the context.
  - That projective content is anaphorically satisfied by a more specific antecedent \( \text{unus propheta} \rightarrow \text{satisfied by Elias}; \text{unus vir Dei} \rightarrow \text{satisfied by (unus) propheta} \)
  - the bare form has no such presupposition attached to it, and would be compatible with a coreferent reading

Unclear Antecedent (if there is one), no Contrast Element

• Matthew 9:18 (with some context). Jesus comes into his hometown and continues doing miracles. People introduced into discourse before these verses: scribes, tax collectors, sinners & Pharisees

(19) 14 tunc accesserunt ad eum discipuli Iohannis dicentes quare nos et Pharisaei ieunamus frequenter discipuli autem tui non ieunant et ait illis Iesus numquid possunt filii sponsi lugere quamdiu cum illis est sponsus venient autem dies cum auferetur ab eis sponsus et tunc ieunabunt 16 nemo autem inmittit commissuram panni rudis in vestimentum vetus tollit enim plenitudinem eius a vestimento et peior scissura

14 Then came to him the disciples of John, saying: Why do we and the Pharisees, fast often, but thy disciples do not fast? 15 And Jesus said to them: Can the children of the bridegroom mourn, as long as the bridegroom is with them? But the days will come, when the bridegroom shall be taken away from them, and then they shall fast. 16 And nobody putteth a piece of raw cloth unto an old garment. For it taketh away the fulness thereof from the garment, and there is made a greater rent. 17 Neither do they put new wine into old bottles. Otherwise the bottles break, and the wine runneth out, and the bottles perish. But new wine they put into new bottles: and both are preserved. 18 As he was speaking these things unto them,

behold a certain ruler came up, and adored him, saying: Lord, my daughter is even now dead; but come, lay thy hand upon her, and she shall live.

There is nothing subsequently that might motivate the presence of unus

Much hinges on the interpretation of princeps here: the New International Version and other translations give this as leader of a synagogue — which might then be a Pharisee

Secondary character in any case

No Contrast, No Antecedent

• Matthew 21:19

(20) 18 mane autem revertens in civitatem esurit 19 et videns fici arborem unam secus viam venit ad eam et nihil inventit in ea nisi folia tantum et ait illi nuncquam ex te fructus nascatur in sempiternum et arefacta est continuo ficulnea

And in the morning, returning into the city, he [Jesus] was hungry. 19 And seeing a certain fig tree by the way side, he came to it and found nothing on it but leaves only. And he saith to it: May no fruit grow on thee henceforward for ever. And immediately the fig tree withered away.

Not obvious what motivates unus here

It seems plausible that this tree was the only fruit tree around

Identity of the fig tree is completely unimportant; merely a pretext for a parable

3.4 Probably ‘Normal’ Cardinals

• What we look at here: things that might be rendered with one in English, or only one, or a single one; where there might be reason to stress cardinality

• John 6:9: story of Jesus feeding five thousand with five loaves. Jesus asks: Where could we buy bread such that all may eat?

(21) 7 respondit ei Philippus ducentorum denariorum panes non sufficiunt eis ut unusquisque modicum quid accipiat 8 dicit ei unus ex discipulis eius Andreas frater Simonis Petri 9 est puer unus hic qui habet quinque panes hordiacios et duos pisces sed haec quid sunt inter tantos 10 dixit ergo Jesus facite homines discumbere erat autem faenum multum in loco discubuerunt ergo viri numero quasi quinque milia

7 Philip answered him: Two hundred pennyworth of bread is not sufficient for them that every one may take a little. 8 One of his disciples, Andrew, the brother of Simon Peter, saith to him: 9 There is a boy here that hath five barley loaves and two fishes. But what are these among so many? 10 Then Jesus said: Make the men sit down. Now, there was much grass in the place. The men therefore sat down, in number about five thousand.

The boy is discourse-new (and a throw-away reference)

Verse 9 might have been translated also as: there is only one boy here/a single boy

Numeral information very salient here — but the original Greek has no cardinal:

(22) Estin φαιδαριον ἱὸς ἰσχεὶ πεντε αρτοὺς κριθινοὺς

is little boy here who has five loaves barley
4 The Nature of the Projective Content

- Can we be more precise about the nature of the projective content? Arguably yes, if we look at Tonhauser et al. (2013)

- They make distinction of two independent properties of projective content:
  - strong contextual felicity
  - obligatory local effect

- Strong Contextual Felicity as defined by Tonhauser et al. (ibid.: 75f.)

  \[(23)\]
  a. \(m\)-positive and \(m\)-neutral contexts: An \(m\)-positive context is an utterance context that entails or implies \(m\). An \(m\)-neutral context is an utterance context that entails or implies neither \(m\) nor \(\neg m\).
  b. Strong Contextual Felicity Constraint: If utterance of trigger \(t\) of projective content \(m\) is acceptable only in an \(m\)-positive context, then \(t\) imposes a strong contextual felicity constraint with respect to \(m\).

AFAIK, interpreted to correspond to (necessary) anaphoricity.

Improbable that the trigger \(unus\ N\) comes with a strong contextual felicity constraint with respect to the existence of a more than one N (see (16) and other examples where it appears in beginning of discourse).

- Obligatory Local Effect as defined by Tonhauser et al. (ibid.: 93)

  \[(24)\] A projective content \(m\) with trigger \(t\) has obligatory local effect if and only if, when \(t\) is syntactically embedded in the complement of a belief-predicate \(B\), \(m\) necessarily is part of the content that is targeted by, and within the scope of, \(B\).

This is potentially tricky to test, but there is one element that should do.

- Not from the Vulgate, but same textual layer:

  \[(25)\] Credo in unum deum [Nicene Creed]
  believe.1SG in UNUS.Acc god.Acc

  (25) should not commit the speaker to the existence of another god.

  Depending on what your idea ‘belief’ here is, (25) should not fall prey to Moore’s paradox

  \[(26)\] There is at least one other god, and I do not believe in it/him/her.

  Not quite sure that this is the sort of belief-predicate Tonhauser et al. (ibid.) have in mind, but:

  - asserting the existence of one god doesn’t make much sense if anybody believes/knows that there is only one (traditional Roman society was polytheistic)
  - assuming that believe creates here an intensional context:
    * there is at least another entity that is a god in (epistemically?) accessible possible worlds (e.g., Zeus, Juno, etc.)
    * these epistemically accessible worlds do not belong to what the speaker assumes to be the set of reality-compatible worlds
• * unus’* projective content should have obligatory local effect

• What exactly is the content of the presupposition?

• Problem (pointed out by one reviewer):

(27) Q: How many books do you have?
   R: I have one book.
   R’: #I have a book.

This should not mean that I have at least two books.

The meaning of the question in (27) according to standard semantics should be something like the following:

(28) \{I have one book, I have 2 books, … I have n books\}

So, alternatives of which one book are under discussion; only they have not been asserted (yet).

• What is the exact content of the projective content of \(\Phi[one \ N]\) (where \(\Phi[one \ N]\) is a proposition containing one \(\ N\), \(N^*\) a plural predicate \(N\), and \(\oplus\) Link’s sum-operator)?

(29) a. \(\exists y[N(y) \land N^*(x \oplus y) \land y \neq x \land \Phi[y]]\)
    b. \(\exists y[N(y) \land N^*(x \oplus y) \land y \neq x \land \diamond [\text{Inertia}] \Phi[y]]\)
    c. \(\exists y[N(y) \land N^*(x \oplus y) \land y \neq x]\)
    d. \(\diamond \exists y[N(y) \land N^*(x \oplus y) \land y \neq x]\)

I would venture that it is about (29b-c) — (29d) would be extremely weak; but even (29b) does not impose many restrictions…

• Maybe that this is rather about salience of other elements than about the effective existence of other elements?

If I assert that there is one \(\ N\), I commit to the idea that the prior possibility of \(nN\) is not o.

• Other take on such presuppositions (see Gauker, 2003): one \(\ N\) will be asserted/assertible when it is important for the discoursive goal to state the cardinality or to exclude other cardinalities

• But: (29b)–(29c) are very weak projective contents (especially given that it can be accommodated), and there will be hardly any context where it would be excluded

• Unique inalienable possession is rather strange with one, and even contrast does not seem to rescue the issue.

(30) a. ?John has one nose.
    b. ?John has one nose, and Jill has another.

(30b) should trivially satisfy partitive presupposition (there is some plural \(X\) such that \(\text{noses}(X)\)), and in (30b), a contrastive element is satisfied. Still, (30) suggests that there should at least be a possibility that John has 2+ noses…

Evidence for (29a)?

---

6Basic idea: it would be ‘normal’ that there could be another entity \(y\) such that \(\Phi[y]\), see (31a).
• However, such a strong requirement does not hold in all cases:

(31)  a. ?I crossed Lake Constance on one boat.
    b. I crossed Lake Constance on one boat, and John crossed it on another.

Assuming that (31a) = \( \Phi \), (31b) does not need to satisfy constraints (29a), nor (29b); only (29c) seems to be warranted

• Other such examples (from COCA)

(32)  a. Bailey has lost one appeal and is mounting another.
    b. The broker, it seemed, had promised the Bayolos one carrier but was delivering another.
    c. But strategy is one thing and truth is another.

(32a) does not presuppose that Bailey has lost another appeal; there is no necessity that the broker promised another carrier in (32b); and (32c) is probably an idiom, but still does not fit well with characterizations along (29a) or (29b)

• Conclusion: projective content may be not as weak as (29c), but it surely needs to be very weak

5 Towards an Explanation of the First Steps in Grammaticalization

• Basic Assumption: indefinite article is a unity cardinal that
  – is used very often (and eventually may become obligatory in some contexts)
  – may develop a different form from the cardinal (\( \checkmark \) English, \( \checkmark \) German, not in Romance …)

• Basic thing to explain:
  – why and how does the increase in use of one/unus happen?

• Idea: there are few semantic obstacles to the use of the cardinal (if the hypothesis here is correct)
  – no textual antecedent required (but possible)
  – accommodable content trivially satisfied in most contexts (under a sufficiently generous notion of ‘context’)

• With a nearby textual antecedent in English, use of one seems to be (close to) obligatory — see example (4)

• Isolated instances of one N in English do seem to presuppose (although in an informative way) the presence of other instances of Ns in the context:

(33) One problem this book will try to solve is to let you form a working solution without it blocking you from finding the ideal solution.\(^7\)

(33) surely suggest that there are either other problems (completely unproblematic), or other problems this book will try to solve

• These isolated instances probably also have textual antecedents — even if they may be far away (in the book/the text)

\(^7\)Example from https://books.google.fr/books?isbn=0128054492.
In languages like contemporary English, non-textual antecedents are difficult with one N:

(31a) I crossed Lake Constance on one boat.

If there is no contrastive element (‘and I came back on another’), a condition that would require a numerical contrast (‘I am so incredibly tall/heavy that it might require n ships to transport me’) or a suitable antecedent, (33) is odd — even though it is a safe bet that there are other ships on the Lake of Constance.

Notice that in most circumstances, the presence of other ships on the lake of Constance is a fact that
– one can take to be obvious given world-knowledge
– has no particular textual impact/salience (and if the common ground is somehow textually constructed, and not just the set of propositions speaker and hearer take to be true, should be excluded from it)

• Scale of resolving presupposition attached to one N:
  1. by anaphora to close textual antecedent
  2. by anaphora to textual antecedent further away
  3. by accommodation to an element relevant to the situation described in the discourse
  4. by accommodation to an element in (presumably shared) world-knowledge

• The scale gives what the indefinite article is used for (or a bare noun; bare nouns in Latin do not exclude unique reference)

• The scale does not give very clear-cut frontiers

• English one seems to be obligatory for (1), and possible for (2) and (3), but not available for (4)

• English a is not available for (1), possible for (2-3), and obligatory for (4)

• Grammaticalization of a unity cardinal can be seen as extending the domain in which it is possible and obligatory to be used

6 Conclusions and Perspectives

• I have tried to show that ‘innovative’ uses of unus in the Vulgate correspond to ± standard uses of the cardinal

• Cardinal expression 1 is extremely similar in meaning to indefinite article, and there are few contexts that definitely exclude one or the other

• Cardinal expression 1 is not a clitic, and its emphatic effect will be the greater the rarer it appears

• Overusing emphasis erodes the effect of the non-clitic, and what is assumed to be the expected form.

Note Translations have been taken from the Douay-Rheims English version, and taken from http://www.latinvulgate.com.
References


The evolution of the French definite article: from strong to weak

Alexandra Simonenko and Anne Carlier
FWO/Ghent University & Université Lille 3/CNRS-UMR STL

FORMAL DIACHRONIC SEMANTICS 2016
University of Konstanz, September 12–14, 2016

1 Goals

• Empirically: Investigate the expansion of the l-article (le, la, les) in Medieval French using parsed corpora MCVF (2010) and Penn Supplement to MCVF (2010).

• Theoretically: Provide support for the hypothesis that the l-article underwent a change from strong to weak definite semantics, which explains its progressive expansion to new contexts of use.

2 Origins and use

• Common assumption: French le, la, les originate from ille (cf. Greenberg’s (1978) hypothesis about the typological origins of definite articles from distal demonstratives, De Mulder and Carlier (2011)).

Classical Latin:

• Anaphoric relations are mostly unmarked for central discourse referents.

• Demonstratives and identity markers are used for time and place indications.

Late Latin:

• Direct anaphoric relations are increasingly marked by demonstratives hic, is, ille and identity markers ipse, idem, even for central discourse referents.

• ≈40% of all the uses of ille are with relative clauses.

• However, the indirect anaphora of the type “book – D author” is never marked by ille.

Like ille in Late Latin, the l-series in Old French is used in order to establish an identity relation with a previously mentioned antecedent:

(1) Buona pulcella fut eulalia. Bel auret corps bellezour anima... La polle... good girl was Eulalia beautiful had body more.beautiful soul... The girl... ‘Eulalia was a pious young woman. She had a beautiful body and even a more beautiful soul. The girl...’ (Ste Eulalie, IX c.)

*This work has been supported by Fonds Wetenschappelijk Onderzoek (Vlaanderen) and grant ANR/DFG-14-FRAL-0006.
†alexandra.simonenko@ugent.be, anne.carlier@univ-lille3.fr
It is also used for first-mention **definites with a relative clause**:

(2) Vint en la cambra [ou eret sa muiler].
    went into the room where was his wife
    ‘He went into the room where his wife was’ (10XX-ALEXIS-V,11.115)

- E.g. Rickard (1989:55), Fournier (2002) argue that the *l*-series kept a deictic semantics up until the end of the XIII c. Unsurprisingly then, nouns without articles are found with generic and kind interpretation, definite pluralities, abstract nouns, and mass nouns (Mathieu (2009) among others).

(3) Por amor Deus et pur mun cher ami...
    for love God and for my dear friend
    “For the love of God and for my dear friend…” (10XX-ALEXIS-V,45.422)

(4) Soleill n’ i luist
    Sun not there shines
    “The sun does not shine there.” (1100-ROLAND-V,78.951)

(5) Kar vasselage par sens nen est folie
    since courage by reason not is folly
    “Since the courage with reason is not a folly.” (1100-ROLAND-V,131.1737)

Cf. Modern French:

(6) Pour *(l’) amour de Dieu...
    for the love of God
    “For the love of God”

(7) *(Le) soleil n’ y brille pas
    the sun not there shines neg
    “The sun does not shine there”

(8) Car *(le) courage avec raison n’ est pas *(la) folie
    car the courage with reason not is neg folly
    “Since the courage with reason is not a folly.”

- However, in Old French the *l*-series was also used in contexts where a bona fide demonstrative would not have been used:

(9) et si en corps a grand torment, I’ anima ’n aurra consolament.
    and if in body has great torment the soul of.it will.have consolation
    “And if the body suffers greatly, the soul will receive consolation on account of it.” (0980-LEGER-V,XXIX.194)

- This seeming “inconsistency” is commonly discussed as the problem of bare nouns in French (e.g. Mathieu 2009, Carlier and de Mulder 2010, Déchaine et al. 2016).

**Question:**

- How to reconcile the hypothesis about the origins of the *l*-series with the fact that it has neither a demonstrative nor a definite distribution in the Medieval period?

- Sub-problem: how to establish the exact distribution of the *l*-series?
3 Proposal

- Originally, the semantics of the $l$-series involved an identity relation with a context-given antecedent – “strong” definite semantics.
- An alternative, “weak” definite semantics emerged, which did not involve an identity relation, but only a uniqueness requirement.
- The new lexical entry gradually won over, which accounts for the expansion of the $l$-series.

3.1 Strong and weak definite articles: a modern example

Austro-Bavarian German has two series of definite articles, “weak” and “strong”.

<table>
<thead>
<tr>
<th>Table 1: Austro-Bavarian strong definite article</th>
</tr>
</thead>
<tbody>
<tr>
<td>M.SG</td>
</tr>
<tr>
<td>NOM</td>
</tr>
<tr>
<td>ACC</td>
</tr>
<tr>
<td>DAT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2: Austro-Bavarian weak definite article</th>
</tr>
</thead>
<tbody>
<tr>
<td>M.SG</td>
</tr>
<tr>
<td>NOM</td>
</tr>
<tr>
<td>ACC</td>
</tr>
<tr>
<td>DAT</td>
</tr>
</tbody>
</table>

Strong articles are used when a DP has an anaphoric or deictic antecedent:

(10) In da Stodtbücherei gibt ’s a Buach über Kanada. Letzens woa I doat und hob ma in det$_w$ townlibrary exists it a book about Canada recently was I there and have me $#s/des$ Buach ausboagt.
det$_w$/det$_s$ book borrowed
‘In the public library, they have a book about Canada. Recently, I was there and borrowed that book.’

(11) Gfoit da $#s’/des$ Haus?
like you det$_w$/det$_s$ house
‘Do you like that house?’ (Pointing to a house)

Or indirect antecedent (example from Standard German):

(12) Hans entdeckte in der Bibliothek einen Roman über den Hudson. Dabei Hans discovered in the library a novel about det Hudson. In the process fiel ihm ein, dass er vor langer Zeit einmal einen Vortrag von dem Autor remembered he.DAT PRT that he a long time ago once a lecture by det$_s$ author besucht hatte.
attended had.

‘Hans discovered a novel about the Hudson in the library. In the process, he remembered that he had attended a lecture by the author a long time ago.’ (Schwarz 2009:229–230)

Or else a restrictive relative clauses:

(13) A jeda Vota fiacht si voa dem Moment wenn s’ ödeste Kind ausziagt.
   Every father dreads that moment when his eldest child moves out.

A weak article is used whenever it is part of the common ground that the extension of a nominal property is a singleton in a given situation:

(14) Wie geht ’s ’n da/*dea Frau?
   How is your wife doing? (Wiltschko 2012:7)

Similar pairs in: Several German dialects (Standard, Rhineland, Cologne), Hausa, Lakhota (Schwarz 2009), Icelandic (Ingason 2016), North Frisian (Ebert 1970).

4 Strong and weak definite semantics (Schwarz 2009)

4.1 Strong definite semantics

- Elbourne (2008) for English demonstratives and Schwarz (2009) for German strong articles propose that the LF of such anaphora-oriented determiners involves a silent individual pronoun.
- The function they denote, if defined, establishes an identity relation between the referent of the silent pronoun and the referent of the DP.

(15) $[i [[D_{strong} s_r] \text{NP}]]$
   where i is an individual pronoun index and $s_r$ a situation pronoun.

(16) $[D_{strong}] = \lambda s . \lambda P_{<s,<e,t>>} . \lambda y : \exists x ![P(x)(s) & x = y] . \forall x ![P(x)(s) & x = y]$
   (adapted from Schwarz (2009))

(17) $[1 \text{ des } s_r \text{ Haus}]^g$ is defined if there exists a house identical to $g(1)$ in the situation $g(r)$,
   if defined, $[1 \text{ des } s_r \text{ Haus}]^g = g(1)$ (i.e. the individual picked out by the index 1)

- By assumption, the assignment function $g$ is defined just in case there is a salient individual in its range to which it can map the index 1.
- For the determiner function to be defined, the individual to which $g$ maps 1 has to be a house in the situation $g(r)$.
- Therefore, for this DP to be felicitous there has to be a salient house in a given situation. This captures the antecedent condition on the distribution of strong articles/demonstratives.

For the indirect anaphora case:

(18) $[D_{strong}] = \lambda s . \lambda P_{<s,<e,t>>} . \lambda y : \exists x ![P(x)(y)(s)] . \forall x ![P(x)(y)(s)]$
   (adapted from Schwarz (2009))
4.2 Weak definite semantics

(19) \[[D_{\text{weak}} \ s_r] \ NP\]
where \(s_r\) is a situation pronoun

(20) \[[D_{\text{weak}}] = \lambda s . \lambda P_{<s,<e,t>} : \exists! x[P(x)(s)] . \iota x[P(x)(s)]\]
(adapted from Schwarz (2009))

(21) \[[s' \ Haus \ s_r] \] is defined iff there exists a unique house in the situation \(g(r)\),
if defined, \([[s' \ Haus \ s_r] = \iota x[[\text{house}][x](g(r))]]\) (i.e. unique house in the situation \(g(r)\))

- For the utterance to be felicitous, in a given situation there has to be a unique house.

Proposal:

- In Medieval French the LF of the \(l\)-series was simplified from strong to weak: the individual
pronoun disappeared (situation arguments removed to clarity).

(22) \[[i \ [[D_{\text{strong}}] \ NP]] \rightarrow [[D_{\text{weak}}] \ [NP]]\]
- Accordingly, the semantics of the \(l\)-series was reanalyzed:

(23) \(\lambda P_{<e,t>} . \lambda y . \exists! x[P(x) \& x = y] \rightarrow \lambda P_{<e,t>} . \exists! x[P(x)] . \iota x[P(x)]\)

- With the disappearance of an individual pronoun, disappears the requirement that there be a
salient individual with the nominal property in a given situation.

- The “new” requirement is that there be a unique individual with the relevant nominal property
in a given situation, not necessarily salient. Roughly, an antecedent requirement is replaced by
a uniqueness/maximality requirement.

- How can we check this hypothesis?

5 Distributional blueprints of strong and weak definite articles

- Let us assume that there is a pragmatic pressure to use an element triggering a stronger
presupposition whenever its presupposition is satisfied (cf. Maximize presupposition of Heim

- Then the pressure to use the \(l\)-forms would arise in different contexts depending on which
presupposition it carries, “strong” or “weak”.

5.1 The strong \(l\)-series

- The pressure to use the strong \(l\)-forms would arise only in contexts providing an antecedent
for the DP in question, that is, with (some) individual-denoting nouns and relational nouns
relating the referent of the DP to an antecedent.

- The weaker alternative – bare noun – would not be blocked from the contexts satisfying unique-

5.2 The weak \(l\)-series

- The pressure to use the weak \(l\)-forms would arise in all contexts satisfying uniqueness/maximality
presupposition, which presumably includes abstract (amour “love”) and mass (argent “money”)
nouns.
5.3 Empirical predictions

With a change from strong to weak semantics, the \( l \)-forms are expected to expand onto the contexts typically satisfying the uniqueness/maximality presupposition and previously left to the zero article.

- The \( l \)-series is predicted to appear at the maximal rate with individual and event denoting nouns and possibly relational nouns from the earliest periods on.
- The \( l \)-series is predicted to appear at a growing rate with noun phrases satisfying the uniqueness/maximality presupposition: abstract and mass nouns.

6 Testing the predictions

6.1 Methodology

- Corpus of the project “Modéliser le changement : les voies du français” (MCVF) and Penn Supplement to MCVF (2010).
- Total of 32 parsed texts from the X to XVI cc. (\( \approx 1 \text{mln words} \)).

```
IP-MAT-SPE
    NP-SBJ   NEG   CL-PP   VJ   ID
      |       |       |     |
NCS  n'     CL   luist  (1100-ROLAND-V.78.951)  DATE 1100
      |       |     |
Soleill  i
```

*The sun does not shine there.*

We classified \( \approx 4000 \) nouns (orthographic types) in the corpus (\( \approx 25\% \) of total) into the following classes:

<table>
<thead>
<tr>
<th>Noun class</th>
<th>e.g.</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. abstract</td>
<td><em>amertume</em> “bitterness”</td>
</tr>
<tr>
<td>b. mass</td>
<td><em>ambre</em> “amber”</td>
</tr>
<tr>
<td>c. individuals and events</td>
<td><em>bataille</em> “battle”, <em>apparition</em> “apparition”</td>
</tr>
<tr>
<td>d. relational</td>
<td><em>anceisur</em> “ancestor”, <em>ambasadour</em> “ambassador”</td>
</tr>
</tbody>
</table>

6.2 Results

- We estimated the probability of having various determiners in samples of NPs corresponding to a–d classes above.
  - We limited ourselves to the subject position.\(^3\)

\(^3\)Prepositional phrases and predicative positions should be considered apart because even in Modern French they oftentimes license bare nouns. Direct objects may be part of idiomatic constructions with light verbs which may not require a determiner until Modern French.
We also excluded NPs with quantificational and possessive modifiers, since they may interact with determiners’ presence. Ended up with \( \approx 1000 \) tokens.

**Figure 1:** Determiners with abstract nouns

**Figure 2:** Determiners with mass nouns

- Figures 2 & 3 are as expected given the hypothesis that the passage to the new weak semantics entails the expansion to the maximality-satisfying contexts.

**Figure 3:** Determiners with individual/event nouns

**Figure 4:** Determiners with relational nouns

- Figure 3 requires interpretation: let us assume there are contexts a) with antecedents and b) without.
  - In the former we expect to find the strong \( l \)-forms at the highest rate from the beginning.
– The latter divide again in two: b-i) satisfying and b-ii) not satisfying uniqueness/maximality. In the former we expect to see the rise of the new weak \( l \)-forms.

– On the assumption that the rate of contexts with antecedents (a) stays the same over time, we expect to see a rise which is due to contexts (b-ii) without antecedents but satisfying maximality. This is tentatively what we see in Figure 4.

- Figure 4: the frequency of the \( l \)-forms is almost maximal and does not change. If these contexts had been considered by the grammar as uniqueness- but not antecedent satisfying, we would have expected a gradual rise of the (weak) \( l \)-forms, contrary to the fact. Preliminary conclusion: such contexts were analyzed as (indirect) antecedent-providing.

To check whether the differences observed are due to a nominal class difference or a random variation in data, we estimate the distribution of the variable \textsc{Determiner} with values \textit{yes} and \textit{no} in our four samples:

\[
P(\text{Determiner} = \text{yes}|\text{Date} = d) = \frac{e^{\alpha + \beta \cdot \text{Date}}}{1 + e^{\alpha + \beta \cdot \text{Date}}}
\]

![Figure 5: Determiner presence with different noun classes](image)

Regression parameters:

<table>
<thead>
<tr>
<th>Noun class</th>
<th>( \alpha )</th>
<th>( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{Relational}</td>
<td>7.4411</td>
<td>-0.0038</td>
</tr>
<tr>
<td>Individual/Event</td>
<td>-1.947</td>
<td>0.0031</td>
</tr>
<tr>
<td>Abstract</td>
<td>-4.563</td>
<td>0.0039</td>
</tr>
<tr>
<td>Mass</td>
<td>-4.768</td>
<td>0.0044</td>
</tr>
</tbody>
</table>
7 Conclusions and discussion

- We hypothesized a passage from strong to weak definite semantics in the history of the \( l \)-forms in French and showed that the attested evolution of the distribution of the \( l \)-forms is compatible with this hypothesis:

- \( l \)-forms rise in frequency in contexts satisfying the maximality presupposition, but not in contexts providing an (indirect) antecedent.

The nature of the reanalysis

What made the change possible? One suggestion:

- The semantics in (18) delivers the antecedent-anchored (directly referential) interpretation demonstratives and strong articles are known for.\(^4\)

(25) Every father dreads that moment. (A specific context-given moment)

- However, it was also noticed that under some conditions DPs with distal demonstratives and strong articles stop being antecedent-anchored \((\text{e.g. Neale 1993, Dever 2001})\).

- This mostly happens in the context of restrictive relative clauses \((\text{Powell 2001, Simonenko 2015})\).

(26) Every father dreads that/the moment \([\text{when his eldest child moves out}]\). (For every father there is a different moment)

To accommodate both readings, (18) can be modified as follows.

(27) \([D_{\text{strong}}] = \lambda s . \lambda P_{<s,<e,t>} . \lambda Q_{<s,<e,t>} : \exists ! x[P(x)(s) & Q(x)(s)] . \iota x[P(x)(s) & Q(x)(s)]\)

- If the second argument is filled by a singleton property denoted by a type-shifted pronoun \((\text{Partee 1987}) - \text{the property of being identical to the referent of the pronoun} - \text{a directly referential reading arises.}\)

- If the second argument corresponds to the denotation of a restrictive relative, the reading is quantificational \((\text{Simoenno 2015})\), characteristic of “weak” definites, (26).

- In the latter case the presupposition of the DP does not carry the requirement that there be a \textit{salient} individual with a nominal property, since there is no pronominal element that the assignment function has to interpret.

(28) \([\text{RelativeClause } [D_{\text{strong}} \ NP]] \rightarrow [D_{\text{weak}} \ [NP \ \text{RelativeClause}]]\)

(29) \(\lambda P_{<e,t>} . \lambda y . \exists ! x[P(x) & x = y] . \iota x[P(x) & x = y] \rightarrow \lambda P_{<e,t>} . \exists ! x[P(x)] . \iota x[P(x)]\)

- The presupposition therefore boils down to a maximality requirement, as in the case of weak articles, which leads to the expansion in use.

- Recall that \(\approx 40\%\) of all the uses of \textit{ille} are with relative clauses.


9
A more intuitive interpretation: That distal demonstratives and articles descending from them allow for an additional modifier in their structure captures the informal intuition that recovering a distal (in time and/or space) referent requires providing additional descriptive content. Furthermore, they can “lose” the pronominal argument because, by virtue of the distal component in their meaning, the referent is normally not present in the immediate context of the utterance.

<table>
<thead>
<tr>
<th></th>
<th>Late Latin</th>
<th>Old French</th>
<th>Preclassical French</th>
<th>Modern French</th>
</tr>
</thead>
<tbody>
<tr>
<td>direct anaphora</td>
<td>ille</td>
<td>li</td>
<td>le</td>
<td>ce</td>
</tr>
<tr>
<td>relative clauses</td>
<td>ille</td>
<td>li</td>
<td>le</td>
<td>le/ce</td>
</tr>
<tr>
<td>relational nouns</td>
<td>Ø</td>
<td>li</td>
<td>le</td>
<td>le</td>
</tr>
<tr>
<td>mass nouns</td>
<td>Ø</td>
<td>Ø</td>
<td>le</td>
<td>le</td>
</tr>
<tr>
<td>abstract nouns</td>
<td>Ø</td>
<td>Ø</td>
<td>le</td>
<td>le</td>
</tr>
</tbody>
</table>

Figure 7: The suggested global cycle of the definite articles, cf. Harris (1980), Mulder and Carlier (2006)

References

Déchaine, Rose-Marie, Monique Dufresne, and Mireille Tremblay. 2016. Determiners and feature hierarchy in Old French. A talk given at the New Ways of Analysing Syntactic Variation (NWASV 2) at the University of Ghent, 19–20 May.


Penn Supplement to MCVF. 2010. Penn Supplement to the MCVF Corpus by Anthony Kroch and Beatrice Santorini.


Synchronic theory and semantic change
Maximization and Middle English *which*-relatives

Nikolas Gisborne and Robert Truswell
n.gisborne@ed.ac.uk, rob.truswell@ed.ac.uk

Konstanz, 12/9/16
We are interested in the emergence of headed $wh$-relative clauses.

There appears to be a robust pathway from correlative to free relative, to nonrestrictive headed relative, to restrictive relative.

Diagnosing restrictiveness is fraught with problems in purely textual data.

In this paper, we draw on insights from formal semantics to establish a distributional diagnostic for nonrestrictive relative clauses in Middle English.

This allows us to correlate the pathway with distributional evidence.

Our case study today is $which$. 
Roadmap

1. Distributional evidence for semantic change
2. Relative clause types
3. English: 3500BC–1500AD
4. Synchronic semantics to the rescue
5. Conclusions
Section 1

Distributional evidence for semantic change
Collocations and meaning

- The grammaticalization literature (e.g. Traugott & Dasher 2002) is exercised with data like (1).

(1)  
  a. I am going to London (to marry Bill).  
  b. I am going to marry Bill.  
  c. If interest rates are going to climb, we’ll have to change our plans.  
  d. *If interest rates will climb, we’ll have to change our plans. (Hopper & Traugott 2003)

- marry Bill is not a place you can go to.  
- interest rates are not the kind of things that can go.  
- So we know that the meaning of go has changed.
What collocations are good for

- Collocational evidence is often able to diagnose primary grammaticalization.
  - Spatial motion $\rightarrow$ (abstract) temporal motion
- Wider set of collocates $\rightarrow$ loss of semantic selectional restrictions $\rightarrow$ bleaching.
- Not all semantic change works like this.
- Secondary grammaticalization may have little direct collocational evidence.
  - Demonstrative $\rightarrow$ definite article
- And yet, distributional evidence is all we have in diachronic semantics.
  - Obligatoriness of article
- The challenge is to relate distributional changes to denotational changes.
Section 2

Relative clause types
Free vs. headed relatives

- A **free relative** is a clause with the external distribution of an NP.
- A **headed relative** is a clause that modifies a noun.
- Both are syntactically subordinate.
- A headed relative can be introduced by an inflecting phrase (a **relative specifier**), an uninflecting particle (a **relative complementizer**), both or neither.

(2)  

\[
\begin{array}{c}
\text{a. The food} & \text{\(\emptyset\)} & \text{which} & \text{she ate} \\
\text{that} & \text{which that}
\end{array}
\]

\[
\text{b. What she ate}
\]

- Indo-European relative specifiers tend to be formed from **demonstratives** or **interrogatives**.
Restrictive vs. nonrestrictive headed relatives

- A **restrictive** relative denotes a property which modifies a nominal property.

  (3) The person who left: $\nu x. [\text{person}'(x) \land \text{left}'(x)]$

- A **nonrestrictive** relative denotes a proposition containing a discourse anaphor.

  (4) The person, who left: $\nu x. [\text{person}'(x)] \bullet \text{left}'(y)$

- A discourse anaphor needs an accessible antecedent (Evans 1980, Sells 1985) $\rightarrow$ nonrestrictive relatives cannot modify opacity-inducing quantifiers.

  (5) *No person, who left
Maximization and free relatives

- English free relatives are definite descriptions (Jacobson 1995), and therefore maximizers.

(6) I ate [what he cooked].

- Two factors can obscure this, but not invalidate it:

(7) I eat [what he cooks].

  2. -ever can indicate ignorance or indifference (von Fintel 2000) regarding the referent of the free relative.

(8) I will eat [whatever he cooks].

Standard analyses of both treat the free relative as a definite description within the scope of a quantifier over situations or worlds.
The *wh*-phrase in English nonrestrictive relatives is a discourse anaphor (Sells 1985).

Discourse anaphors are maximizing (Evans 1980).

This yields contrasts like (9) (Sells 1985: 19).

(9)  
   a. Each farmer owns some sheep, which the State buys in the Spring. (→ state buys all the sheep)  
   b. Each farmer owns some sheep that the State buys in the Spring. (→ state may not buy all the sheep)

So free relatives and nonrestrictive relatives both involve maximization, but in different ways.

Free relative: maximal individual.  
Nonrestrictive relative: proposition about maximal individual.

We are looking at a free > nonrestrictive pathway.
Section 3

English: 3500BC–1500AD
Context

- The emergence of headed *which*-relatives is part of a wider set of changes in English:
  - Old English demonstrative relative constructions abruptly disappear.
  - *Wh*-forms are gradually co-opted in their place.
- Free relatives provided the source for headed *wh*-relatives (Truswell & Gisborne 2015).
- It is tempting to attribute the emergence of headed *wh*-relatives to the loss of demonstrative relatives.
- However, *wh*-relatives have emerged in other Germanic languages without anterior loss of demonstrative relatives.
- The *wh*-relative strategy emerges repeatedly across the Indo-European family.
- We can understand this better by tracking the history of *wh*-forms, rather than the history of relative clauses (Gisborne & Truswell 2016).
Prehistory: Early IE correlatives

- English *wh*-forms and cognates are descended from PIE \(k^w i-/k^w o-\).
- Original functions: probably interrogative and (restricted) indefinite (e.g. Belyaev & Haug 2014).
- Belyaev & Haug: bipartite asyndetic conditional structure + *wh*-indefinite \(\sim\) correlative.

(10) \[kuiš=an=šan \text{ EGlR-pa tornai} \ n=an \]
\[WH=him=PTCL\text{ back lets PTCL=him} \]
šakuwanzi
they.imprison
‘If anyone lets him back, they will imprison him.’ \(\sim\)
‘Whoever lets him back, they will imprison him.’
(Garrett 2008, conditional ‘back-formation’ ours)

- Early IE did not have embedded relatives (Clackson 2007); later headed *wh*-relatives descend from structures like (10).
Correlatives are rare (< 3% of languages in Dryer 2013) and overrepresented in IE (De Vries 2002).

Correlatives with interrogative forms are even rarer.

Headed *wh*-relatives are just as rare.

<table>
<thead>
<tr>
<th></th>
<th>IE</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Wh</em>-RC</td>
<td>19 (47.5%)</td>
<td>3 (2.3%)</td>
</tr>
<tr>
<td>Other</td>
<td>21 (52.5%)</td>
<td>129 (97.7%)</td>
</tr>
</tbody>
</table>

Table 1: Headed *wh*-relatives in 172 languages (based on De Vries 2002)

We’re investigating a secondary grammaticalization pathway which recurs across IE but only very rarely in other languages.
On contact

- Comrie (1998): *wh*-relatives are a European, not an IE phenomenon.
  - Also attested in neighbouring unrelated languages.
- However, fine details of varieties in contact are rarely similar.
  - Middle English vs. Medieval French (Sakalauskaite 2016).
  - Early Modern Icelandic vs. ENHG (Youmerski 2016).
- Plausible contact situations aren’t always in evidence.

  (11) de fout *wie* hun eigenlijk maken
  the mistake who they actually make
  ‘the mistake which they actually make’
  (Johan Cruyff, via Boef 2012)

- So contact can’t explain everything.
- (See also Poplack et al. 2012 on French P-stranding.)
Universal $\leadsto$ definite wh-correlatives (Belyaev & Haug 2014);
Loss of multiple correlatives (unattested in English written record);
Generalization from clause-initial $\leadsto$ clause-peripheral position.
By the start of the written history of English, correlatives have morphed into left-dislocated free relatives + resumption.
OE free *wh*-relatives
Clause-initial, generalizing, *swa* obligatory

(12) [Swa hwylc eower swa næfð nane synne on
So which you.GEN.PL so NEG.have no sin in
him], awyrpe se ærest ænne stan on hy
him, cast.out.SBJ he first one stone on her
‘He that is without sin among you, let him first cast a
stone on her.’ (coaelhom,+AHom_14:214.2117, c.990)

(13) Soðlice [swa hwar swa Israhela bearn wæron], þar
Truly so where so Israel’s children were, there
wæs leohunt.
was light
‘all the children of Israel had light in their dwellings.’
(cootest,Exod:10.23.2788, c.1050)
OE free *wh*-relatives

Clause-final, optionally generalizing, *swa* optional

(14) Fyres  
    gecynd is þæt hit fornymð  [swa hwaet swa  
Fire.\textsc{gen} nature is that it consumes so what so  
him \textsc{dat} near \ is  
‘Fire’s nature is that it consumes whatever is near it.’

    (cocathom1,+ACHom\_I,\_22:360.152.4446, c.990)

(15) Gemyne,  [hwaet Sanctus Paulus cwæð]  
    Remember what Saint Paul said  
‘Remember what Saint Paul said.’

    (cogregdC,GDPref\_and\_3\_[C]:15.207.28.2739, c.1075)

- Presence of *swa*, not position, determines interpretation.
- *Swa* $\approx$ -ever (Truswell & Gisborne 2015).
- OE free relatives are definite descriptions, as described above.
Latent structural ambiguity

- Clause-final definite free relatives could in principle be used appositively.

(16) \[ \ldots \text{NP}_i \ldots \text{FR}_i \]

- This permits the following reanalysis.

(17) \[ \ldots [\text{NP} \ldots t_i] \ldots \text{RC}_i \]

(18) Þa cwæð ic to him, æteowe me [þa byrigeles [hwar ic þe
Then said I to him show me the tomb where I you
leigde]].
laid
‘Then I said to him, “Show me the tomb where I laid you”.’

Se Hælend me þa beo þære rihnhand genam and me ut
The Saviour me then by the right hand took and me out
lædde [hwar ic hine byrede]
led where I him buried
‘The Saviour then took me by the right hand and led me out to
where I buried him’ (conicodC,Nic_[C]:149.161–2,c.1150)
Early Middle English free relatives

- Various aspects of the OE free relative system disintegrated in early ME.
- *Which*-FRs almost never occur with explicit indicators of generality (*se, ever*) after 1200.
- Bare *which*-FRs can be interpreted as generalizing.

(19) a. beo he hwuch-se eauer beo be.SBJ he which-so ever be.SBJ ‘whichever he may be’ (cmhali-m1,152.352)

b. Bo wuch ho bo (OwlNight,116.1378.751)

- *What*-FRs behave much as in OE through Middle and Early Modern English: generalizing with *se* or *ever*; often definite without.

- In other words, *which* largely leaves the FR system before entering the HR system. *What* apparently specializes as free relativizer in its absence.
Early Middle English headed relatives

- Demonstrative relatives largely disappeared with the collapse of case inflection c.1100.
- But *wh*-relatives weren’t a direct replacement (Gisborne & Truswell 2016).
  - *where* and *there* coexisted for c.200 years.
  - Argumental *se*-relatives disappeared 100 years before argumental *wh*-relatives emerged.
- The first *wh*-relatives emerged in the low-frequency, low accessibility shadows, c.1150.
- Headed relatives with *which* followed c.1350, then *whom* (c.1400), and *who* (c.1500).
- All of this coexisted with stable, high-frequency relativization with *that* and ∅.
Demonstrative and interrogative relatives over time

Proportion of relatives with filled Spec, YCOE + PCMEP + PPCME2

Red = *wh*-rels, NP gaps; Blue = *wh*-rels, PP gaps.
Early *which/whom/who*-relatives

(20) he is emperour of him-zelue. ðet is of his bodye: and of he is emperor of himself that is of his body and of his herte. [huiche he demþ and halt ine guode payse] his heart which he deems and holds in good weight huerof he deþ his wyl. whereof he does his will

(cmayenbi-M2,85.1658, 1340)

(21) But he [whom God hath sent], spekith the wordis of God but he whom God hath sent speaks the words of God

(cmntest-M3,3,20J.234, c.1395)

(22) This declaryth the Mayster of the storyes [who so lyste this declares the master of the stories who so wants to se it].

to see it

(cmfitzja-M4,A5R.71, 1495)
The first headed relatives are all clause-final.
They all seem nonrestrictive.
This allows for a minimal specification of the reanalysis, in terms of scope of the maximization operator. Restrictive relatives would be further from the source construction, in that they do not involve maximization.

(23) a. $\iota x. (\text{boy}'(x) \land \text{saw}'(j, x))$
    b. $\lambda P. [P(\iota x. (\text{boy}'(x)))](\lambda y. \text{saw}'(j, y))$
    c. $\lambda x. (\text{boy}'(x) \land \text{saw}'(x))$

(There is a change, contra De Vries 2002: appositive relatives denote propositions; free relatives typically denote individuals).

But do we know that they’re all nonrestrictive?
A robust indicator of restrictiveness: only restrictive relatives can occur under opacity-inducing quantifiers.

There are no such examples with *which*-relatives prior to c.1450.

(24) and anone he saw he was in a wylde mounteyne whych was closed with the se nyghe all aboute, that he myght se no londe aboute hym [whych myghte releve hym], but wylde land about him which might relieve him but wild bestes.

But such examples are rare anyway — insufficient data to distinguish real from accidental gaps.
Plentiful low-quality data

➤ Textbook examples of restrictiveness often work like this:

(25)  a. A car which I bought last year . . .  
      b. The car, which I bought last year, . . .

➤ This might suggest that indefinite antecedents correlate in some way with restrictiveness.
➤ They don’t, and given the Kamp/Heim treatment of indefinites, we shouldn’t expect them to.
➤ And intuitions about restrictiveness break down in the face of corpus examples.

(26)  þæ cwæð  ic to him, æteowe me [þæ byrigeles  
      Then said  I to him show  me the tomb  
      [hwar  ic þe leigde]].  
      where I  you laid  
      ‘Then I said to him, “Show me the tomb where I laid you”.’
We need more good data

- There is plentiful evidence for the endpoints of the change:
  - OE wh-relatives are all free;
  - Early Modern English restrictive wh-relatives are well-attested.
- There is a natural series of reanalyses:
  - Free $\rightarrow$ nonrestrictive (clause-final, maximizing);
  - Nonrestrictive $\rightarrow$ restrictive (take nominal antecedents, distinction often unclear).
- But direct evidence for the nonrestrictive-only stage is limited to intuitions and the few examples with opacity-inducing quantifiers.
- We shouldn’t be convinced by the natural story unless it’s supported by more robustly attested data.
Section 4

Synchronic semantics to the rescue
A new generalization

Head nouns in nonrestrictive relatives only

- A relative clause of the form *which N IP* is nonrestrictive.
- The proportion of nonrestrictive *which*-relatives correlates with the frequency of *which N*-relatives.

We will not derive this from first principles, but the following considerations make the generalization natural.

1. *Wh*-phrases in nonrestrictive relatives are maximizing by virtue of being interpreted as discourse anaphors (Evans 1980, Sells 1985).
2. Any ‘head noun’ is interpreted internal to a maximizing relative, and often also pronounced RC-internally (Grosu & Landman 1998).
Internal interpretation of head nouns

- Maximizing relatives: amount relatives, free relatives, some internally-headed relatives, correlatives.
- Grosu & Landman’s generalization: a head noun in maximizing relatives is interpreted internal to the relative.

\[(27) \quad \text{I read the books that there were on the table:} \]
\[\quad \text{‘I read the unique individual composed of } d \text{-many books s.t. } d \text{ is the maximal amount s.t. there are } d \text{-many books on the table.’}\]

- *Books* does dual duty: I read books (RC-external), but also the predicate *books* is one of the restrictors that determine the restrictor of MAX (RC-internal).
- (Grosu & Landman have machinery in place to ensure that *books* need only be interpreted in one position, even if it does two jobs.)
Nonrestrictive relatives and head nouns

Nonrestrictive relatives involve $\text{MAX}$ in a different way, but still use a nominal restrictor in the scope of $\text{MAX}$ (as with other discourse anaphors; Evans 1980, Elbourne 2001).

(28) If a man owns a donkey, he always beats it.
[[always$_{s_1}$ if a man($s_1$) owns($s_1$) a donkey($s_1$)]$s_2$ he man($s_1$) beats($s_2$) it donkey($s_1$)](Elbourne 2001: 250)

(29) I read the books, which were on the table.
I read the books $\wedge$ they books were on the table.
Restrictive relatives and head nouns

- Standard accounts of restrictive relative semantics (e.g. Heim & Kratzer 1998) involve conjunction of predicates.

(30) I read the books that were on the table.
    ‘I read the $x: \text{book'}(x) \land \text{on'}(x, t)$’

- Although nothing goes wrong truth-conditionally if N is also interpreted within the restrictive relative, this is redundant.

(31) ‘I read the $x: \text{book'}(x) \land \text{book'}(x) \land \text{on'}(x, t)$’

- In sum:
  - Head nouns are interpreted inside maximizing relatives (Grosu & Landman 1998).
  - Although nonrestrictive relatives use $\text{MAX}$ differently, we still expect head nouns to be interpreted inside them (Sells 1985, Elbourne 2001).
  - Head nouns inside restrictive relatives are redundant, and so probably not there.
Plentiful high-quality data

Internal realization of head nouns

- Early headed *which*-relatives frequently have a full *which*-NP, not just pronominal *which*.

(32) How Kyng Arthure ʒaf bataile to þe Emperour, [in þe how King Arthur gave battle to the Emperor in the whiche bataile þe Emperoure was slayn].

This reflects the likely source of headed *which*-relatives in free relatives (almost always of the form *which* *N*).

- If the head noun is pronounced RC-internally, it must be interpreted there.
  - E.g. no QR-like mechanism to get *N* out of the RC.
- Therefore pronunciation of *N* within RC implies interpretation of *N* within RC.
  - ... which implies nonrestrictive interpretation.
- RCs without overt head nouns could be restrictive or nonrestrictive.
Rise of restrictive *which*-relatives

**Graph Description**
- **X-axis (Year):** The years range from 1300 to 1900, marked at intervals of 100 years.
- **Y-axis (Frequency):** The frequency scale ranges from 0.00 to 1.00, with intervals of 0.25.
- **Data Representation:**
  - Blue line: Frequency of *which* in negative opaque contexts.
  - Red line: Frequency of *N with which*.
- **Legend:**
  - Frequency of which in negative opaque contexts (blue).
  - Frequency of N with which (red).

**Observations**
- The blue line (frequency of *which* in negative opaque contexts) starts below 0.25 in the early 1300s, rises gradually, and peaks around the early 1700s before declining in the mid-1800s.
- The red line (frequency of *N with which*) starts above 0.25 in the early 1300s, decreases steadily, and stays below 0.25 from the mid-1700s onwards.

**Analysis**
- The graph illustrates the historical trend of the use of *which* in restrictive relative clauses and *N with which* in negative opaque contexts.
Correlation *which N* vs. opacity-inducing quantifiers
Section 5

Conclusions
Diagnosing nonrestrictiveness is easy now

- *Which N* → nonrestrictive.
- But *which N* is visible, unlike restrictiveness.
- And the classification of examples according to whether the *which*-phrase contains a noun is crisp, unlike classifications according to restrictiveness.
The entire pathway is visible

- We now have distributional evidence for each step in a complex series of semantic changes.
  - Erosion of *swa* . . . *swa*;
  - Loss of *which N*;
  - Co-occurrence with *no N*, etc.
Synchronic formal semantics can generate new distributional hypotheses

- There is no common-sense reason to associate presence of N with nonrestrictiveness.
- It is only because of the work of Evans, Sells, Heim, Kadmon, etc. that we can propose this distributional diagnostic.
References I


A semantic solution to the progressive passive puzzle

George Walkden

University of Manchester

Konstanz, 12th September 2016
Structure of this talk

- First I’ll introduce the progressive passive puzzle, and previous attempts at solutions
- Then I’ll talk about the semantics of English BE, particularly ‘agentive BE’
- Finally I’ll present a diachronic scenario that links the two
The progressive passive

Formed with:

▶ a finite form of \textit{be}
▶ then a present participle (\textit{being})
▶ then the past participle of the lexical verb

First attested in the second half of the 18th century (van Bergen, 2013, 182). An early example:

(1) ODE To a PIG, while his Nose \textit{was being bored}.

The puzzle

- The progressive – with **be** as auxiliary – is available from at least Middle English.
- The passive – with **be** as auxiliary – is available from at least Middle English.
- So why is the combination of the two not attested until 1761? (cf. Denison, 1993, 440, Warner, 1995, 533)

(diachronic puzzle – not like the *present perfect puzzle* of Klein, 1992)

Weinreich et al. (1968, 102): “What factors can account for the actuation of changes? Why do changes in a structural feature take place in a particular language at a particular time, but not in other languages with the same feature, or in the same language at other times?”
Previous explanations: functional

- Visser (1963–73, §2158): “the urge, permanently inherent in English as an analytic language, to signal separately every separate shade of meaning, function or connotation”. Local cause: analogy with being + adjective. BUT this construction is not earlier (Denison, 1993, 442).

- Denison (1993, 440): these forms “fill a system gap in the patterning of English verbs”. Local cause: regularization of do-support meant that all verbs with the NICE properties complemented by another verb were auxiliaries except progressive be. BUT do-support is virtually categorical by 1750 before the progressive passive emerges.

- Denison (1998, 149–150): earlier form, the passival (The house was building), with inanimate subject, became ambiguous when the active progressive began to take inanimate subjects. BUT the chronology isn’t right for that (Hundt, 2004a; Kranich, 2010). 91 of 130 early examples of progressive passive have non-human subjects (Hundt, 2004b).
Previous explanations: formal

- Traugott (1972, 178): “an excellent example of simplification by generalization”. BUT why in the 18th century?
- Warner (1995, 1997): reanalysis after which auxiliaries are frozen forms lacking the morphological interrelationships of full verbs. Local cause: loss of *thou* (removing last vestige of inflection on modals) and fixing of *do*-support.
- Cowper & Hall (2012): restructuring of tense and aspect system (two syntactic heads rather than one). Local cause: reanalysis of *-en* as exponent of passive rather than resultativity. Resultative *be*-perfect in decline throughout eighteenth century. (Must assume accidental homophony between past participle *-en* and passive *-en*.)
Towards a new explanation

- Most of these approaches propose a local cause – some more plausible than others.
- Warner and Cowper & Hall also link the change to other syntactic changes occurring at the same time.
- But none of the approaches considers the semantics of either the progressive or of \textit{be} in detail.
- Warner (1997, 162): “Semantic and functional contrasts ... are not central determinants in the same sense as the availability of morphosyntactic categories.”
The semantics of **BE**

- Traditional (Aristotelian/Russellian) approach: **BE** is semantically completely vacuous, a verbal expletive needed to bear tense.
- Partee (1987): $\lambda P. P$ (the identity function)
- However, **BE** carries agentive implications in the progressive (Partee, 1977).

(2) a. Mary is noisy.
   b. Mary is being noisy.
   c. *Mary is being awake.
   d. *Mary is being a murderer.
   e. *The river is being noisy.
   f. *The children are being quiet because they are asleep.

How to account for this?
Two BEs (or not two BEs)

- Partee (1977): there is an additional [+active] BE that is homophonous (cf. also Dowty, 1979).
- Problems (Rothstein, 1999, 360–362):
  - little supporting evidence (stipulative)
  - also doesn’t capture certain parallels and differences between the progressive and small clauses
  - Unlike other verbs, BE does not inherently belong to any of the Vendler (1967) aspectual classes
  - BE + AP may be assigned to stative, activity, or achievement depending on contextual factors
Ingredients of Rothstein’s account

- \( \text{BE} \) packages non-atomic states into atomic eventualities.
  \[ \lambda S \lambda e. \exists s \in S : e = l(s) \]
- \( S \) is a set of states and \( l \) is a locating function.
- Adjectives (before predicate abstraction): \( \lambda s. A(s) \land \text{Arg}_1(s) = x \)
- Verb classes (leaving accomplishment aside):
  - Statives: \( \lambda P \lambda e. P(e) \)
  - Activities: \( \lambda P \lambda e. (DO(P))(e) \)
  - Achievements: \( \lambda P \lambda e. (BECOME(P))(e) \)
- \( \text{BE} + \text{AP} \) is then \( \lambda e. \exists s [A(s) \land \text{Arg}_1(s) = x \land e = l(s)] \)
Stative vs. activity BE

- Example adjective: *obnoxious* (Rothstein, 1999, 404)

- Stative: \( \lambda P \lambda e. P(e) \left( \lambda e. \exists s[OBNOSIOUS(s) \land Arg_1(s) = x \land e = l(s)] \right) \)
  \[
  = \lambda e. \exists s[OBNOSIOUS(s) \land Arg_1(s) = x \land e = l(s)]
  \]

- Activity: \( \lambda P \lambda e. (DO(P))(e) \left( \lambda e. \exists s[OBNOSIOUS(s) \land Arg_1(s) = x \land e = l(s)] \right) \)
  \[
  = \lambda e. (DO(\exists s[OBNOSIOUS(s) \land Arg_1(s) = x \land e = l(s)]))
  \]
The progressive

- Stative verbs are incompatible with the progressive (Dowty, 1979, Rothstein, 1999, 402 among many others).
  
  (Rothstein, 1999, 409–416 discusses the semantic details of this, based on the theory of the progressive in Landman, 1992.)

(3) *John is knowing French.

- But BE can have an activity reading.
The underspecification of *be* is clearly not universal. Other languages don’t permit an activity reading.

(4) Hij is een dwaas aan het zijn.
    he is a fool in the be.INF
    ‘He is being a fool.’

So has the activity reading of *be* in English always been available? Or is it a more recent development?
A diachronic scenario

- I hypothesize that English be originated as a lexically stative verb – as Rothstein argues is the case for other verbs such as resemble – and underwent semantic bleaching during the late 18th century to become underspecified for aspectual class.

- This predicts that the rise of the progressive passive should be concomitant with the rise of other uses of be with an activity reading.

- ...Is this true? How can we test it?
The literature generally agrees that the progressive passive precedes uses of *being* with nominal or adjectival complements (see e.g. Denison, 1993).

Visser (1963–73, 1954) argues that the opposite is true – but his much earlier examples are misanalysed.

However, all this literature is based on the ‘find the earliest example’ school of historical linguistics, without use of modern corpus methods. Is the claim correct?
Corpus investigation

- CLMET 3.0 (Diller et al., 2011): a 34-million-word POS-tagged corpus covering the period 1710–1920.
- Search query: (is|are|am|was|were) being\_VBG followed by manual inspection to root out false positives. (There will be false negatives too.)
- Problem: difficult (both in theory and in practice) to identify the envelope of variation. Not just the passival.
## Results

<table>
<thead>
<tr>
<th>Prog. pass.</th>
<th>Per million</th>
<th>With NP or AP</th>
<th>Per million</th>
</tr>
</thead>
<tbody>
<tr>
<td>1710–1780</td>
<td>1</td>
<td>0.10</td>
<td>0</td>
</tr>
<tr>
<td>1780–1850</td>
<td>68</td>
<td>6.03</td>
<td>1</td>
</tr>
<tr>
<td>1850–1920</td>
<td>638</td>
<td>50.55</td>
<td>5</td>
</tr>
<tr>
<td>Overall</td>
<td>707</td>
<td>20.56</td>
<td>6</td>
</tr>
</tbody>
</table>

(5) You will be glad to hear ... how diligent I have been, and am being.

(CLMET 3.0; Letter from John Keats, July, 1819)

(6) ... that one of their own fraternity is being launched into eternity

(CLMET 3.0; Memoirs of Henry Hunt, 1820–2)
Discussion

- At first it looks as if the hypothesis is false – not only do the earliest examples of \( \text{be} + \text{being} + \text{NP/AP} \) start much later, but the progressive passive is vastly more common than \( \text{be} + \text{being} + \text{NP/AP} \) in the later period.

- But things are not this simple, as \( \text{be} + \text{being} + \text{NP/AP} \) is much less common in present-day English too.

- Random sample of 500 of 17,399 hits for the same query in the British National Corpus yields only 22 examples that are not progressive passives!

- It would be good to test for a Constant Rate in the sense of Kroch (1989), but we need a handle on the other competing variants.
Conclusions

- I’ve proposed that the absence of the progressive passive before Late Modern English was not due to an arbitrary gap in a paradigm, but rather due to the absence of the relevant reading of \( \text{be} \).
- The emergence of the progressive passive would thus be the automatic corollary of a lexical semantic change that permitted \( \text{be} \) to receive non-stative readings.
- Due to the much lower overall frequency of \( \text{be} + \text{being} + \text{NP/AP} \), and the difficulty of finding other variants, it’s hard to say whether this is true or not.
Thanks for listening!


References III


EVOLUTIONARY MODELING FOR SEMANTIC CHANGE

IGOR YANOVICH

TUTORIAL AT Formal Diachronic Semantics (FoDS) AT Konstanz

SEPTEMBER 12, 2016

EXTENDED HANDOUT

1. A POPULATION PERSPECTIVE ON SEMANTIC CHANGE

• Some familiar changes:

  (1) resultative (state after an event)
      ⇒ perfect (an event happened, leading to a state)
      ⇒ past (just a past event, no backshifting)

  (2) be able to (physically)
      ⇒ can (given the circumstances)
      ⇒ may (according to the rules)

  (3) be going to PLACE (directional)
      ⇒ be going to ACTION (future)

• The familiar goals of diachronic semantics:
  – identify the meanings involved
  – describe their presence/absence at different periods
  – examine cross-linguistic data on similar constructions
  – hypothesize pathways of change
  – characterize them in terms of what exactly happens to the meanings

• Doing familiar diachronic analyses, we focus on broad stages.
  Those stages occupy large time periods (usually several centuries each).
  We implicitly take them to be monolithic, uniform.

• This is not a bad practice.
  But we abstract away from the messiness of real-life distributions.

Messiness I: diversity within speaker populations.

• Different speakers at the same time may have different grammars.
As an example, [Devitt, 1989] and [Nevalainen and Raumolin-Brunberg, 2003] trace several changes in Early Modern English at the by-speaker level.

- \( t \rightarrow d \) in Scots English ([Devitt, 1989]): replacement of archaic \(-t\) past forms with London \(-d\) forms
- \( ye \rightarrow you \) ([Nevalainen and Raumolin-Brunberg, 2003]): replacement of archaic nominative \( ye \) with innovative \( you \) (the accusative form)

• Important points:
  
  - Even a “simple” change often takes a long time to happen
  - In the middle of a change, we have people with “mixed” output, but also some nearly-categorical speakers.
  - People with mixed output do not cluster neatly around a central value.

• In principle, change may proceed via misacquisition in childhood or via shifting your language in adulthood.
  - Recently, there’s been a new focus in historical sociolinguistics on changes that happen in the adulthood (e.g. [Raumolin-Brunberg, 2009])

Messiness II: diversity of the present linguistic forms.

• Example from resultative-perfect-perfective:
  
  - [Condoravdi and Deo, 2014]: Early Vedic \(-ta\) only had result readings, while in Late Vedic it could express both result and existential perfect.
  
  \[ \Rightarrow \text{a single form becomes responsible for two types of readings} \]
– [McFadden and Alexiadou, 2010]: Early Middle English *have* and *be* + participle were resultatives, then in Late Middle English *have* expands into existential perfects, but continues to be used in resultatives

⇒ one form starts to cover two types of meanings, but the other one remains restricted

• Natural questions: when the same form covers several meanings, what’s the distribution of those meanings? How does that distribution change over time?

• Example from modals:

<table>
<thead>
<tr>
<th></th>
<th>ability</th>
<th>circumstantial</th>
<th>deontic</th>
<th>epistemic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present-Day English</td>
<td>–</td>
<td>15%</td>
<td>11%</td>
<td>66%</td>
</tr>
<tr>
<td>Early Modern English</td>
<td>6%</td>
<td>35%</td>
<td>6%-12%</td>
<td>17%</td>
</tr>
<tr>
<td>Late Middle English</td>
<td>19%</td>
<td>53%</td>
<td>2%</td>
<td>3%</td>
</tr>
<tr>
<td>Early Middle English</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
</tbody>
</table>

Percentages for Present-Day English from [Coates, 1983]; for Early Modern and Late Middle English, from [Gotti et al., 2002]; presence or absence in Early Middle English, from [MED, 2002].

• The change in the distributions can clearly be a part of the larger change itself.

⇒ It is not enough to think in broad terms of presence/absence of specific meanings.

• **Exercise 1:** can you give an example of changing “percentage distributions” for some change you know well?

**The population perspective.**

• **Empirical questions:**
  – Given a population of speakers, how do the shares of different “speaker types” change over time?
  – Given a “population” of actual utterances with some linguistic items, how do the shares of different alternative items develop diachronically?

• Answering these empirical questions does not **replace** understanding broad stages of change. But it would provide us with more empirical context for the broad-stage analyses.

• **Theoretical questions:**
  – What are the laws governing quantitative trajectories of change?
  – Can we learn about any substantive features of specific linguistic changes from the observed trajectories? I.e. can we falsify some substantive theories?

• Answering these theoretical questions requires special mathematical apparatus.
2. **Mathematical apparatus: population genetics, evolutionary game theory**

- An analogy: to understand natural language semantics, higher-type logics were needed
  - When scholars tried to understand quantifiers with only first-order logic tools, they saw non-compositionality.

  \( \forall x : \text{girl}(x) \rightarrow \exists y : \text{poem}(y) \land \text{wrote}(y)(x) \)

  The contribution of “every” is not a constituent in the formula!

- With higher-type representations available, it is easy to see that 4 is compositional.

  But it took Montague to arrive at that realization. Moreover, the relevant mathematical apparatus needs to be learned by aspiring semanticists.

- Evolutionary modeling for language change:
  - Much useful machinery has already been invented by biologists, economists, and others. (Don’t need to invent the bike.)
  - Naturally, linguistic applications would often have a different shape than in other sciences.
    \( \Rightarrow \) Sometimes we’ll need to develop the apparatus further. (Maybe a new style of bike.)

- Types of apparatus available:
  - **Population genetics.** Very useful, but almost not known among linguists.
  - **Evolutionary game theory.** Very useful, has been adapted for linguistic purposes multiple times.
  - **Agent-based modeling.** Strictly speaking, this is a methodology, not a mathematical framework: a particular type of simulation. Has been widely used in linguistics.

    A particularly popular branch is Iterated Learning, used both in computational experiments and in learning experiments with people. At least in a simple case, mathematically amounts to a well-studied population-genetics model, [Reali and Griffiths, 2010].

- Today, we will mostly discuss population genetics, and a little of evolutionary game theory.
3. Wright-Fisher: an evolutionary model from population genetics

- We will now define a simple population-genetic model. After we learn its structure, we will discuss why that model is applicable not only to biological evolution, but also to different linguistic settings.

- Imagine a population of $n$ individuals, reproducing asexually, thus each child has only one parent.

  Each individual belongs to a particular genetic type. We denote them by $A$, $B$, etc.

  The usual biological setup is to consider different alleles at the same locus: the different versions of the same gene, i.e. distinct nucleotide sequences at the same segment in the DNA. But $A$, $B$ and so on can also represent full genome types, or morphological types, or whatever other inheritable object.

  We write the population state as a vector $x = (x_A, x_B, \ldots)$, where each element records the number of individuals of the corresponding type.

- For simplicity, assume that generations do not overlap. This leads to discrete time: the individuals at generation $t$ are the parents of all individuals at generation $t + 1$.

  Again for simplicity, assume that the population size is constant.

  For example, imagine a species that occupies a limited amount of space that can maximally be exploited by $n$ individuals, and assume that the whole space is always exploited.

- Finally, formulate a probabilistic rule that determines which individuals reproduce, and how many times. A popular rule is: there is $n$ reproduction events per generation, and the parent for each of them is chosen randomly from the current population.

  In effect, the parents of the new generation are drawn with replacement. This results in a variant of the classical model of population genetics called Wright-Fisher model. This variant can be called neutral (we’ll return to this term later.)

  \[(5) \text{ Neutral Wright-Fisher model:} \]
  \[\begin{align*}
  &\text{Population size } n \\
  &\text{Reproduction rule: parents are independently drawn with replacement} \\
  &\text{For two types } A \text{ and } B: \quad x'_A := \text{Bin}(n, \frac{x_A}{n})
  \end{align*}\]

  In words, the number of $A$s in the next generation is binomially distributed, with the size $n$ of the population and the current share $\frac{x_A}{n}$ of $A$ as the parameters. That number $x'_A$ is a random variable, and the binomial distribution as above provides the probabilities that $x'_A = 0$, that $x'_A = 1$, and so forth to $x'_A = n$. 
• The binomial distribution $Bin(n, p)$ is the distribution of the total number of events of interest in $n$ independent trials with probability of such an event $p$.

For us, the trial is selecting a parent for a fixed child. The event of interest is that the parent will be an $A$. The alternative outcome is that the parent is a $B$. We do not care about the identity of the parents, only about their type. It is then clear that population size is the $n$ parameter of $Bin$, and the share of $A$, namely $x_A/n$, is the $p$ parameter.

• Examples of binomial distributions: share of $A$ is 30% in all three graphs, but $n$ differs.

---

• Sample trajectories of the neutral Wright-Fisher model

In these examples, $n = 100$ or 1000, starting state: 50% $B$

---

• Exercise 2:
  – What will happen if we increase $n$? If we decrease $n$?
  – What will happen if we start from a different state? 20% $B$s? 95% $B$s?
  – What if we run the process with large $n$ (e.g. 10000) and a very small percentage of $B$s?
  – What about small $n$ and the same low percentage of $B$s?
• The random force that we just observed is called **genetic drift**.

In a finite population, the percentages of different variants will change just due to the random nature of reproduction. No systematic forces are present, and yet the shares of different alleles change.

In an infinite population, things are different (we can get an intuition about that by observing the changes as \( n \) gets closer to the limit of \( \infty \).) In infinity, all randomness is averaged out, and change trajectories become deterministic. There is no genetic drift in this setting, and if we start with 50% Bs, we’ll always have the same share in the neutral model.

• Now we introduce **natural selection**.

Suppose \( A \) and \( B \) affect the reproductive success of their bearers. This can happen in many ways: a “better” \( B \) can help its bearer to survive to maturity; it may help it to find and exploit better sources of food; it may make its reproduction more efficient, allowing for more reproductive events...

But all those different features in the end interest us only insofar as they affect the number of offspring. We thus lump them together into a single quantitative feature called **fitness**. By definition, the expected number of offspring is proportional to a parent’s fitness. So fitness is a cumulative metric for overall reproductive success.

• In the simple case, each type \( A, B, \text{etc.} \) has fixed fitness \( f_A, f_B, \text{etc.} \).

In a more complex case, fitness depends on the current state of population. For example, it can be profitable to be a \( B \) in a predominantly \( A \) population. Or it could be much better to be a \( B \) when there are many other \( B \)s. Formally, in this setting \( f_A \) is not a constant, but a function from the current population state into numbers.

\[ (6) \text{ Wright-Fisher with fitness:} \]

- Population size \( n \)
- Fitnesses \((f_A, f_B, \ldots)\)
- Reproduction rule: parents are independently drawn with replacement proportionally to their type’s fitness
- For two types \( A \) and \( B \):
  \[ x'_A := Bin(n, \frac{x_A f_A}{x_A f_A + x_B f_B}) \]

• In neutral Wright-Fisher, the second parameter of the distribution in the formula was \( x_A/n \), or the share of \( A \). Note that \( n = x_A + x_B \) for two types.

In Wright-Fisher with fitness, the probability of getting an \( A \) is not just the share of \( A \), as it is proportional to \( A \)’s fitness.
To get an intuitive understanding of formula $x_A f_A + x_B f_B$, consider $f_A = 1.5$ and $f_B = 1$, and a population with equal shares $x_A = x_B$. By proportionality, each $A$ should have a $\frac{1.5}{1.5} = 1$ better chances to become a parent than a $B$. It is equivalent to a situation where the fitness would be equal for $A$ and $B$, but there would be 1.5 times more $A$ than in our population. This means the total population size would be $1.5 x_A + x_B$, and the share of $A$ would be $\frac{1.5 x_A}{1.5 x_A + x_B}$. Generalizing this argument, we get $\frac{x_A f_A}{x_A f_A + x_B f_B}$, where $x_A f_A$ and $x_B f_B$ may be viewed as effective shares of $A$ and $B$.

- Sample trajectories of the Wright-Fisher model with fitness

In these examples, $n = 1000$, starting state: 50% $B$, fitnesses $f_A = 1$, $f_B = 1.01$ or 1.1

- Exercise 3:
  - What will happen if $f_A = 2$, $f_B = 2.02$?
  - What will happen if $f_A = 1$, $f_B = 1.001$ with $n = 100$ (as in the pictures above)?
  - What will happen if $f_A = 1$, $f_B = 1.001$ with $n = 10000$?
  - What will happen if $f_A = \frac{x_A}{n}$ and $f_B = \frac{x_B}{n}$? (fitness positively correlated with shares)
  - What will happen if $f_A = 1 + 0.05 \times \frac{x_A}{n}$ and $f_B = 1 + 0.05 \times \frac{x_B}{n}$?
  - What will happen if $f_A = \frac{x_B}{n}$ and $f_B = \frac{x_A}{n}$? (fitness negatively correlated with shares)

- This simple model already shows the limitations of working with simple trajectories.

  For example, the two pictures above are different. But they come from stochastic processes. How do we know they do not show fluke runs?
• One way to get a real handle on the process’s behavior is to explore large numbers of runs systematically. Here is how to do it in R:

```r
n_gen = 200  
n_runs = 100

runs_fB_101 = lapply(1:n_runs, function(x)  
{tmp=run_n_generations(c(5000,5000), function(x)  
      stochastic_selection_mutation_ConstantSize(x,  
         generate_Q_drift(0,0), function(x) c(1,1.01)), n_gen);  
      print(x); tmp})

length(runs_fB_101) #shows the number of runs  
dim(runs_fB_101[[12]]) #shows the dimensions of matrix representing an arbitrary run

#get the # of B organized as a matrix with rows representing generations,  
#    columns representing runs  
runs_fB_101_Bshares = sapply(runs_fB_101, function(x) x[,2]/10000)

runs_fB_101_averageBs = sapply(1:n_gen, function(i) sum(runs_fB_101_Bshares[i,])/100)  
runs_fB_101_B_sds = sapply(1:n_gen, function(i) sd(runs_fB_101_Bshares[i,]))

dev.new()  
plot(runs_fB_101_averageBs, xlab="Generation", ylab="Average share of Bs",  
     xlim=c(1,n_gen), ylim=c(0,1), type="l", col="darkolivegreen", lwd=2,  
     main="n=10000, fA=1, fB=1.01")  
lines(1:n_gen, rep(0.5,n_gen), col="gray80")  
lines(runs_fB_101_averageBs+runs_fB_101_B_sds, col="darkolivegreen1", lty=3, lwd=2)  
lines(runs_fB_101_averageBs-runs_fB_101_B_sds, col="darkolivegreen1", lty=3, lwd=2)
```

• To simplify this, I defined a single function that draws such plots given the evolutionary parameters:

```r
compute_and_display_average_Bshare(in_state = c(5000,5000),  
    fitness_function=function(x) c(1,1.1), main="n=10000, fA=1, fB=1.1")

#and for fB=1.01 and small n=100:
compute_and_display_average_Bshare(in_state = c(50,50),  
    fitness_function=function(x) c(1,1.01), main="n=10000, fA=1, fB=1.01")
```
• In the models so far, new types could not arise. Once the population lost all As or all Bs, it will never get them again.

But biological evolution — and linguistic evolution, too — is not just the story of loss. It is also the story of innovation: new types entering the population.

The process bringing in new variants is called **mutation**.

Note a terminological complication: in the molecular setting, mutation is a change in some nucleotide in the DNA, but in population-genetic models, it is simply a change from one type to another.

• When our types are alleles (recall that this is how variants of the same gene are called), there is a very large number of variants just one nucleotide change away from any A. Moreover, the rate of molecular mutation is so small that on relatively short time scales, it’s very improbable that the same nucleotide in the sequence will be hit by a mutation the second time.

That is why for biologists, popular models assume that a gene has infinite length, and therefore mutation cannot create the same new types twice, and it cannot undo an earlier mutation.

For the linguistic case, this is an inappropriate model. So from the start, we will consider a different type of model, where both forward and backward mutations are possible.

• Introducing mutation:
  - Without mutation, parent’s type $X$ is faithfully copied by the child.
  - If mutation happens, the child’s type changes to some $Y$.
  - Mutations happen with a constant probability. $r_{A\rightarrow B}$ is the probability that an $A$ parent gives rise to a $B$ child, and similarly for $r_{B\rightarrow A}$.
  - The probability that a mutation does not happen when an $A$ parent reproduces is $1 - r_{A\rightarrow B}$, trivially.
(7) Wright-Fisher with fitness and mutation:

- Population size $n$
- Fitnesses ($f_A, f_B, ...$)
- Mutation rates $r_{A \rightarrow B}, r_{B \rightarrow A}, ...$
- Reproduction rule: parents are independently drawn with replacement proportionally to their type’s fitness; with corresponding probabilities, their children get a mutation
- For two types $A$ and $B$:

$$x'_A := \text{Bin}(n, \frac{x_A f_A}{x_A f_A + x_B f_B} \times (1 - r_{A \rightarrow B}) + \frac{x_B f_B}{x_A f_A + x_B f_B} \times r_{B \rightarrow A})$$

To derive the formula above, we go by cases. First, the parent of a random child could be an $A$ or a $B$. This happens with probabilities $\frac{x_A f_A}{x_A f_A + x_B f_B}$ and $\frac{x_B f_B}{x_A f_A + x_B f_B}$, respectively. Second, conditional on the parent being an $A$, the probability that the child will be also an $A$ is $1 - r_{A \rightarrow B}$: the probability that an out-mutation out of $A$ did not happen. Conditional on the parent being a $B$, the probability that the child is an $A$ is $r_{B \rightarrow A}$, the probability of a mutation into $A$ from $B$. Together, this gives us the probability that a single child will be an $A$: it is $\frac{x_A f_A}{x_A f_A + x_B f_B} \times (1 - r_{A \rightarrow B})$ (the case of an $A$ parent) plus $\frac{x_B f_B}{x_A f_A + x_B f_B} \times r_{B \rightarrow A}$ (the case of a $B$ parent). This probability is the second parameter of the binomial distribution characterizing $x'_A$.

- Let’s see what happens when fitness is equal for $A$ and $B$, but mutations are asymmetric:

  ![Graphs showing population dynamics with different mutation rates](image)

- Exercise 4:
  - Let $r_{A \rightarrow B} = 10^{-3} = 0.001$, $x_A = n = 100$, fitnesses be equal. What is the probability that there will be at least one $B$ in the next generation?

    Hint: it will be hard to actually compute the number, so go for a formula instead.

  - Assume $r_{A \rightarrow B} = 10^{-3} = 0.001$, $r_{B \rightarrow A} = 10^{-4} = 0.0001$, fitnesses are equal, $n = 100$, $x_A = 99$, $x_B = 1$.

    Write down the probability that in the next generation, there won’t be any $B$, in terms of a binomial random variable as in 7.
• For the conditions we just saw in the exercise, the probability that one $B$ will get eliminated in the next generation is $\approx 33\%$. For two $Bs$, that probability is $12\%$.

For comparison, in a drift-only process those probabilities are $37\%$ and $13\%$. The presence of a mutation bias towards $B$ makes it easier for an invading $B$ to survive.

• Here is how mutational asymmetry may look like when averaged. Fitness asymmetry graph also given (on the right). The solid line is the average number of $Bs$ across 100 simulation runs. The dotted line is at one standard deviation on each side of the average line. Individual dots show the position of individual runs in the state space.

• Let’s look at how our model departs from the all-$A$ state under different mutation rates:
• Consider \( x_B = 1, x_A = 99 \). This is a typical situation that can arise if the system starts in the all-\( A \) state.

When there is no mutation, the chance that \( B \)s will get eliminated at the next generation is \( \approx 37\% \). When \( r_{A \rightarrow B} = 0.001, r_{B \rightarrow A} = 0.0001 \) (as in the lower-row picture above), that probability is \( \approx 33\% \). When we proportionally increase the mutation rates to \( r_{A \rightarrow B} = 0.01, r_{B \rightarrow A} = 0.001 \), the probability that the \( B \) will get eliminated immediately is just \( \approx 14\% \).

• This illustrates the balance between the directional forces of asymmetric mutation and differential fitness, and the non-directional force of genetic drift.
  – Genetic drift always favors pure states — states where there is only one type.
  – Mutation creates a possibility to get out of a pure state.
  – After an initial mutation happened, if mutation and fitness are weak, the process is largely driven by drift. But if at least one of them is strong, the process develops directionally.

• We can see the effects of this in the average share of \( B \) plots, as before:

![Graphs showing the average share of B plots](image)

• Note how the dots in the plots above indicate the process’s bias towards pure states. That bias is stronger in the graph on the right, where mutation is slower.

• We can also see the same effects if we plot several trajectories of the share of \( B \)s from individual simulated runs (they come from the same simulation as the graphs above):
Again, once the process takes off from the pure state, it moves erratically, not obviously directionally. This is one effect of genetic drift. At the same time, the pure states are favored, attracting the trajectory (the attraction is stronger on the right graph.) This is another effect of genetic drift.

- To finish our discussion of the mechanics of the Wright-Fisher model from population genetics, we look at what happens when the population is infinite.

We already mentioned that with infinite populations, all randomness averages out, and instead of stochastic trajectories, we have deterministic ones. For the two processes above driven by mutation bias, those deterministic trajectories are shown below.

Compare these two with the trajectories in the previous two pairs of graphs: the evolutionary parameters are the same for them, except for the population size.
You will notice that the average share of Bs in the finite-population process develops similarly to the B share in the infinite-population, deterministic process. But the variability is much greater in the finite-population process.
4. Linguistic interpretations of population-genetic models

- Evolutionary models in population genetics, like the simple Wright-Fisher model that we just saw, were formulated for biological evolution. But they are actually pure mathematical objects. The biological interpretation is not the only one for them.

The model we studied can be applied to any process where we have reproduction, or copying, between elements from different time slices (=generations), effects of differential proneness to be copied (=fitness), and copying with mistakes/changes (=mutation). Those three elements create a situation where evolutionary models of the above type may be useful.

Of course, the model above also has more specific assumptions — for example, it presupposes that generations are non-overlapping and that there is no geographical structure in the population. Those assumptions will not be met by most biological cases. But that is not to say that the model should be discarded. Instead, biologists do the following: they study the behavior of a simple model first, and then build a more complex model whose behavior they can then compare to the simple model’s behavior. Sometimes it will turn out that there is not much analytical gain from going to the more complex model. Other times the behavior of the simple and complex models will be so different that the simple one will not be practically useful. But the bottom line is that by itself, the fact that some model’s assumptions do not match our understanding of the reality is fine, as long as we keep it in mind.

- Linguistic interpretation 1A:
  - Evolution operates on speakers’ grammars. Namely, one individual in the model is the grammar (type) of one speaker.
  - Grammar is selected during language acquisition.
    * If we simplify and assume that each speaker has a single privileged “linguistic parent” — an individual from the previous generation with most influence, for example — then the child by default copies that parent’s grammar type.
    * We can also make the model a bit more complex and work with multiple linguistic parents for each child (say, one hundred). Modeling can then be guided by biological models for diploid organisms (like humans; we have two parents, and we inherit two copies of most genes, one from each parent).
  - Different grammar types have different potential to reproduce. For example, some are more functionally beneficial (e.g., induce less cognitive load), and thus better than others. As another example, some can be more fashionable because of the social
class they are currently associated with. We capture those heterogeneous sources of

differential reproductive potential through fitness.

– Sometimes a child will innovate, acquiring a grammar type they never observed. This

is captured by mutation.

⇒ This real-world interpretation of reproduction, fitness and mutation is different from the

one in biology. But it fits the mathematical model.

We are not saying that linguistic evolution is just like biological evolution. We are only

saying that the mathematical model can be applied (bearing in mind that it is a simplified

conceptualization of the reality of language change).

In fact, for linguistics we need to use a different subtype of that model than in biology.

When the model is about genes, there are almost infinitely many different alleles, and

mutation will almost never create the same variant again. In linguistics, there are only that

many grammar types. Therefore the parametrization of mutation should be different than

in biology, but it still fits the general framework of population-genetic models.

• **Linguistic interpretation 1B:**

  – Evolution units are speaker grammars at a time slice. Thus each real-world speaker

    corresponds to a sequence of grammars going through many “generations” of the model.

  – Grammar changes not only in childhood, but also in adulthood.

  – The linguistic parent of the current time slice of some speaker is her own previous

    time slice. By default, the grammar of the previous time slice is just maintained. But

    sometimes, the speaker switches to the grammar by some other speaker.

  – Fitness determines how beneficial different grammars are. The beneficial ones will be

    copied by other speakers more frequently.

  – Sometimes, a speaker may also change the grammar not by copying somebody else’s,

    but by internal innovation. This is mutation.

• **Linguistic interpretation 2** (“usage-based”, cf. [Croft, 2000]):

  – Units of evolution are tokens of linguistic features from actually made utterances.

  – Time is discretized: utterances from, say, the same day are treated as one generation.
– Tokens of linguistic features get stored in speakers’ memory, and when new utterances need to be made, speakers select for each feature one exemplar from the memory and copy it into their production. However, we do not model speakers’ memories directly in this simple model: the population of the model models the combined memory of all speakers in the speech community.

– Some features are more prone to be copied, which we capture through fitness. Again, the reason may be comparative cognitive benefits, social prestige, etc.

– When creating a new copy, the speaker may misproduce, thus changing the type of the copied feature. Additionally, when a produced feature is perceived, the hearer may misanalyze it, and thus store in the memory a different type than the one the speaker copied from. In either case, we have mutation.

• The bottom line: there are different reasonable linguistic interpretations of the model we studied today. The model is therefore useful.

• A possible pipeline:

  (1) Identify the variants involved in your linguistic change of interest.

  (2) Think about fitness and mutation in your case.

  At the least, make an educated guess about the reasonable ranges for their values.

  To do it properly, one would try to estimate those values from data: from experiments, from corpora. This is not trivial, and no work of this type has yet been done, but we should get there.

  (Last academic year, I have developed the preliminaries for an inferential framework that would allow me to estimate evolutionary parameters from texts. It is highly non-trivial, but I hope this academic year, I’ll be able to actually do some inference of this type.

  (3) Study the behavior of your model under those realistic ranges of parameters.

But what exactly does “study” mean in this context? Let’s discuss the options.
5. Different goals of evolutionary modeling

• **Goal 1: Proof of concept modeling.**

  Suppose you have a general theory of how some linguistic change works: what are the meanings, what are the available switches between states, how likely are different switches on the individual level, etc.

  You can then write your assumptions into an evolutionary model, and examine its behavior. Then you’ll have one of the two cases:

  – The desired evolutionary behavior emerges. This means your theory passes the proof-of-concept check.
  
  – The desired evolutionary behavior fails to emerge. This falsifies your idea. Either your theory of that change is wrong, or your evolutionary model is not a good fit for the actual process.

  ⇒ The value of modeling is the potential to falsify the substantive theory of the change.

• Example: [Deo, 2015] on the imperfective-progressive cycle

  Deo works with an infinite population of units that represent speaker grammars of four types. Fitnesses of the four types are given by a game-theoretic analyses of communication between speakers. E.g., how well would the speaker that does not have a progressive form understand on average a speaker with the progressive in their grammar? With some assumptions about communication and frequencies of real-life utterances, we can get a numerical measure of that, representing the predictions of Deo’s particular substantive theory of the progressive-imperfective cycle.

  Children can learn a different grammar than spoken by their linguistic parent by virtue of learning biases. Deo’s educated guess about the magnitude of those biases is captured in mutation rates she assumes.

  • The model is then run to check the predictions. They turn out to be not particularly appealing: the intermediate grammars of the cycle never get spoken by the whole population.

  • However, [Yanovich, 2015] shows that the problem is the infinite population assumption. Once we switch to a finite-population model, quite reasonable trajectories emerge, so we pass the proof of concept.

• **Goal 2: new predictions.**

  Interestingly, the trajectories [Yanovich, 2015] obtains using the model of [Deo, 2015] with a finite population are not exactly what we expected given the prior knowledge. But in fact we do not have enough data to decide whether our prior expectations or the new predictions based on Deo’s analysis are correct. More fine-grained data on progressive-imperfective changes is needed.
In effect, evolutionary modeling can generate novel empirical predictions, forcing us to check properties we never thought about before.

In this way, evolutionary modeling is similar to any other theory. E.g., a novel semantic analysis will often make subtle predictions about contexts that nobody looked at in detail in the past. What did not seem significant under the old understanding of the world can become highly relevant in the light of new theories.

- **Goal 3: Selection of theories.**

  Full theories in evolutionary modeling consist of two components: the substantive part and the evolutionary part. Together, the two components make a range of predictions.

  Suppose we have alternative substantive parts (e.g., different semantic analyses), or alternative evolutionary parts (i.e. different assumptions about how linguistic evolution works in a particular case). We can then implement both and check their predictions. Often, those predictions will differ. We can then see which set of predictions is better given what we know about the actual language change.

  In diachronic semantics, I am not aware of evolutionary modeling of this type.

  For phonetic change, [Baxter et al., 2009] is an example, testing the dialect formation theory by [Trudgill, 2004] that denies any effective role for social prestige or social identity (and thus is controversial). Baxter and colleagues conclude that under their evolutionary framework coupled with an implementation of Trudgill’s theory, the evolutionary parameters under which data from New Zealand English could be derived, would be completely unrealistic. They take it to falsify Trudgill’s theory.

- **Goal 4: Equilibrium selection.**

  A popular type of analysis in evolutionary game theory is based on equilibrium analysis, as opposed to full-scale evolutionary modeling.

  In population genetics, we essentially abstract away from how fitness is determined. We just take differential fitness for granted.

  In game theory, fitness (called payoffs or utility in that theory) is the very thing we focus on. Agents of different types $A$, $B$ and so on interact with each other. Different pairings result in different payoffs for the players: e.g., $A$ can get a large payoff when playing with $A$, but a small payoff when playing with $B$. Payoffs can be asymmetric. In the particular type of game called Signalling Game, the two players are Speaker and Hearer, and the payoffs represent communicative success minus production effort.
We get from individual game rounds to evolutionary game theory by assuming that very
many agents interact with each other in very many random pairings, and computing the
average payoffs which they obtain. The idea is that types with better payoffs are more
attractive, and therefore get copied more frequently. We can determine more successful
types simply by analyzing payoff matrices as above. For proper evolutionary analysis, we
have to also consider evolutionary trajectories, and use a lot more apparatus — which is
analogous to that used in population genetics. [?] is a good overview driving the point
home that equilibrium analysis needs to consider evolutionary trajectories if we don’t want
to be misled.

[Jäger, 2007] is an example of equilibrium analysis in linguistics. He defines a game for
studying argument alignment (e.g. ergative, accusative, split-ergative, and so on.)

He shows that only a few of the logically possible alignment patterns represent game-
theoretic Nash equilibria: in this context, those pairings of speaker and hearer behavior
(that is, grammars) that optimally fit each other. If speakers and hearers adjust their
behavior towards optimality, they can never converge on a grammar that is not a Nash
equilibrium. Being such an equilibrium is thus a necessary condition for having a good
probability of actually occurring and surviving for a long time.

A crucial parameter is Jäger’s system is the cost of maintaining case markers. We can
assume that the structure of the language as a whole determines that cost — at any rate,
it is independent of the alignment patterns themselves and their utility. Depending on
the cost, different alignments become Nash equilibria, thus demonstrating why different
languages may make different choices of alignment.

As the second analytical step, Jäger applies dynamic analysis, examining which of the
“nice” grammars would also arise and be retained for a long time under mutations in a
finite population. As the result, a lesser, and more empirically correct, set of alignment
grammars is found to be favored by evolution.
6. Conclusion

• There are many possible specific goals for evolutionary modeling.

• The evolutionary apparatus is just a technical tool — similarly to, e.g., lambda calculus. It is useful for helping us to derive specific consequences from our assumptions.

• It is not the thing which you absolutely must do — just as you might not need intensional semantics, or formal pragmatics, or alternative semantics, or statistical tests, depending on your project.

• However, there are issues about the dynamical behavior of linguistic systems for addressing which it is essential to use evolutionary analysis:

  – Language does not switch from stage to stage overnight. We need to address the diversity in the speaker population, and how that diversity progresses.

  – Words and constructions are usually not in one-to-one correspondence with meanings. But frequencies of different readings for a given item may change over time. We need to address not just the presence/absence of different readings, but also their changes in frequency.

    For example, we know that epistemic meanings of modals arise from circumstantial readings. If this is all there is to it, then we could imagine a situation where a modal has 98% ability readings and 2% circumstantial ◊ readings, and gives rise to epistemic readings. But what we observe in English where counts were done, is that epistemics only arise for reals when the circumstantial percentage is large. So frequency distributions of meanings do affect what categorical changes can happen.

⇒ There is an important place for evolutionary modeling in the study of semantic change.
REFERENCES


