TENSE USE IN DISCOURSE AND DIALOGUE: CROSS-LINGUISTIC SEMANTICS THROUGH PARALLEL CORPORA

Henriëtte de Swart
ZAS, Berlin, December 8, 2021
Time in Translation team at Utrecht University

Bert Le Bruyn  Martijn van der Klis  Jos Tellings  Martín Fuchs

http://time-in-translation.hum.uu.nl/
Time in Translation project (2017-2022)

- **Aim**: write a semantics of the PERFECT.
- **Constraint**: take into account cross-linguistic variation in the distribution and meaning of the HAVE PERFECT.
- **Methodology**: Translation Mining.
  - Existing parallel corpora like *EuroParl*.
- **Core languages of today’s talk**: English, Swedish, Dutch, German, Spanish and French.

We are grateful to NWO for their financial support of the project 'Time in Translation' (# 360-80-070).
Roadmap

• **Background**: why the (present) perfect, literature on the *HAVE* perfect, cross-linguistic variation, a form-based approach.


• **Semantics**: ‘classical’ perfect readings, compositional semantics (hodiernality/past time reference) and dynamic semantics (narrativity) are all ingredients of a cross-linguistically robust semantics of the perfect.

• **Pragmatics**: perfect as part of the spoken language grammar, association of the perfect with sentences types. Ratio between interrogatives/declaratives changes along the implicational hierarchy.
Background
Why the **PERFECT**?

- Extensive research on the English *Present Perfect*.
- Morphosyntax: auxiliary (*have*) + past participle (*HAVE* perfect, Dahl & Velupillai 2013).
- Core of perfect meaning: past event + state with current relevance.

1) Mary has visited Paris.                  [experiential perfect]
   (so she knows things about Paris that have current relevance)
2) Mary has moved to Paris.                [resultative perfect]
   (she currently lives in Paris)
3) Mary has been living in Paris since 2010. [continuative perfect]
   (she currently lives in Paris)
4) Malcolm X has just been assassinated.   ['hot news' perfect]
State of the art in the literature on the PERFECT


- No agreement on the semantics and pragmatics of the (present) perfect as a cross-linguistic tense-aspect category (Ritz 2012).
Cross-linguistic variation: narrative sequences

  - Le concierge s’est penché vers elle, lui *a parlé*, mais elle *a secoué* la tête, *a bredouillé* quelque chose, et *a continué* de pleurer avec la même régularité. [French]
  - El conserje se *inclinó* hacia ella, le *habló*, pero la mujer *sacudió* la cabeza, *musitó* algo y *siguió* llorando con la misma regularidad. [Spanish]
  - The caretaker *leant* over and *spoke* to her, but she *shook* her head, *mumbled* something and *went* on sobbing with the same regularity as before. [English]
- Use parallel corpora to spell out the cross-linguistic variation in the distribution of labour between *PERFECT* and (*SIMPLE/PERFECTIVE*) *PAST*. 

Utrecht University
1st TinT parallel corpus: Albert Camus (1942). *L’Étranger*

- Source language French: liberal use of the *Passé Composé*.
- Translation problem: other languages make a more restricted use of their *PERFECT*.
- *Translation Mining*: use translations to map out cross-linguistic similarities and differences in distribution across contexts.
- From language use to grammar: determine the criteria that *PERFECT* use is sensitive to in French, Italian, German, Dutch, English, Greek.

Van der Klis et al. 2020, 2021/2022
Sliding scale: implicational hierarchy

- Once a context switches from **PERFECT** to **PAST** in a particular language, the verb form remains a **PAST** tense in the next language: creates a subset relation.
- Not a dichotomy (Schaden 2009), but a sliding scale from more liberal to more restricted **PERFECT** languages (Van der Klis et al. 2020, 2021/22).
Linguistic principles governing variation
Reflections on parallel corpus research

- Data driven approach, so results might be skewed due to the specific corpus.
- Camus is said to make a ‘special’ use of the passé composé in L’Étranger — potential translation bias towards extended PERFECT use in target languages.
- Can we reproduce the insights in a translation corpus with a different source language, preferably one that has a ‘classical’ PERFECT?
- Replicate the findings from Camus in a different corpus with a different source language to validate the Translation Mining methodology.
Discourse and dialogue

• ‘Special’ literary use of the Passé Composé in Camus, but traditional story telling in Rowling.
• HP corpus can be divided into two parts:
  i. narrative discourse in which the narrator tells the story and
  ii. dialogue in which the characters talk to each other.
• Preview of results: the distribution of verb forms in discourse and dialogue is very different.
Data collection: splitting discourse and dialogue

- Convert the original and its translations into electronically readable documents.
- Select sentences containing a finite verb form from chapters 1+17 (first and last).

<table>
<thead>
<tr>
<th>dialogue</th>
<th>discourse</th>
<th>dialogue</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘You can’t blame them,’</td>
<td>said Dumbledore gently</td>
<td>‘We’ve had precious little to celebrate for eleven years’</td>
</tr>
</tbody>
</table>

- Algorithm that selects as dialogue text parts that appear between quotation marks. Remaining text is discourse.

Software created by the Digital Humanities Lab of Utrecht University. Code available via [https://time-in-translation.hum.uu.nl/](https://time-in-translation.hum.uu.nl/)
Annotation and data selection

- Align the English sentences with their translations (TimeAlign).
- Select the finite verbs in the original, match them with the translations, and specify the tenses (language specific morpho-syntactic labels): *present perfect*, *voltooid tegenwoordige tijd*, *Perfekt*, *pretérito perfecto compuesto*, etc.
- Create a scenario: choose languages (English, Dutch, German, Spanish, French), restrict the dataset to complete datapoints and finite indicative verb forms in all languages (including imperative).
- Dataset: 481 discourse contexts (scenario #533) and 320 dialogue contexts (scenario #534).
## Tense distribution in narrative discourse: verb forms per language

### English

<table>
<thead>
<tr>
<th>Label</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>simple past</td>
<td>414</td>
</tr>
<tr>
<td>past continuous</td>
<td>37</td>
</tr>
<tr>
<td>past perfect</td>
<td>24</td>
</tr>
<tr>
<td>simple present</td>
<td>4</td>
</tr>
<tr>
<td>past perfect continuous</td>
<td>1</td>
</tr>
<tr>
<td>future in the past</td>
<td>1</td>
</tr>
</tbody>
</table>

### German

<table>
<thead>
<tr>
<th>Label</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Präteritum</td>
<td>449</td>
</tr>
<tr>
<td>Präsens</td>
<td>6</td>
</tr>
<tr>
<td>Plusquamperfekt</td>
<td>26</td>
</tr>
</tbody>
</table>

### Spanish

<table>
<thead>
<tr>
<th>Label</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>pretérito indefinido</td>
<td>332</td>
</tr>
<tr>
<td>pretérito imperfecto</td>
<td>124</td>
</tr>
<tr>
<td>pretérito pluscuamperfekt</td>
<td>18</td>
</tr>
<tr>
<td>presente</td>
<td>3</td>
</tr>
<tr>
<td>estar (imperfecto) + gerundio</td>
<td>2</td>
</tr>
<tr>
<td>futuro en el pasado</td>
<td>1</td>
</tr>
<tr>
<td>ir + gerundio</td>
<td>1</td>
</tr>
</tbody>
</table>

### French

<table>
<thead>
<tr>
<th>Label</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>passé simple</td>
<td>321</td>
</tr>
<tr>
<td>imparfait</td>
<td>130</td>
</tr>
<tr>
<td>plus-que-parfait</td>
<td>24</td>
</tr>
<tr>
<td>passé récent du passé</td>
<td>3</td>
</tr>
<tr>
<td>présent</td>
<td>2</td>
</tr>
<tr>
<td>passé antérieur</td>
<td>1</td>
</tr>
</tbody>
</table>

### Dutch

<table>
<thead>
<tr>
<th>Label</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>ovt</td>
<td>453</td>
</tr>
<tr>
<td>vvt</td>
<td>23</td>
</tr>
<tr>
<td>ott</td>
<td>4</td>
</tr>
<tr>
<td>ovtt</td>
<td>1</td>
</tr>
</tbody>
</table>

### Swedish

<table>
<thead>
<tr>
<th>Label</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>imperfect</td>
<td>449</td>
</tr>
<tr>
<td>pluskvamperfekt</td>
<td>27</td>
</tr>
<tr>
<td>presens</td>
<td>5</td>
</tr>
</tbody>
</table>
Tense in discourse: charts per language

Colour coding: dark green is SIMPLE PAST or PERFECTIVE PAST, light green is IMPERFECTIVE PAST, red is PAST PERFECT, orange is PRESENT.
Tense in dialogue: charts per language

Colour coding: orange is PRESENT, blue is PRESENT PERFECT, dark green is SIMPLE PAST or PERFECTIVE PAST, light green is IMPERFECTIVE PAST, purple is FUTURE.
### Tense distribution in dialogue: verb forms per language

<table>
<thead>
<tr>
<th>Language</th>
<th>Label</th>
<th>IT</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>simple present</td>
<td>IT</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>simple past</td>
<td>IT</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>present perfect</td>
<td>IT</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>imperative</td>
<td>IT</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>simple future</td>
<td>IT</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>present continuous</td>
<td>IT</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>past continuous</td>
<td>IT</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>present perfect continuous</td>
<td>IT</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>past perfect</td>
<td>IT</td>
<td>1</td>
</tr>
<tr>
<td>German</td>
<td>Präsen</td>
<td>IT</td>
<td>148</td>
</tr>
<tr>
<td></td>
<td>Perfekt</td>
<td>IT</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>Präteritum</td>
<td>IT</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Imperativ</td>
<td>IT</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Futur I</td>
<td>IT</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Plusquamperfekt</td>
<td>IT</td>
<td>3</td>
</tr>
<tr>
<td>Spanish</td>
<td>presente</td>
<td>IT</td>
<td>126</td>
</tr>
<tr>
<td></td>
<td>pretérito indefinido</td>
<td>IT</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>pretérito imperfecto</td>
<td>IT</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>pretérito perfecto compuesto</td>
<td>IT</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>imperativo</td>
<td>IT</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>futuro imperfecto</td>
<td>IT</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>futuro próximo</td>
<td>IT</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>estar (imperfecto) + gerundio</td>
<td>IT</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>estar (presente) + gerundio</td>
<td>IT</td>
<td>1</td>
</tr>
<tr>
<td>French</td>
<td>présent</td>
<td>IT</td>
<td>130</td>
</tr>
<tr>
<td></td>
<td>passé composé</td>
<td>IT</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>imparfait</td>
<td>IT</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>impératif</td>
<td>IT</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>futur proche</td>
<td>IT</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>futur simple</td>
<td>IT</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>plus-quo-parfait</td>
<td>IT</td>
<td>2</td>
</tr>
<tr>
<td>Dutch</td>
<td>ott</td>
<td>IT</td>
<td>153</td>
</tr>
<tr>
<td></td>
<td>ovtt</td>
<td>IT</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>vtt</td>
<td>IT</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>imperativ</td>
<td>IT</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>ovt</td>
<td>IT</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>vtt</td>
<td>IT</td>
<td>1</td>
</tr>
<tr>
<td>Swedish</td>
<td>présens</td>
<td>IT</td>
<td>141</td>
</tr>
<tr>
<td></td>
<td>imperfectt</td>
<td>IT</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>perfekt</td>
<td>IT</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>imperativ</td>
<td>IT</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>futurum</td>
<td>IT</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>supinum</td>
<td>IT</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>plusquamperfekt</td>
<td>IT</td>
<td>1</td>
</tr>
</tbody>
</table>
Conclusions based on descriptive statistics

• The dominant colour in the discourse charts is green: stories are told in the PAST tense (Fleischman 1990), either SIMPLE PAST or alternation PERFECTIVE/IMPERFECTIVE PAST.

• The dominant colour in the dialogue charts is orange. It shows up in combination with blue, green and purple: dialogue is anchored in the here and now (PRESENT), looking backward to the past (PERFECT/PAST) and forward to the future (FUTURE).

• The association between tense use and formal structure is statistically significant.
Association between tense use and discourse/dialogue

• We can use association tests per language to show the statistical significance of differences in tense use between narrative discourse and dialogue.
• We focus on the five most frequently annotated TA-labels across these two registers: PRESENT, (PERFECTIVE) PAST, IMPERFECTIVE PAST, PERFECT, PAST PERFECT.
• For English, we find a significant association between register and tense use: $\chi^2(3) = 295.17, p < 0.001$. The effect size is large (Cramér’s $V = 0.65$). Idem for other languages.
• Of course, dialogues are strongly speech time oriented ~ influence of the PRESENT?
• Even if we remove all PRESENT uses from the equation, we still find significant associations between register and tense use. But we now find the effect sizes indicate only a moderate effect for English, Swedish, and Spanish.
## Association tests

Effect size increases with the implicational **PERFECT** hierarchy

<table>
<thead>
<tr>
<th>Language</th>
<th>df</th>
<th>$\chi^2$</th>
<th>$p$</th>
<th>Cramér’s V</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>3</td>
<td>295.17</td>
<td>&lt; 0.001</td>
<td>0.65</td>
</tr>
<tr>
<td>Swedish</td>
<td>3</td>
<td>347.93</td>
<td>&lt; 0.001</td>
<td>0.68</td>
</tr>
<tr>
<td>Dutch</td>
<td>3</td>
<td>465.89</td>
<td>&lt; 0.001</td>
<td>0.78</td>
</tr>
<tr>
<td>German</td>
<td>3</td>
<td>521.05</td>
<td>&lt; 0.001</td>
<td>0.83</td>
</tr>
<tr>
<td>Spanish</td>
<td>4</td>
<td>318.29</td>
<td>&lt; 0.001</td>
<td>0.65</td>
</tr>
<tr>
<td>French</td>
<td>4</td>
<td>589.06</td>
<td>&lt; 0.001</td>
<td>0.88</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Language</th>
<th>df</th>
<th>$\chi^2$</th>
<th>$p$</th>
<th>Cramér’s V</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>2</td>
<td>88.57</td>
<td>&lt; 0.001</td>
<td>0.39</td>
</tr>
<tr>
<td>Swedish</td>
<td>2</td>
<td>95.84</td>
<td>&lt; 0.001</td>
<td>0.39</td>
</tr>
<tr>
<td>Dutch</td>
<td>2</td>
<td>224.04</td>
<td>&lt; 0.001</td>
<td>0.61</td>
</tr>
<tr>
<td>German</td>
<td>2</td>
<td>314.68</td>
<td>&lt; 0.001</td>
<td>0.72</td>
</tr>
<tr>
<td>Spanish</td>
<td>3</td>
<td>85.16</td>
<td>&lt; 0.001</td>
<td>0.37</td>
</tr>
<tr>
<td>French</td>
<td>3</td>
<td>420.98</td>
<td>&lt; 0.001</td>
<td>0.82</td>
</tr>
</tbody>
</table>

Table 1: Results of the association tests between register and tense use. All languages show a significant association between register and tense use.

Table 2: Results of the association tests between register and tense use when removing all *present* uses. All languages show a significant association between register and tense use.
The **PERFECT** belongs to the spoken language grammar

- No **PERFECT** use in the discourse part of HP, exclusively in dialogue.
- One author, one translator per language, so this is *intraspeaker* variation.
- The association tests show a significant different, effect size goes up with the implicational hierarchy.
- **Conclusion**: even though the **PERFECT** may appear in narrative sequences in the spoken language (French, Italian, Catalan – see Camus corpus), it has not replaced the **PERFECTIVE PAST** in the written language.
- **Hypothesis**: the **PERFECT** is a tense-aspect category that belongs to the spoken language grammar.
Validation of the TinT methodology in the discourse part of *Harry Potter*
From data on individual languages to multilingual comparison: n-tuples

- So far: verb forms per language.
- Multilingual comparison requires investigation of distribution of tense forms over individual datapoints (contexts).
- Multilingual comparison based on n-tuples: combination of language specific verb forms in each context.
- In the scenario, we chose 6 languages, so we have 6-tuples of verb forms.
- For example:
  
  <simple past, imperfekt, ovt, Präteritum, pretérito indefinido, passé simple>
  <past continuous, imperfekt, ovt, Präteritum, pretérito imperfect, imparfait>
### 6-tuples: cross-linguistic distribution over contexts

<table>
<thead>
<tr>
<th>English</th>
<th>German</th>
<th>Spanish</th>
<th>French</th>
<th>Dutch</th>
<th>Swedish</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>simple past</td>
<td>Präteritum</td>
<td>pretérito indefinido</td>
<td>passé simple</td>
<td>ovt</td>
<td>imperfekt</td>
<td>312</td>
</tr>
<tr>
<td>simple past</td>
<td>Präteritum</td>
<td>pretérito imperfecto</td>
<td>imparfait</td>
<td>ovt</td>
<td>imperfekt</td>
<td>79</td>
</tr>
<tr>
<td>past continuous</td>
<td>Präteritum</td>
<td>pretérito imperfecto</td>
<td>imparfait</td>
<td>ovt</td>
<td>imperfekt</td>
<td>26</td>
</tr>
<tr>
<td>simple past</td>
<td>Präteritum</td>
<td>pretérito indefinido</td>
<td>imparfait</td>
<td>ovt</td>
<td>imperfekt</td>
<td>11</td>
</tr>
<tr>
<td>past perfect</td>
<td>Plusquamperfekt</td>
<td>pretérito pluscuamperfecto</td>
<td>plus-que-parfait</td>
<td>vvt</td>
<td>pluskvamperfekt</td>
<td>9</td>
</tr>
<tr>
<td>simple past</td>
<td>Präteritum</td>
<td>pretérito imperfecto</td>
<td>passé simple</td>
<td>ovt</td>
<td>imperfekt</td>
<td>4</td>
</tr>
<tr>
<td>past perfect</td>
<td>Plusquamperfekt</td>
<td>pretérito pluscuamperfecto</td>
<td>passé récent du passé</td>
<td>vvt</td>
<td>pluskvamperfekt</td>
<td>3</td>
</tr>
<tr>
<td>past perfect</td>
<td>Plusquamperfekt</td>
<td>pretérito imperfecto</td>
<td>plus-que-parfait</td>
<td>vvt</td>
<td>pluskvamperfekt</td>
<td>3</td>
</tr>
<tr>
<td>past perfect</td>
<td>Präteritum</td>
<td>pretérito indefinido</td>
<td>plus-que-parfait</td>
<td>vvt</td>
<td>pluskvamperfekt</td>
<td>2</td>
</tr>
<tr>
<td>simple past</td>
<td>Präteritum</td>
<td>pretérito indefinido</td>
<td>passé simple</td>
<td>ovt</td>
<td>imperfekt</td>
<td>2</td>
</tr>
<tr>
<td>past continuous</td>
<td>Präteritum</td>
<td>pretérito imperfecto</td>
<td>plus-que-parfait</td>
<td>vvt</td>
<td>pluskvamperfekt</td>
<td>2</td>
</tr>
<tr>
<td>simple present</td>
<td>Präsens</td>
<td>presente</td>
<td>imparfait</td>
<td>ovt</td>
<td>presens</td>
<td>2</td>
</tr>
<tr>
<td>past continuous</td>
<td>Präteritum</td>
<td>futuro en el pasado</td>
<td>imparfait</td>
<td>ovt</td>
<td>imperfekt</td>
<td>1</td>
</tr>
<tr>
<td>past continuous</td>
<td>Präteritum</td>
<td>estar (imperfecto) + gerundio</td>
<td>imparfait</td>
<td>ovt</td>
<td>imperfekt</td>
<td>1</td>
</tr>
</tbody>
</table>
Multidimensional Scaling

- From data on individual languages to multilingual comparison through *Multidimensional Scaling* (Wälchli & Cysouw 2012).
- N-tuples: combination of language specific verb forms per context.
- N-tuples constitute the input to *Multidimensional Scaling*: dissimilarity matrix.
- MDS is a statistical algorithm that searches for dimensions of variation. For all languages, MDS tries to place the same (language-specific) verb forms close to each other on the map.
- Output: scatter map in two-dimensional space.
- Colour coding per language shows the cross-linguistic distributional patterns.
MDS visualisation of the cross-linguistic distribution of tense forms over contexts: English map
MDS visualisation of the cross-linguistic distribution of tense forms over contexts: English map
German map
Spanish map
French map
2 dimensions of variation determine tense distribution in discourse

- **Progressive/imperfective/perfective aspect**
- **Temporal reference**

- Simple past
- Past continuous
- Past perfect
- Simple present
- Past perfect continuous
- Future in the past
English map
Spanish map
Interactive interface: going back and forth between maps and underlying data

- Point the mouse at a dot, to see the tuple + number of fragments + example.

Click on the dot to see the fragments.

Select the fragment to see the underlying datapoint.

79 fragments, for example:
56720: As usual when he was indoors, Hagrid looked too big to be allowed simple past, ovt, Präteritum, imperfect, imperfecto, imperfecto

79 fragments with labels: simple past, ovt, Präteritum, imperfect, imperfecto, pretérito imperfecto

<table>
<thead>
<tr>
<th>Document</th>
<th>Sentence XML-id</th>
<th>Target words</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.xml</td>
<td>s3.1</td>
<td>was</td>
</tr>
<tr>
<td>1.xml</td>
<td>s3.2</td>
<td>was</td>
</tr>
<tr>
<td>1.xml</td>
<td>s3.3</td>
<td>had</td>
</tr>
<tr>
<td>1.xml</td>
<td>s3.3</td>
<td>was</td>
</tr>
<tr>
<td>1.xml</td>
<td>s3.4</td>
<td>had</td>
</tr>
<tr>
<td>1.xml</td>
<td>s4.1</td>
<td>had</td>
</tr>
<tr>
<td>1.xml</td>
<td>s4.3</td>
<td>were</td>
</tr>
<tr>
<td>1.xml</td>
<td>s4.3</td>
<td>pretended</td>
</tr>
<tr>
<td>1.xml</td>
<td>s4.3</td>
<td>was</td>
</tr>
</tbody>
</table>
Discourse: SIMPLE PRESENT, SIMPLE PAST/PERFECTIVE PAST with events, SIMPLE PAST/IMPERFECTIVE PAST with states; freedom of translation

a. When Mr and Mrs Dursley woke up on the dull, grey Tuesday our story starts, there was nothing about the cloudy sky outside to suggest that strange and mysterious things would soon be happening all over the country. [English]

b. När mr och mrs Dursley vaknade den dystra, gråa tisdag då vår berättelse börjar, fanns det ingenting hos den molniga himlen utanför som antydde att konstiga och mystiska saker snart skulle hånda runt omkring i landet. [Swedish]

c. Toen meeneer en mevrouw Duffeling wakker werden, op de duffe grijze dinsdag waarop ons verhaal begint, zou je bij het zien van de bewolkte hemel absoluut niet gezegd hebben dat die dag overal in het land vreemde en geheimzinnige dingen zouden gaan gebeuren. [Dutch]

d. Als Mr und Mrs Dursley an dem trüben und grauen Dienstag, an dem unsere Geschichte beginnt, die Augen aufschlugen, waren an dem wolkenverhangenen Himmel draußen kein Vorzeichen der merkwürdigen und geheimnisvollen Dinge zu erkennen, die bald überall im Land geschehen sollten. [German]

e. Nuestra historia comienza cuando el señor y la señora Dursley se despertaron un martes, con un cielo cubierto de nubes grises que amenazaban tormenta. Pero nada había en aquel nublado cielo que sugeriera los acontecimientos extraños y misteriosos que poco después tendrían lugar en toda la región. [Spanish]

f. Lorsque Mr et Mrs Dursley s'éveillèrent, au matin du mardi où commence cette histoire, il faisait gris et triste et rien dans le ciel nuageux ne prévoyait que des choses étranges et mystérieuses allaient bientôt se produire dans tout le pays. [French]
Narrative sequence: alternation of PERFECTIVE PAST (with events), PAST CONTINUOUS/IMPERFECTIVE PAST (with ongoing activities); incomplete tuples.

a. At half past eight, Mr Dursley picked up his briefcase, pecked Mrs Dursley on the cheek and tried to kiss Dudley goodbye but missed, because Dudley was now having a tantrum and throwing his cereal at the walls. [English]

b. Klockan halv nio tog mr Dursley upp sin portfölj, gav mrs Dursley en hastig kyss på kinden och försökte pussa Dudley adjö men missade, för Dudley hade just ett raseriutbrott och slängde omkring flingorna i köket. [Swedish]

c. Om halft negen pakte meneer Duffeling zijn koffertje, gaf mevrouw Duffeling een kus op haar wang en probeerde Dirk ook een zoentje te geven, maar dat mislikte omdat Dirk net een woedeaanval had en zijn cornflakes tegen de muur smeet. [Dutch]

d. Um halb neun griff Mr Dursley nach der Aktentasche, gab seiner Frau einen Schmatz auf die Wange und versuchte es auch bei Dudley mit einem Abschiedskuss. Der ging jedoch daneben, weil Dudley gerade einen Wutanfall hatte und die Wände mit seinem Haferbrei bewarb. [German]

e. A las ocho y media, el señor Dursley cogió su maletín, besó a la señora Dursley en la mejilla y trató de despedirse de Dudley con un beso, aunque no pudo, ya que el niño tenía un berrinche y estaba arrojando los cereales contra las paredes. [Spanish]

f. A huit heures et demie, Mr Dursley prit son attaché-case, déposa un baiser sur la joue de Mrs Dursley et essaya d'embrasser Dudley, mais sans succès, car celui-ci était en proie à une petite crise de colère et s'appliquait à jeter contre les murs de la pièce le contenu de son assiette de céréales. [French]
Summary of findings: narrative discourse

• **Perfective/Imperfective/Simple Past** tenses dominate in narrative discourse, as expected (Fleischman 1990).
• The distribution of verb forms in narrative discourse confirms claims made about tense use in narration in the DRT/SDRT literature (Kamp & Rohrer (1983), Partee (1984), Hinrichs (1986), Lascarides & Asher (1993), etc.
• No (substantial) cross-linguistic variation is found: this is as expected.
• **Conclusion**: cross-linguistic stability in narrative discourse validates the *Translation Mining* methodology.
• **No** Perfect use in the discourse component of HP in any language. Focus on dialogue.
Replication of the implicational hierarchy in the dialogue part of *Harry Potter*
## Tense use in dialogue: descriptive statistics

<table>
<thead>
<tr>
<th>Tense Type</th>
<th>English</th>
<th>German</th>
<th>Spanish</th>
<th>French</th>
<th>Dutch</th>
<th>Swedish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>117</td>
<td>148</td>
<td>126</td>
<td>130</td>
<td>153</td>
<td>141</td>
</tr>
<tr>
<td>Past</td>
<td>114</td>
<td>54</td>
<td>85</td>
<td>0</td>
<td>78</td>
<td>114</td>
</tr>
<tr>
<td>Present Perfect</td>
<td>26</td>
<td>78</td>
<td>24</td>
<td>99</td>
<td>58</td>
<td>26</td>
</tr>
<tr>
<td>Imperative</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>22</td>
<td>22</td>
<td>20</td>
</tr>
<tr>
<td>Imperfect</td>
<td>0</td>
<td>0</td>
<td>39</td>
<td>48</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Future</td>
<td>19</td>
<td>14</td>
<td>13</td>
<td>8</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>Recent Future</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Present Progressive</td>
<td>13</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Past Perfect</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Past Progressive</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Past Participle</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Present Perfect Progressive</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Multidimensional scaling: English map
Multidimensional scaling: two meaning dimensions

- future/non-future time reference
- non-past/past time reference

- imperative
- simple future
- simple present
- present perfect
- simple past
- present continuous
- present perfect
Unpacking dimensions of variation

• Preview of results: in all languages, the PRESENT PERFECT lives in the right part of the map (non-future temporal reference).
• ‘Classical’ uses live in the upper right corner (non-future, non-past temporal reference).
• More liberal languages extend the PERFECT domain towards the lower right corner of the map (past temporal reference of events).
Multidimensional scaling: English map
Perfect: English map (26 present perfect)
Perfect: Swedish map (26 perfekt)
Perfect: Spanish map (24 pretérito perfecto compuesto)
Perfect: Dutch map (58 voltooid tegenwoordige tijd)
+27 blue dots in the lower right corner of the map
Perfect: German map (78 Perfekt)
27+15 blue dots in the lower right corner of the map
Perfect: French map (99 passé compose)
27 + 15 + 12 blue dots in the lower right corner of the map
Intermediate conclusions

• The HP maps confirm the subset relation from the Camus dataset: English/Swedish/Spanish $\subseteq$ Dutch $\subseteq$ German $\subseteq$ French.
• There is a bit more translation noise in the HP dataset than in the Camus dataset.
• Implicational hierarchy: more liberal PERFECT languages maintain a presence in the upper right corner, but extend the blue domain towards the lower right corner of the map.
• They become more PAST oriented.
• We still need to show that they acquire a PERFECTIVE PAST flavour (reporting an event in the past).
Present, perfect, past: English map
Present, perfect, past: Swedish map
Present, perfect, perfective/imperfective past: Spanish map
Present, perfect, past: Dutch map
+27 blue dots in the lower right corner of the map
Present, perfect, past: German map
27+15 blue dots in the lower right corner of the map
Present, perfect, imperfective past: French map
27 + 15 + 6 blue dots in the lower right corner of the map
Intermediate conclusions: cross-linguistic distribution

- Implicational hierarchy: competition between PERFECT and (PERFECTIVE) PAST
- Are the linguistic features underlying the distribution the same as in the Camus dataset? Do they help to distinguish Spanish, Swedish and English?
- Based on Camus, we expect that resultative/existential meanings are shared by all languages (‘core’ perfect readings).
- Based on Camus we predict the following differences:
  - French/German ~ stative verbs
  - German/Dutch ~ narrativity
  - Dutch/Spanish ~ past time reference
  - Spanish/English ~ hodiernality
Semantic ingredients of the PERFECT
‘Core’ perfect readings: resultative

7 contexts in which all languages use a PERFECT, 3 clear resultative readings.

a. ‘See what I have become?’ the face said. [English]
b. "Ser du vad det har blivit av mig?" sade ansiktet. [Swedish]
c. "Zie je wat ervan me geworden is?" zei het gezicht. [Dutch]
d. "Siehst du, was aus mir geworden ist?", sagte das Gesicht. [German]
e. - ¿Ves en lo que me he convertido? - dijo la cara -. [Spanish]
f. “Tu vois ce que je suis devenu?” dit le visage. [French]
‘Core’ perfect reading: universal

7 contexts in which all languages use a perfect, 3 clear instances of
an universal/negative/negative existential reading.

a. My dear Professor, I’ve never seen a cat sit so stiffly.’
   [English]

b. "Kära professor, jag har aldrig sett en katt sitta så stelt.'
   [Swedish]

c. ‘M'n beste professor, ik heb nog nooit een kat zo stijfjes zien zitten.’
   [Dutch]

d. "Mein lieber Professor, ich habe noch nie eine Katze so steif dasitzen sehen.'
   [German]

e. - Mi querida profesora, nunca he visto a un gato tan tieso.
   [Spanish]

f. “Mon cher professeur, je n’ai jamais vu un chat se tenir d’une manière aussi raide.”
   [French]
‘Core’ perfect reading: existential

- Overall cross-linguistic stability in the part of the blue domain in the upper right corner of the map, yet minor variations across languages: translation choices.
- Out of 26 cases of the English *present perfect*, there are 11 contexts in which 4 out of 5 translations maintain the PERFECT from the original.
- E.g. state that lasted for a while and is now over: different construction in French.

a. ‘You’ve had nearly fifteen minutes, now OUT’, she said. [English]
b. Ni *har haft* nästan femton minute, nu ska ni ut", sade hon bestämt. [Swedish]
c. “Jullie *hebben* bijna een kwartier de tijd *gehadt*, wegwezen!” zei ze gedecideerd. [Dutch]
d. “Ihr *habt* fast fünfzehn Minuten *gehhabt*, nun aber RAUS” sagte sie bestimmt. [German]
e. - Ya *habéis estado* quince minutos, ahora FUERA - dijo con severidad. [Spanish]
f. “Ça *fait* presqu’un quart d’heure, maintenant.” [French]
‘Core’ perfect reading from the literature on English, but not necessarily for other languages: continuative

Only 1 context in the HP dataset with a clear continuative interpretation, but familiar observation from the literature.

- **a.** ‘How long *have I been* in here?’ [English]
- **b.** "Hur länge *har* jag *varit* här inne?" [Swedish]
- **c.** "Hoe lang *lig* ik hier al?" [Dutch]
- **d.** "Wie lange *bin* ich schon hier?" [German]
- **e.** - ¿Cuánto tiempo hace que *estoy* aquí? [Spanish]
- **f.** “Ça fait combien de temps que je *suis* là?” [French]
Cross-linguistic differences: English/Swedish

• The English present perfect is never translated by means of an Imperfekt in Swedish, but presens in 7 cases (resultative/result state alternation). In 2 out of 3 instances of the present perfect continuous the Swedish translator uses a perfekt; the remaining context displays a supinum (4 supinum cases in dataset).
• Universal readings do not necessarily require the present perfect in English, or in other languages, but the Swedish translator sometimes uses a perfekt.
• Overall cross-linguistic variation makes this datapoint hard to interpret.

a. But he [never] wanted you dead. [English]
b. Men han [har aldrig velat att du ska dö]. [Swedish]
c. Maar hij [heeft nooit je dood gewild]. [Dutch]
d. Aber er [wollte nie, dass Sie sterben]. [German]
e. Pero nunca [quiso que estuvieras muerto]. [Spanish]
f. Mais il n’a jamais [voulu vous tuer pour autant]. [French]
Cross-linguistic differences Spanish/English: hodiernality

• In Spanish, the *pretérito perfecto compuesto* is used to refer to an event in the hodiernal past, even if the original is in the *simple past*, which the Swedish translation maintains.
• Only a handful of examples illustrating hodiernality in the HP dataset, so weak contrast, but we found more evidence in the Camus dataset.

a. What *did* you *just* see?’ [English]
b. Vad var det du *såg* nyss?” [Swedish]
c. Wat *heb* je *gezien?” [Dutch]
d. Was *hast* du *gesehen?” [German]
e. ¿Qué es lo que *has visto?” [Spanish]
f. Qu’est-ce que vous *avez vu?” [French]
Cross-linguistic differences Dutch/Spanish: past time reference

27 contexts in which Dutch uses the vtt, but Spanish switches to the pretérito indefinido: pre-hodiernal past time reference.

a. Dumbledore *gave* me the day off *yesterday* ter fix it. [English]
b. Dumbledore *gav* mej ledigt hela dan i går för å fixa de. [Swedish]
c. Perkamentus *heb* me gisteren vrijaf *gegeven* om 't te regelen. [Dutch]
d. Dumbledore *hat* mir gestern dafür *freigegeben*. [German]
e. Dumbledore me *dio* libre el día de ayer para hacerlo. [Spanish]
f. Dumbledore m’a *accordé* un jour de congé hier pour le préparer. [French]
Cross-linguistic differences German/Dutch: narrativity

HP dialogue mostly has short turns, but some longer turns report sequences of events. 15 contexts in which German uses the Perfekt, but Dutch switches to the onvoltooid verleden tijd (Le Bruyn et al. 2019).

a. Your friend Miss Granger accidentally knocked me over as she rushed to set fire to Snape at that Quidditch match. She broke my eye contact with you. [English]

b. Din väninna miss Granger slog omkull mig av en händelse då hon rusade i väg för att sätta eld på Snape vid den där quidditchmatchen. Hon bröt min ögonkontakt med dig. [Swedish]

c. Juffrouw Griffel lied me per ongeluk omver toen ze Snape in brand wilde steken tijdens die Zwerkbalwedstrijd en daardoor verbrak ze mijn oogcontact met jou. [Dutch]

d. Ihre Freundin Miss Granger hat mich versehentlich umgerempelt, als sie beim Quidditch-Spiel zu Snape hinüberrannte, um ihn anzuzünden. Sie hat meinen Blickkontakt zu Ihnen unterbrochen. [German]

e. Tu amiga, la señorita Granger, accidentalmente me atropelló cuando corría a prenderle fuego a Snape, en ese partido de Quidditch. Y rompió el contacto visual que yo tenía contigo. [Spanish]

f. Votre amie, Miss Granger m’a bousculé par accident quand elle s’est précipitée pour mettre le feu aux vêtements de Rogue, pendant le match de Quidditch. A cause d’elle, j’ai perdu le contact visuel avec vous. [French]
Cross-linguistic differences French/German: stative verbs

• Much of the variation between French and German relates to translation noise: the choice the French translator makes to present the English *simple past* as a state or an event, and the German translator choosing a *Präteritum* in narrative sequences.

• In 6 contexts, the difference clearly relates to aspectual class: French can report a stative verb in the *passé composé*, but with stative verbs, German switches to the *Präteritum*. Familiar observation from the Camus dataset.

a. Ah - your father happened to leave it in my possession and I thought you might like it.’
   [English]

b. A, din pappa råkade lämna den i mitt förvar och jag tänkte att du kanske skulle tycka om den.
   [Swedish]

c. Die had je vader toevallig bij mij in bewaring gegeven en ik dacht dat jij hem misschien zou willen hebben. [Dutch]

d. "Ah, es traf sich, dass ihn dein Vater mir anvertraut hat, und ich dachte, dir gefiele er vielleicht." [German]

e. Resulta que tu padre me la habia dejado y pensé que te gustaría tenerla.
   [Spanish]

f. Il se trouve que ton père l’avait laissée en ma possession et j’ai pensé que tu aimerais peut-être l’avoir. [French]
Conclusions about the PERFECT scale in dialogue

• **Subset relation** English/Swedish ⋒ Spanish ⋒ Dutch ⋒ German ⋒ French. Established for the Camus corpus, and replicated in the dialogue part of HP, but more translation noise.

• **Implicational hierarchy**: extended PERFECT semantics, competition between PERFECT and (PERFECTIVE) PAST.

• **Semantic features** responsible for variation are the same across the two datasets: lexical aspect (stative verbs), compositionality (hodiernal/past time reference), dynamic semantics (narrativity).
Remaining issue: discourse vs. dialogue

• Why is the \textit{PERFECT} a tense-aspect category of the spoken language grammar in all languages under investigation?
• Investigate the pragmatics of the \textit{PERFECT}. 
Pragmatic ingredients of the PERFECT
The **PERFECT** belongs to the spoken language grammar

- Why is the **PERFECT** a tense-aspect category of the spoken language grammar in all languages under investigation?
- Portner (2003): the **PERFECT** is used to answer the question under discussion.
- Nishiyama & Koenig (2010): **PRESENT PERFECT** may be used for topic negotiation.
- Both uses are attested in the HP dataset; cross-linguistic stability.
Answering the question under discussion

Question under discussion provided in context; qud relates to the utterance situation (‘here and now’).

Professor McGonagall asking Dumbledore:
And I don't suppose you're going to tell me why you’re here, of all places?’

a. I've come to bring Harry to his aunt and uncle.  
   [English]
b. Jag har kommit för att ta med mig Harry till hans moister och morbor.  
   [Swedish]
c. Ik kom Harry aflveren bij zijn oom en tante.  
   [Dutch]
d. Ich bin gekommen, um Harry zu seiner Tante und seinem Onkel zu bringen.  
   [German]
e. He venido a entregar a Harry a su tía y su tío.  
   [Spanish]
f. Je suis venu confier Harry à sa tante et à son oncle.  
   [French]
Topic negotiation

Discourse context: Mrs Dursley came into the living-room carrying two cups of tea. It was no good. He’d have to say something to her. He cleared his throat nervously.

a. Er - Petunia, dear - you haven’t heard from your sister lately, have you?  
   [English]
b. Hrrm ... Petunia, älskling ... du har väl inte hört ifrån din syster på sistone?  
   [Swedish]
c. Eh - Petunia, schat - je hebt de laatste tijd toch toevallig niet iets van je zus gehoord?  
   [Dutch]
d. Ahm - Petunia, Liebes - du hast in letzter Zeit nichts von deiner Schwester gehört, oder?  
   [German]
e. Petunia, querida, ¿has sabido últimamente algo sobre tu hermana?  
   [Spanish]
f. “Pétunia, ma chérie”, dit-il, “tu n'as pas eu de nouvelles de ta sœur récemment?”  
   [French]
Association with sentence type

• Answering a question under discussion ~ declarative.
• Negotiating a new topic ~ interrogative.
• Both functions are attested in the HP dataset, so how are PERFECTS distributed over sentence types?
• Annotation for sentence type within the dialogue part of HP: interrogative, declarative, exclamatory, imperative.
Distribution of **PERFECTS** over sentence types (English)

<table>
<thead>
<tr>
<th></th>
<th># verb forms</th>
<th># PERFECTS</th>
<th>% PERFECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>interrogative</td>
<td>54</td>
<td>7</td>
<td>13%</td>
</tr>
<tr>
<td>declarative</td>
<td>228</td>
<td>17</td>
<td>7.5%</td>
</tr>
<tr>
<td>exclamatory</td>
<td>23</td>
<td>2</td>
<td>8.7%</td>
</tr>
</tbody>
</table>

- Higher ratio of **PERFECT** use in interrogatives than in declaratives. The odds of **PERFECT** use in questions in the English original are 1.73 times higher as opposed to statements.
- Low numbers, but in line with the analysis of the Switchboard corpus, for which we found that **PERFECT** use in questions is 1.5 times higher as opposed to statements (Tellings et al. 2019).
- No **PERFECT** use in imperatives, we leave exclamatory sentences aside.
Distribution of **PERFECT** over sentence type across languages

More important than the absolute numbers is the change in relative frequency: the path towards extended **PERFECT** use leads to higher percentages of **PERFECT** use in interrogatives as well as declaratives in Dutch. In German and French, the increase is larger in declaratives, changing the odds.

<table>
<thead>
<tr>
<th>PERFECT use</th>
<th>interrogatives</th>
<th>declaratives</th>
<th>odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>7</td>
<td>17</td>
<td>1.74</td>
</tr>
<tr>
<td>Swedish</td>
<td>7</td>
<td>17</td>
<td>1.74</td>
</tr>
<tr>
<td>Spanish</td>
<td>6</td>
<td>17</td>
<td>1.47</td>
</tr>
<tr>
<td>Dutch</td>
<td>15</td>
<td>38</td>
<td>1.66</td>
</tr>
<tr>
<td>German</td>
<td>16</td>
<td>57</td>
<td>1.18</td>
</tr>
<tr>
<td>French</td>
<td>17</td>
<td>76</td>
<td>0.95</td>
</tr>
</tbody>
</table>
More important than the absolute numbers is the change in relative frequency: the path towards extended PERFECT use leads to higher percentages of PERFECT use in interrogatives as well as declaratives in Dutch. In German and French, the increase is larger in declaratives, changing the odds.

<table>
<thead>
<tr>
<th>PERFECT use</th>
<th>interrogatives</th>
<th>declaratives</th>
<th>odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>7</td>
<td>17</td>
<td>1.74</td>
</tr>
<tr>
<td>Swedish</td>
<td>7</td>
<td>17</td>
<td>1.74</td>
</tr>
<tr>
<td>Spanish</td>
<td>6</td>
<td>17</td>
<td>1.47</td>
</tr>
<tr>
<td>Dutch</td>
<td>15</td>
<td>38</td>
<td>1.66</td>
</tr>
<tr>
<td>German</td>
<td>16</td>
<td>57</td>
<td>1.18</td>
</tr>
<tr>
<td>French</td>
<td>17</td>
<td>76</td>
<td>0.95</td>
</tr>
</tbody>
</table>
Distribution of PERFECT over sentence type across languages

More important than the absolute numbers is the change in relative frequency: the path towards extended PERFECT use leads to higher percentages of PERFECT use in interrogatives as well as declaratives in Dutch. In German and French, the increase is larger in declaratives, changing the odds.

<table>
<thead>
<tr>
<th>PERFECT use</th>
<th>interrogatives</th>
<th>declaratives</th>
<th>odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>7</td>
<td>17</td>
<td>1.74</td>
</tr>
<tr>
<td>Swedish</td>
<td>7</td>
<td>17</td>
<td>1.74</td>
</tr>
<tr>
<td>Spanish</td>
<td>6</td>
<td>17</td>
<td>1.47</td>
</tr>
<tr>
<td>Dutch</td>
<td>15</td>
<td>38</td>
<td>1.66</td>
</tr>
<tr>
<td>German</td>
<td>16</td>
<td>57</td>
<td>1.18</td>
</tr>
<tr>
<td>French</td>
<td>17</td>
<td>76</td>
<td>0.95</td>
</tr>
</tbody>
</table>
Distribution of PERFECT over sentence type across languages

More important than the absolute numbers is the change in relative frequency: the path towards extended PERFECT use leads to higher percentages of PERFECT use in interrogatives as well as declaratives in Dutch. In German and French, the increase is larger in declaratives, changing the odds.

<table>
<thead>
<tr>
<th>PERFECT use</th>
<th>interrogatives</th>
<th>declaratives</th>
<th>odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>7</td>
<td>17</td>
<td>1.74</td>
</tr>
<tr>
<td>Swedish</td>
<td>7</td>
<td>17</td>
<td>1.74</td>
</tr>
<tr>
<td>Spanish</td>
<td>6</td>
<td>17</td>
<td>1.47</td>
</tr>
<tr>
<td>Dutch</td>
<td>15</td>
<td>38</td>
<td>1.66</td>
</tr>
<tr>
<td>German</td>
<td>16</td>
<td>57</td>
<td>1.18</td>
</tr>
<tr>
<td>French</td>
<td>17</td>
<td>76</td>
<td>0.95</td>
</tr>
</tbody>
</table>
Distribution of PERFECT over sentence type across languages

More important than the absolute numbers is the change in relative frequency: the path towards extended PERFECT use leads to higher percentages of PERFECT use in interrogatives as well as declaratives in Dutch. In German and French, the increase is larger in declaratives, changing the odds.

<table>
<thead>
<tr>
<th>PERFECT use</th>
<th>interrogatives</th>
<th>declaratives</th>
<th>odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>7</td>
<td>17</td>
<td>1.74</td>
</tr>
<tr>
<td>Swedish</td>
<td>7</td>
<td>17</td>
<td>1.74</td>
</tr>
<tr>
<td>Spanish</td>
<td>6</td>
<td>17</td>
<td>1.47</td>
</tr>
<tr>
<td>Dutch</td>
<td>15</td>
<td>38</td>
<td>1.66</td>
</tr>
<tr>
<td>German</td>
<td>16</td>
<td>57</td>
<td>1.18</td>
</tr>
<tr>
<td>French</td>
<td>17</td>
<td>76</td>
<td>0.95</td>
</tr>
</tbody>
</table>
Increased PERFECT use in interrogatives from Dutch onwards

Preceding discourse: Harry is in the hospital after his fight with Squirrel/ Voldemort, and tells his friends about his adventures, after they got separated.

‘That's what I said, but Dumbledore thinks that - what was it ? - “ to the well-organised mind, death is but the next great adventure”. ‘I always said he was off his rocker,’ said Ron, looking quite impressed at how mad his hero was.

‘So what happened to you two?’ said Harry.

‘ Well, I got back all right,’ said Hermione.

a. ‘So what happened to you two?’ [English]
b. "Så vad hände med er båda?" [Swedish]
c. "En hoe is het jullie vergaan?" [Dutch]
d. "Und was ist mit euch geschehen?" [German]
e. - ¿Yqué os pasó a vosotros dos? – [Spanish]
f. “Et vous, qu’est-ce que vous est arrivé?” [French]

Discourse topic: adventures in the tower, past time reference, topic continuity.
**Increased PERFECT use in interrogatives**

- Nishiyama & Koenig (2010): PERFECT can be used for topic negotiation (association with interrogatives). Existential PERFECT, current relevance.
- Other types of questions:
  - information seeking questions related to an ongoing discourse topic with past time reference,
  - presupposition bearing questions (e.g. involving a cleft construction: further details on definite past event).
- Pragmatic variation across languages: English *present perfect* is infelicitous with presupposed events (Michaelis 1994).
Pragmatic variation across languages

Michaelis (1994): pragmatically presupposed events cannot be reported in the *Present Perfect*, but require the *Simple Past*. Also in Swedish, not in the other languages in the dataset.

a. Il a bégayé un peu: « On l'*a couverte*, mais je dois dévisser la bière pour que vous puissiez la voir.» [Fr]
b. He stuttered a bit. 'We *covered* her up. But I was to unscrew the coffin to let you see her.' [English]
c. Han stammade lätt: "De *lade* på locket, men jag ska ta ur skruvarna så kan ni se henne." [Swedish]
d. Hij stotterde een beetje: 'Ze *hebben* het deksel er op *gedaan*, maar ik moet het losschroeven om u in de gelegenheid te stellen uw moeder te zien.' [Dutch]
e. Er hat ein bißchen herumgestottert: «Man *hat* sie *zugemacht*, aber ich muß den Sarg nur aufschrauben, damit Sie sie sehen können.» [German]
f. Tartamudeó un poco: « La *hemos cubierto*. Pero desatornillaré el féretro para que pueda usted verla ». [Spa]
Presupposition bearing questions: cleft construction

The cleft-construction introduces the past event of telling as presupposed. The question targets the identity of the agent of the telling event.

a. It must have made sense to Dumbledore, though, because he put it back in his pocket and said, ‘Hagrid's late. I suppose it was he who told you I'd be here, by the way?’ ‘Yes,’ said Professor McGonagall.  
   [English]

b. Apropå det var det väl han som talade om för er att jag skulle vara här?”  
   [Swedish]

c. Waarschijnlijk heeft hij gezegd dat ik hier zou zijn?’  
   [Dutch]

d. Übrigens nehme ich an, er hat Ihnen erzählt, dass ich hierherkommen würde?”  
   [German]

e. Imagino que fue él quien le dijo que yo estaría aquí, ¿no?  
   [Spanish]

f. Au fait, j’imagine que c'est lui qui vous a dit que je serais ici?”  
   [French]
Main clause question: the presence of the troll is presupposed, past time reference (preceding discourse). Harry checks the identity of the agent of the letting in event.

Note the cleft constructions in the Swedish, Spanish and French translations.

Embedded question: rhetorical question (no answer expected, none given) related to the same event.

Note the switch to an assertive sentence in the German translation.

“You’re too nosy to live, Potter. Scurrying around the school at Halloween like that, for all I knew you’d seen me coming to look at what was guarding the Stone.”

“You let the troll in?”

“Certainly. I have a special gift with trolls - you must have seen what I did to the one in the chamber back there?” Unfortunately, while everyone else was running around looking for it, Snape, who already suspected me, went straight to the third floor to head me off - and not only did my troll fail to beat you to death, that three-headed dog didn’t even manage to bite Snape’s leg off properly.”

a. “You let the troll in?”
   Certainly. I have a special gift with trolls - you must have seen what I did to the one in the chamber back there?”

   [English]

b. "Var det ni som släppte in trollet?"
   "Javisst. Jag har en speciell förmåga att handkas med troll - du måste ha sett vad jag gjorde med trollen i rummet där borta?"

   [Swedish]

c. "Hebt u die trol binnengelaten?"
   "Jazeker. Ik heb een speciale gave wat trollen betreft - je hebt zeker wel gezien wat ik met dat exemplaar hiernaast heb gedaan?"

   [Dutch]

d. "Sie haben den Troll hereingelassen?"
   Gewiss. Ich habe ein glückliches Händchen, wenn es um Trolle geht. Sie haben ja gesehen, was ich mit dem in der Kammer dort hinten angestellt habe.

   [German]

e. - ¿Usted fue el que dejó entrar al troll? –
   - Claro. Yo tengo un don especial con esos monstruos. ¿No viste lo que le hizo al que estaba en la otra habitación?

   [Spanish]

f. “C’est vous qui avez fait entrer le troll?”
   “Bien sûr. Vous avez dû constater ce que j’ai fait à celui qui se trouve dans l’autre salle, là-bas?”

   [French]
Increased perfect use in declaratives from Dutch onwards: past time reference.

- Portner (2003): the perfect answers the qud.
- Not all answers to the qud are conveyed in the present perfect, only those where the qud has current relevance.
- Discourse topic: events in the faraway past.
- Lifetime effects.

> ‘Well, they did rather detest each other. Not unlike yourself and Mr Malfoy. And then, your father did something Snape could never forgive.’
> ‘What?’
> ‘He saved his life.’
> ‘What?’

a. ‘He saved his life.’ [English]
b. "Han räddade hans liv." [Swedish]
c. "Hij heeft zijn leven gered." [Dutch]
d. "Er hat sein Leben gerettet." [German]
e. - Le salvó la vida. [Spanish]
f. “Il lui a sauvé la vie.” [French]
Extended PERFECT use in interrogatives and declaratives (from Dutch onwards)

- Topic negotiation in interrogatives requires an existential PERFECT reading.
- With existential PERFECKTS, the current relevance of the perfect state resides in the pragmatics (Nishiyama & Koenig 2010).
- **Hypothesis**: increased PERFECT use in more liberal languages implies a wider range of questions, because of the English presuppositionality constraint.
- Interaction with past time reference confirmed by extended PERFECT use in declaratives.
Increased PERFECT use in declaratives (German and French)

• In German and French, the odds are changing, because the frequency of PERFECT use in declaratives goes up faster than in interrogatives.
• Data indicate: sequences of PERFECT sentences more frequent in declaratives than in interrogatives.
• Portner/Nishiyama & Koenig focus on isolated PERFECT sentences. In order to account for sequences of PERFECT sentences we appeal to the rhetorical relation of Elaboration (de Swart 2007).
Increased PERFECT use in declaratives: the rhetorical relation of Elaboration

• De Swart (2007): the PERFECT creates an Elaboration structure in which the speech time S, or the larger utterance situation provides the topic under discussion, and the sentence in the PERFECT elaborates on this topic.

‘I met him when I travelled around the world. A foolish young man I was then, full of ridiculous ideas about good and evil. Lord Voldemort showed me how wrong I was. There is no good and evil, there is only power, and those too weak to seek it... Since then, I have served him faithfully, although I have let him down many times. He has had to be very hard on me.’

• If we find a series of sentences in the PERFECT, which elaborate a particular speech situation, the rhetorical relation between those sentences is the relation of Continuation.
Continuation creates a list reading

- The relation of Continuation is a very weak rhetorical relation: it just connects the sentences in an Elaboration structure.
- Continuation has no temporal implications in and of itself: sequences of perfect sentences in English, Swedish, Dutch, Spanish elaborating a topic describe an unordered set of events (list reading).

a. Since then, I have served him faithfully, although I have let him down many times. He has had to be very hard on me.” [English]

b. Sen dess har jag tjänat honom troget, fastän jag har svikit honom många gånger. Han har blivit tvungen att vara mycket hård mot mig. [Swedish]

c. Sindsdien heb ik hem trouw gediend, maar ook vaak teleurgesteld. Hij heeft me soms hard moeten aannemen. [Dutch]

d. Desde entonces le he servido fielmente, aunque muchas veces le he fallado. Tuvo que ser muy severo conmigo. [Spanish]
From Continuation to temporal structure

- De Swart (2007): **Continuation** does not induce temporal structure, but it is not incompatible with temporal structure: cross-linguistic variation. English, Swedish, Dutch, Spanish: temporal relations between **PERFECTS** are blocked (list reading only).
- German, French: temporal relations between **PERFECTS** are allowed: from continuation to narrative sequences through ‘natural’ connections between events: scripts, scenarios.

a. Your friend Miss Granger accidentally **knocked** me over as she **rushed** to set fire to Snape at that Quidditch match. She **broke** my eye contact with you. [English]

b. Din väninna miss Granger **slog** omkull mig av en händelse då hon **rusade** i väg för att sätta eld på Snape vid den där quidditchmatchen. Hon **bröt** min ögonkontakt med dig. [Swedish]

c. Juffrouw Griffel **liep** me per ongeluk omver toen ze Snape in brand **wilde** steken tijdens die Zwerkbalwedstrijd en daardoor **verbrak** ze mijn oogcontact met jou. [Dutch]

d. Ihre Freundin Miss Granger **hat** mich versehentlich **umgerempelt**, als sie beim Quidditch-Spiel zu Snape **hinüberrannte**, um ihn anzuzünden. Sie **hat** meinen Blickkontakt zu Ihnen **unterbrochen**. [German]

e. Tu amiga, la señora Granger, accidentalmente me **atropelló** cuando **corría** a prenderle fuego a Snape, en ese partido de Quidditch. Y **rompió** el contacto visual que yo tenía contigo. [Spanish]

f. Votre amie, Miss Granger m’ **a bousculé** par accident quand elle **s’est précipitée** pour mettre le feu aux vêtements de Rogue, pendant le match de Quidditch. A cause d’elle, j ’**ai perdu** le contact visuel avec vous. [French]
Series of **PERFECTS** in declaratives: rhetorical relation of **Elaboration** – list reading or narration

a. He *does not forgive* mistakes easily. *When I failed* to steal the Stone from Gringotts, he *was* most displeased. He *punished* me... *Decided* he would have to keep a closer watch on me...

b. Han *föröver* inte lätt några misstag. *När jag misslyckades* med att stjäla Stenen från Gringotts *blev* han mycket missnöjd. Han *straffade* mig ... *beslöt* att han måste bevaka mig närmare "...

c. Hij *vergeeft* niet snel. *Toen het me niet lukte* om de Steen te stelen bij Goudgrijp, *was* hij heel erg kwaad. Hij *heeft me gestraft* ... *besloten* me nauwlettender in het oog te houden "...

d. Fehler *vergibt* er nicht so einfach. *Als es mir nicht gelungen ist*, den Stein aus Gringotts zu stehlen, *war* er höchst ungehalten. Er *hat mich bestraft* ... und *beschlossen*, mich näher im Auge zu behalten "...

e. No *perdona* fácilmente los errores. *Cuando fracasé* en robar esa Piedra de Gringotts, se *disgustó* much. Me *castigó*... *decidió* que tenía que vigilarme muy de cerca... [Spa]

b. Il ne *pardonne* pas facilement les erreurs. *Le jour où je n’ai pas réussi* à voler la Pierre, à Gringotts, il *était* très mécontent. Il m’a *puni*. Et il *a décidé* de me surveiller de plus près... [Fre]
Conclusions on the pragmatics of the **PERFECT**

- ‘Classical’ **PERFECT use** associated with answering a question under discussion (Portner 2003, declarative) and topic negotiation (Nishiyama & Koenig 2010, interrogative). In both cases: existential **PERFECT**, current relevance. Cross-linguistically stable pragmatics.

- Starting with Dutch, languages extended **PERFECT use** across both sentence types. **Pragmatic ingredients of variation**: presuppositionality, life time effects, interaction with past time reference.

- More liberal uses German and French: **series of **PERFECT** sentences** in declaratives: rhetorical relation of **Elaboration** (de Swart (2007), with possible temporal relations between **PERFECTS** ∼ narrative sequencing.

- The three pragmatic roles of the **PERFECT** require the utterance situation as a deictic centre: association with **spoken language**, not written language.
Overall conclusions

• The *HAVE PERFECT* is subject to substantial *cross-linguistic variation*, but languages with a more liberal *PERFECT* subsume the classical *PERFECT* uses (subset relation) ~ extended *PERFECT* semantics rather than switch to a *PERFECTIVE PAST*.

• **Semantic ingredients** of the *PERFECT* scale: lexical semantics (stative verbs), compositional semantics (past time reference), dynamic semantics (narrative sequences).

• **Pragmatic ingredients** of the *PERFECT*: answering the QUD (declaratives), topic negotiation (interrogatives), presuppositionality constraint, rhetorical relation of *Elaboration* (sequences of declarative *PERFECT* sentences).

• The *PRESENT PERFECT* is a tense-aspect category that belongs to the *spoken language grammar* in all languages with a *HAVE PERFECT*. The three pragmatic roles of the *PERFECT* associate with spoken language (deictic centre).