
Time in Translation: Past, Present & Future

Time in Translation



Write a semantics of the PERFECT.

- > Five languages (3 Germanic, 2 Romance)
 - > Account for distribution at the sentence level, in narration and in dialogue
 - > Methodology: back and forth between theory and data



Translation Mining (corpus-based)

Time in Translation



Henriëtte de Swart



Bert Le Bruyn



Martijn van der Klis (PhD)



Anja Goldschmidt (Post-Doc)



Jos Tellings (Post-Doc)

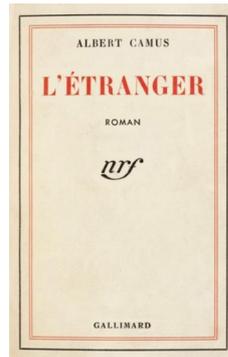


Year 1: Experimenting with Translation Mining

Translation Mining

Data from translation corpora allow us to compare tense/aspect systems of languages and study them in detail.

Example : translations of Camus' *L'Étranger*, Chapter 1



FRENCH

Perfect

Imperfective

Present

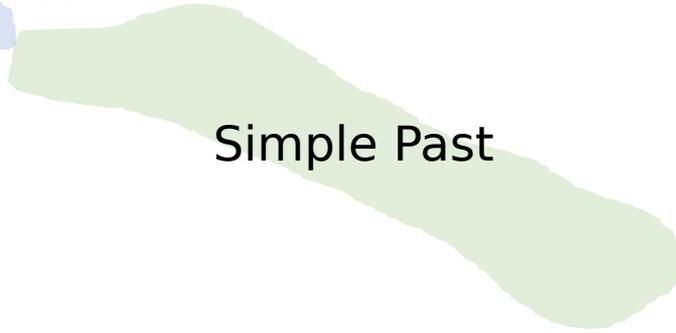
PluPerfect

ENGLISH

Perfect



Simple Past



SPANISH

Perfect

Simple Past

DUTCH

Perfect

Simple Past

GERMAN

Perfect

Simple Past

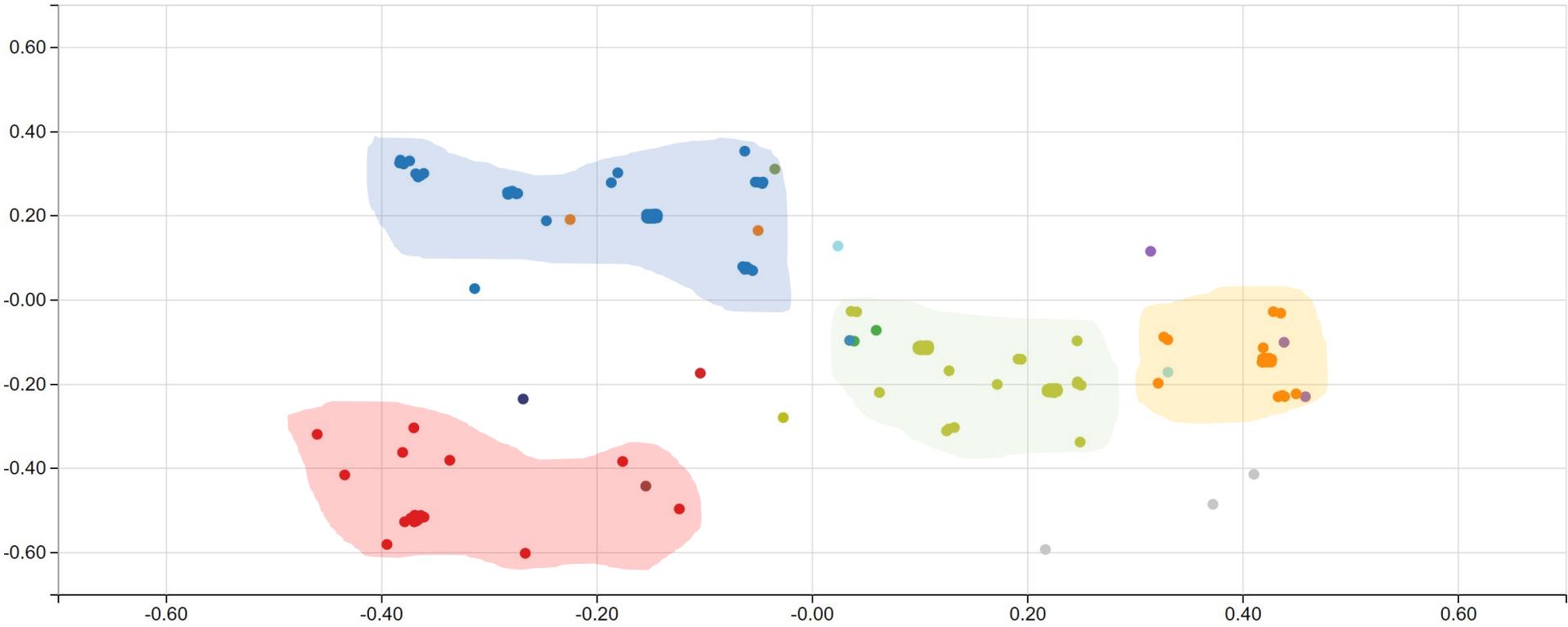
Translation Mining

Using data from translation corpora to compare tense/aspect systems of languages and study them in detail.

Example : translations of Camus' *L'Étranger*, Chapter 1

> Bigger picture but also strong grip on the data

FRENCH



French

passé composé (fr)

1.xml - 24774

J' **ai lu** le dossier de votre mère .

Translations

German

Perfekt

Ich **habe** die Akte Ihrer Mutter **gelesen** .

English

present perfect

I **'ve read** your mother 's file .

Spanish

pretérito perfecto compuesto

He leído el expediente de su madre .

Italian

passato prossimo

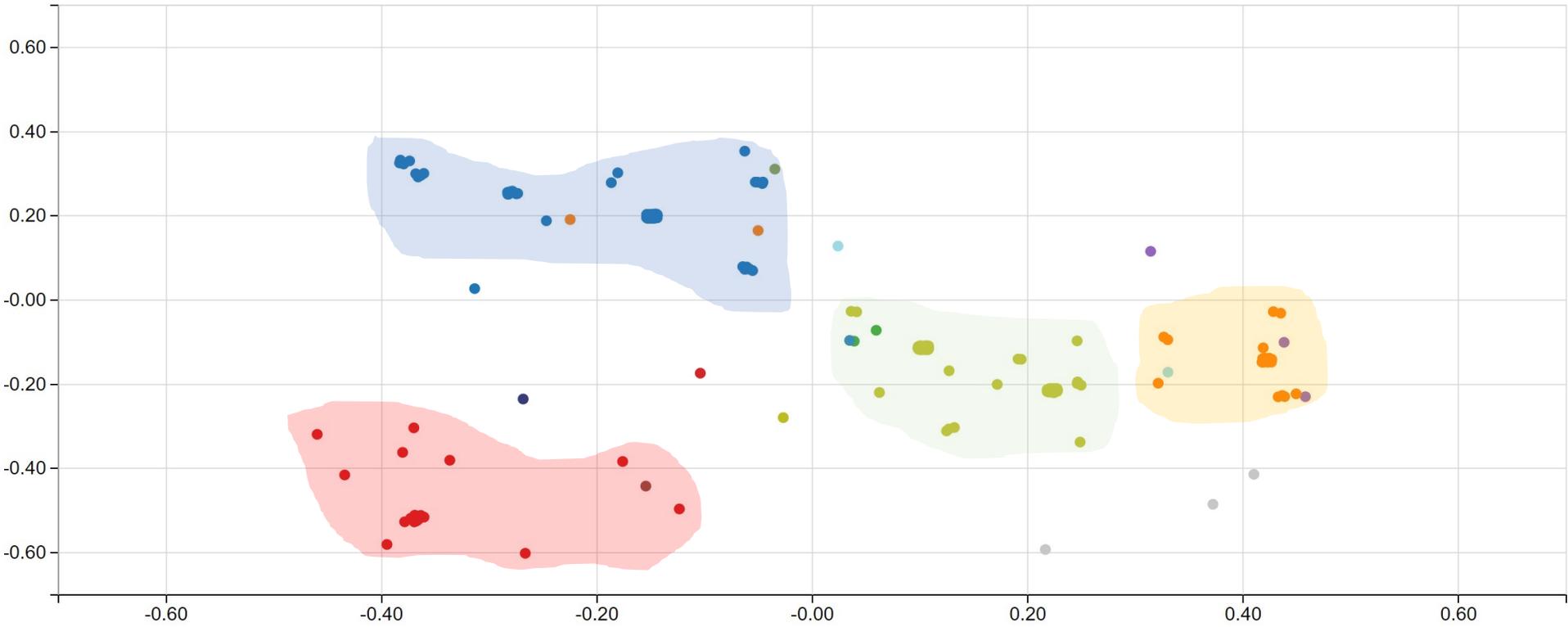
Ho letto la pratica di vostra madre .

Dutch

vtt

Ik **heb** uw moeders map **doorgelezen** .

FRENCH



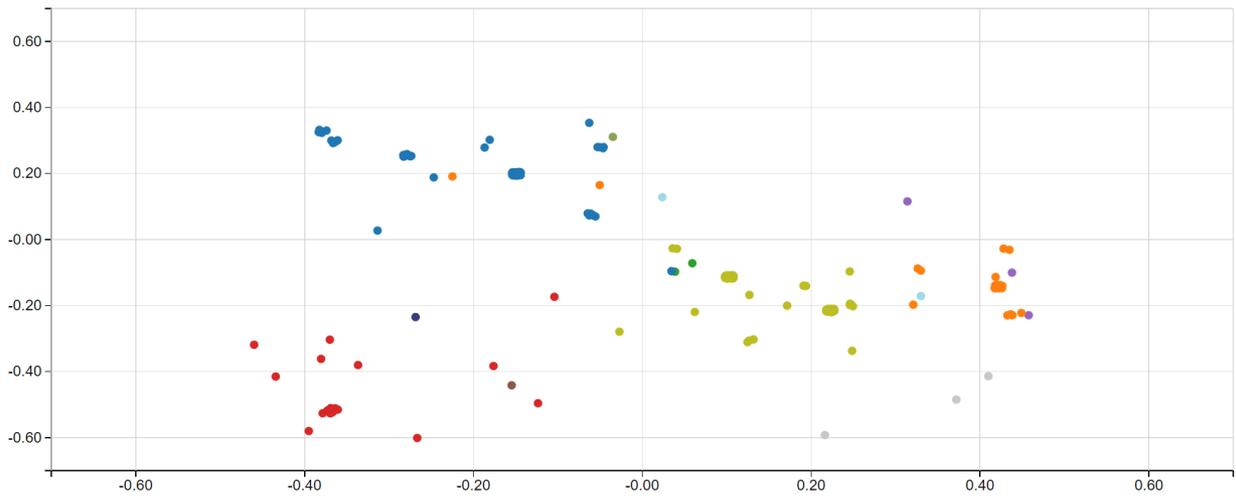
MDS visualization (scenario *Camus - all tenses - chapter 1*)

Stress:

This scatter plot shows the results Multidimensional Scaling of the dissimilarity matrix of tense tuples. Please see the **project summary** for more details.

- passé simple
- futur simple
- futur proche
- conditionnel passé
- conditionnel présent
- passé composé
- présent
- plus-que-parfait
- futur antérieur
- imparfait

futur proche du passé



Filters

Below, you can change the labels and the dimensions displayed on the x- and y-axis. Click the green button to confirm your selection and update the scatter plot.

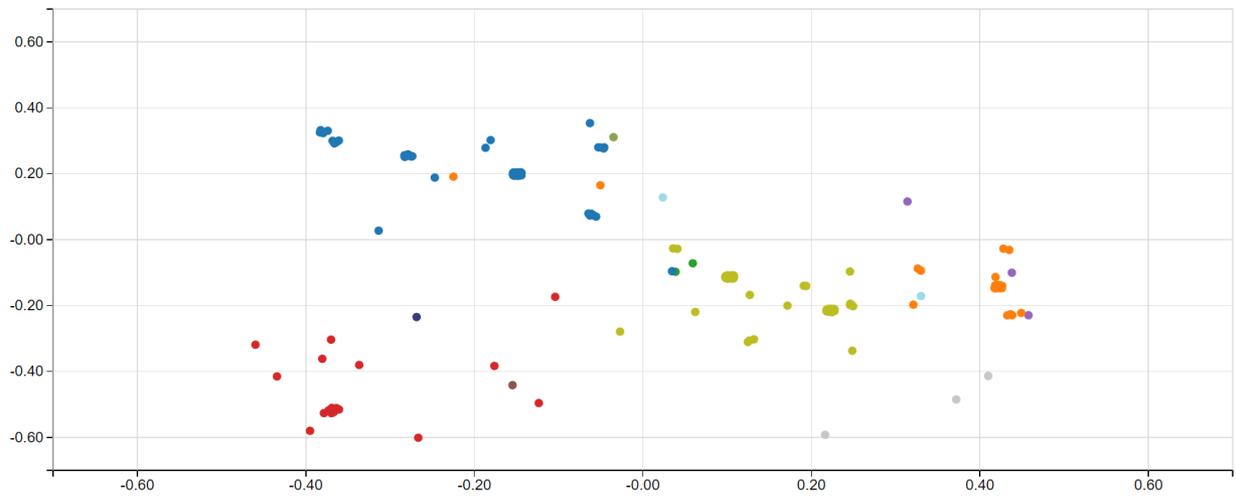
Language: Dimension on x-axis: Dimension on y-axis:

MDS visualization (scenario *Camus - all tenses - chapter 1*)

Stress:

This scatter plot shows the results Multidimensional Scaling of the dissimilarity matrix of tense tuples. Please see the **project summary** for more details.

- passé simple
- futur simple
- futur proche
- conditionnel passé
- conditionnel présent
- passé composé
- présent
- plus-que-parfait
- futur antérieur
- imparfait
- futur proche du passé



Filters

Below, you can change the labels and the dimensions displayed on the x- and y-axis. Click the green button to confirm your selection and update the scatter plot.

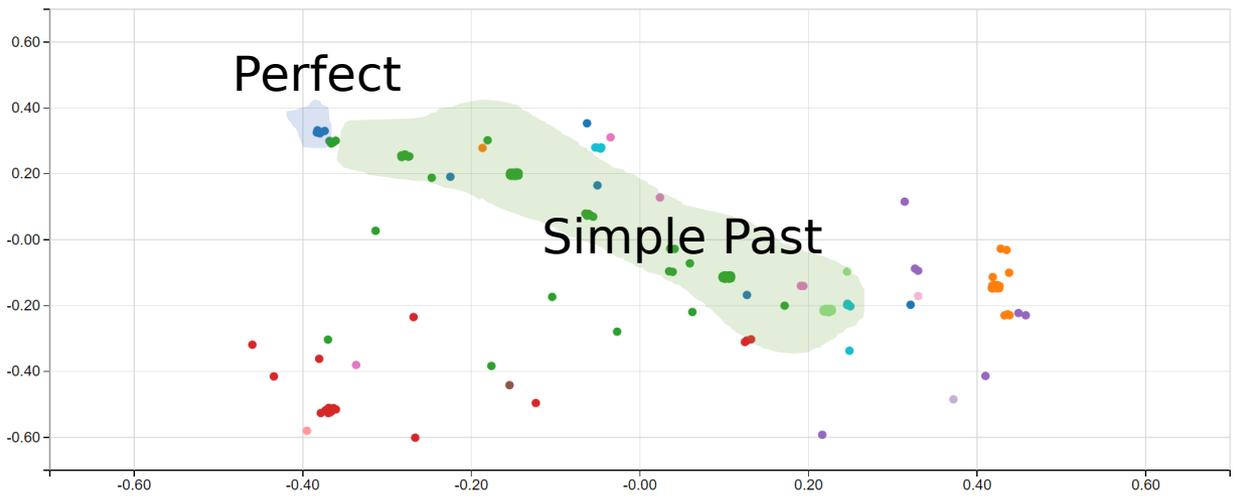
Language: Dimension on x-axis: Dimension on y-axis:

MDS visualization (scenario *Camus - all tenses - chapter 1*)

Stress:

This scatter plot shows the results Multidimensional Scaling of the dissimilarity matrix of tense tuples. Please see the **project summary** for more details.

- past perfect continuous
- simple future
- future perfect
- future in the past
- future in the past continuous
- past continuous
- present participle
- present perfect
- simple present
- simple past
- past perfect
- future continuous



Filters

Below, you can change the labels and the dimensions displayed on the x- and y-axis. Click the green button to confirm your selection and update the scatter plot.

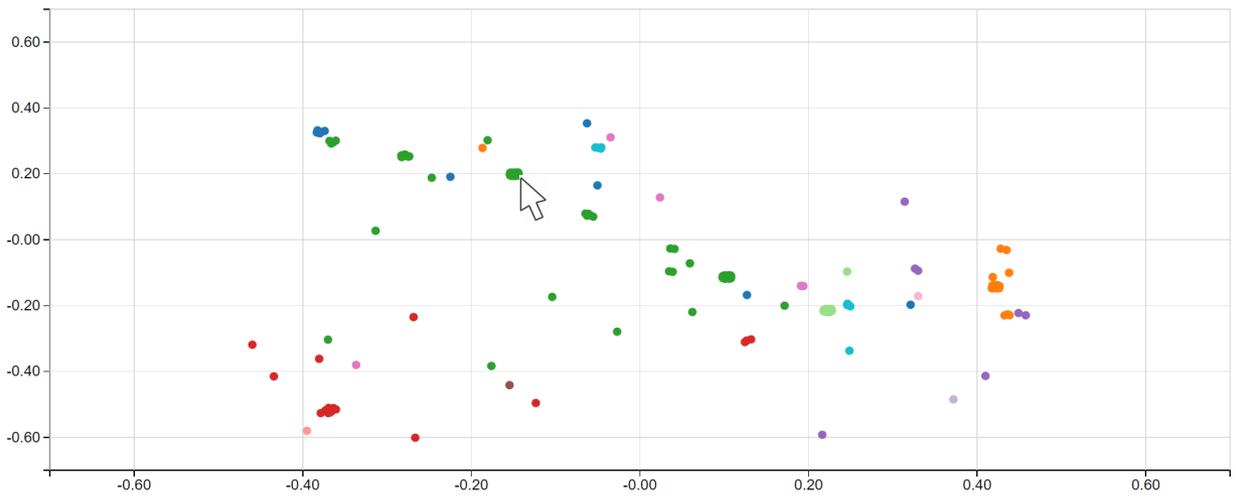
Language: Dimension on x-axis: Dimension on y-axis:

MDS visualization (scenario *Camus - all tenses - chapter 1*)

Stress:

This scatter plot shows the results Multidimensional Scaling of the dissimilarity matrix of tense tuples. Please see the **project summary** for more details.

- past perfect continuous
- simple future
- future perfect
- future in the past
- future in the past continuous
- past continuous
- present participle
- present perfect
- simple present
- simple past
- past perfect
- future continuous



Filters

Below, you can change the labels and the dimensions displayed on the x- and y-axis. Click the green button to confirm your selection and update the scatter plot.

Language: Dimension on x-axis: Dimension on y-axis:

Fragment overview

Source

French **passé composé** **1.xml**
J' ai dormi pendant presque tout le trajet .

Translations

German **Perfekt**
Ich habe fast während der ganzen Fahrt geschlafen .

English **simple past**
I slept almost all the way .

Spanish **pretérito indefinido**
Dormí durante casi todo el trayecto .

Italian **passato prossimo**
Ho dormito quasi tutto il percorso .

Dutch **vtt**
De hele rit bijna heb ik geslapen .

Polish
No annotation available.

Russian
А когда проснулся , то оказалось , что голова моя лежит на плече какого-то военного , моего соседа ; он мне улыбнулся и спросил , издалека ли я еду .

Ukrainian
No annotation available.

Mandarin **le1**
我几乎睡了一路。

Hebrew **past**
שמתי כמעט כל הדרך .

Greek **aorist**
Κουάθηκα ανεδόν σ ' όλη τη διαδρομή

Estonian **preteeritum**
Tukkusin peaaegu kogu tee

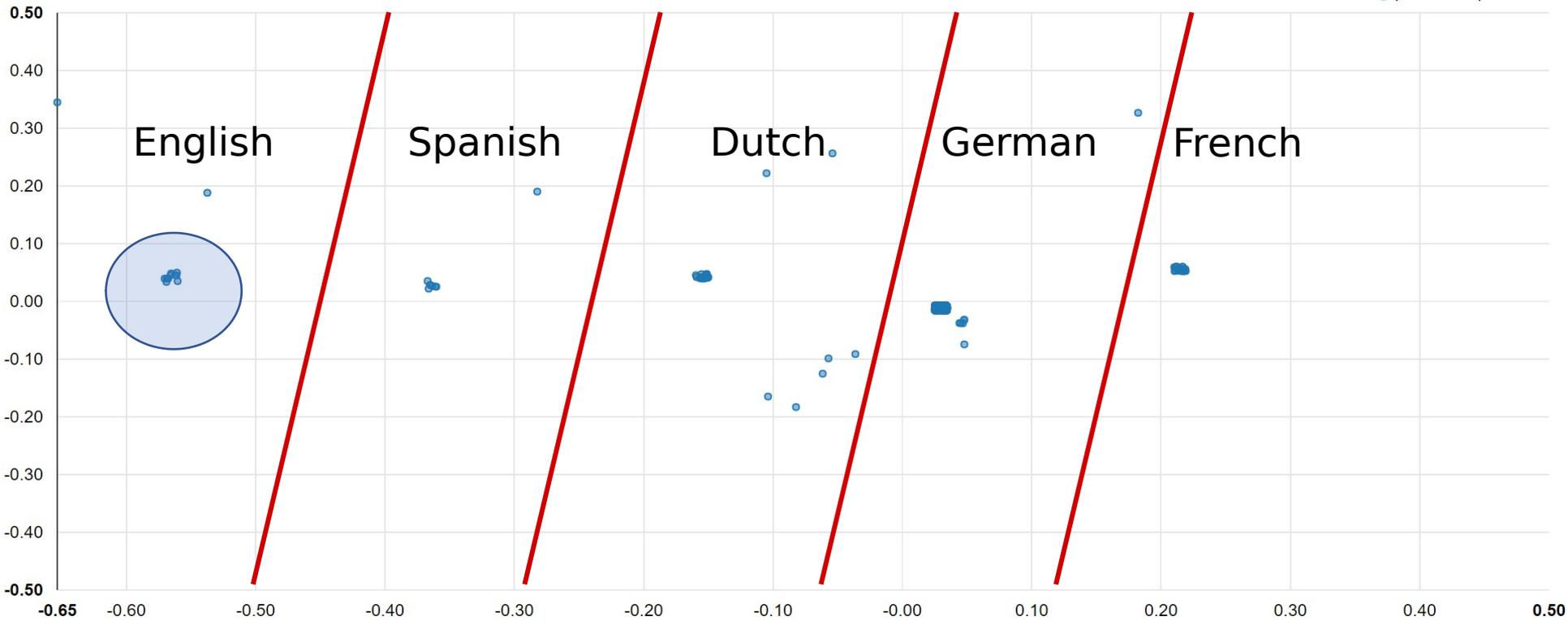
Translation Mining

Using data from translation corpora to compare tense/aspect systems of languages and study them in detail.

Example: translations of Camus' *L'Étranger*, Chapter 1

- > Bigger picture but also strong grip on the data
- > Sliding scale

English < Spanish < Dutch < German < French



English

I've read your mother's file.

We've transferred her to our little mortuary.

I've taken it upon myself to make the necessary arrangements.

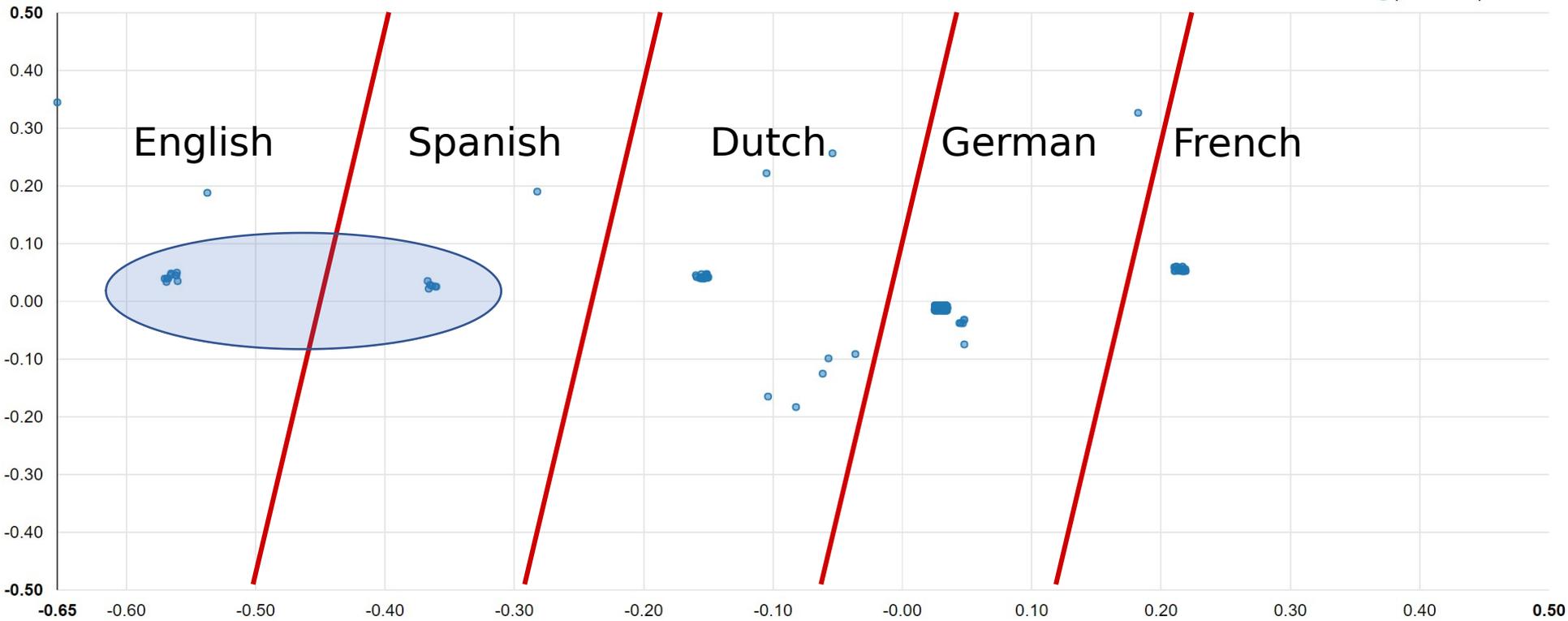
I saw them more clearly than I've ever seen anyone.

In eight years they haven't changed their route.

Classical perfects Past event

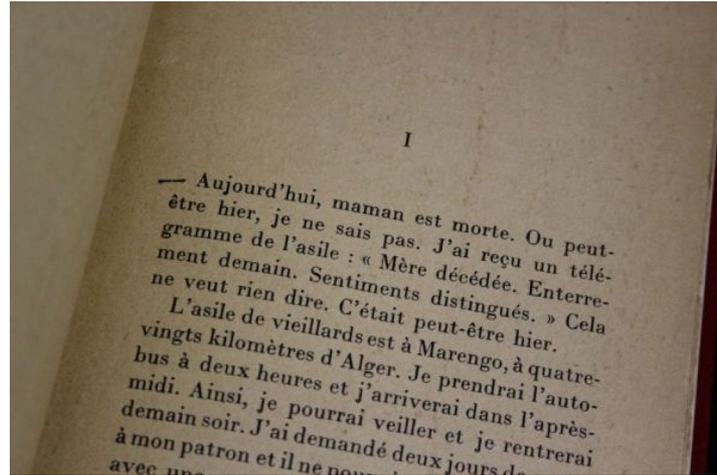
State holding at the present /
current relevance

English < Spanish < Dutch < German < French



English < Spanish

Pragmatics



Mother **died** today.

Hoy, mamá **ha muerto**.

green simple past

blue present perfect

English < Spanish

Pragmatics

We **covered** her up.

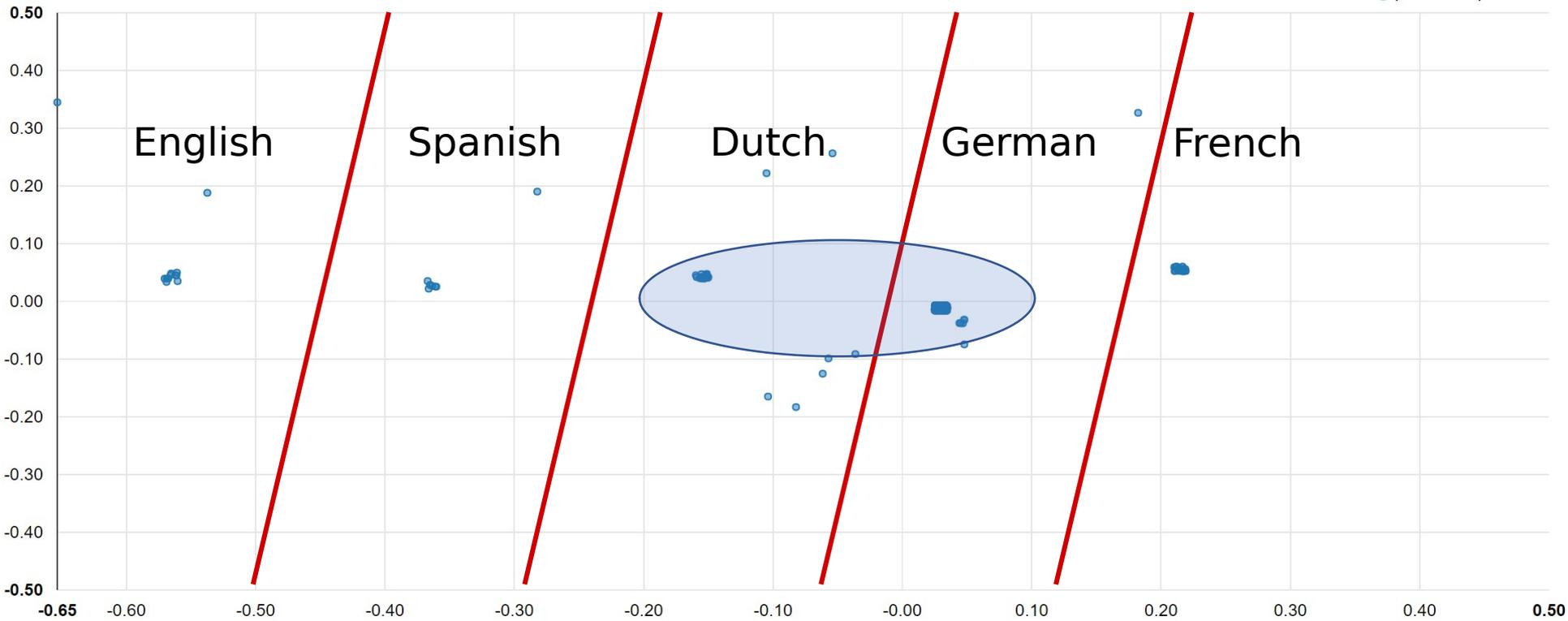
La **hemos cubierto**.



Look mommy, I **drew** a picture of you!

Look mommy, I **ve drawn** a picture of you!

English < Spanish < Dutch < German < French



Dutch < German

After that everything **happened** so quickly...

Todo **pasó** después con tanta precipitación...

Daarna **is** alles zo snel... **gegaan**...

Danach **ist** alles so überstürzt... **abgelaufen**...

I **slept** almost all the way.

Dormí durante casi todo el trayecto.

De hele rit bijna **heb** ik **geslapen**.

Ich **habe** fast während der ganzen Fahrt **geschlafen**.

Dutch < German

Discourse semantics

After that everything **happened** so quickly...

Todo **pasó** después con tanta precipitación...

Daarna **is** alles zo snel... **gegaan**...

Danach **ist** alles so überstürzt... **abgelaufen**...

I **slept** almost all the way.

Dormí durante casi todo el trayecto.

De hele rit bijna **heb** ik **geslapen**.

Ich **habe** fast während der ganzen Fahrt **geschlafen**.

I 'm very fond of white coffee , so I **accepted** and he **came** back a few minutes later with a tray.

Como me gusta mucho el café con leche, **acepté**, y, al cabo de un momento, **volvió** con una bandeja .

Aangezien ik dol op koffie ben **nam** ik zijn aanbod aan, en een ogenblik later **kwam** hij terug met een blad.

Da ich Milchkaffee sehr gern mag, **habe** ich **angenommen**, und er **ist** nach einer Weile mit einem Tablett **zurückgekommen**.

Reflections

We have experimented with several parallel corpora (EuroParl, DCEP, OpenSubtitles, ...)

We find that translations of novels provide us with the best control over the data (one writer per novel, one translator per novel, clear source and target language, ...), especially because we assume we don't know the linguistic variables yet. EuroParl has become our corpus of choice to check predictions as soon as we think we have identified a relevant variable.

Reflections

We are currently looking into the (statistical) status of findings in novel-based corpora and the way our findings relate to those in typological work that uses a similar methodology.

Translation Mining has been about surveying ingredients. The analysis still has to come :-)

Next to a corpus component, the project also contains a native speaker component :-)

Year 2: Exploring dialogue

Two insights from the literature

Portner (2003)

The use of the perfect **presupposes** that the proposition it is in is a (partial or complete) answer to the question under discussion.

How are you? I've been diagnosed with cancer.
?you? I was diagnosed with cancer.
cancer.

What happened? Mom died.

Two insights from the literature

Nishiyama & König (2010)

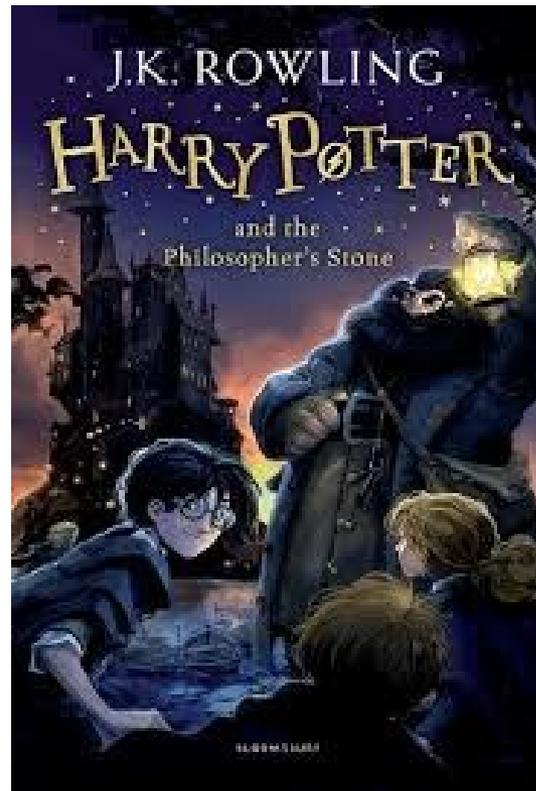
Speakers sometimes use a perfect at the beginning of a conversation to set up a topic.

Have you seen *Dances with Wolves*?

? Did you see *Dances with Wolves*?

Where were you on 9/11?

Back to the data

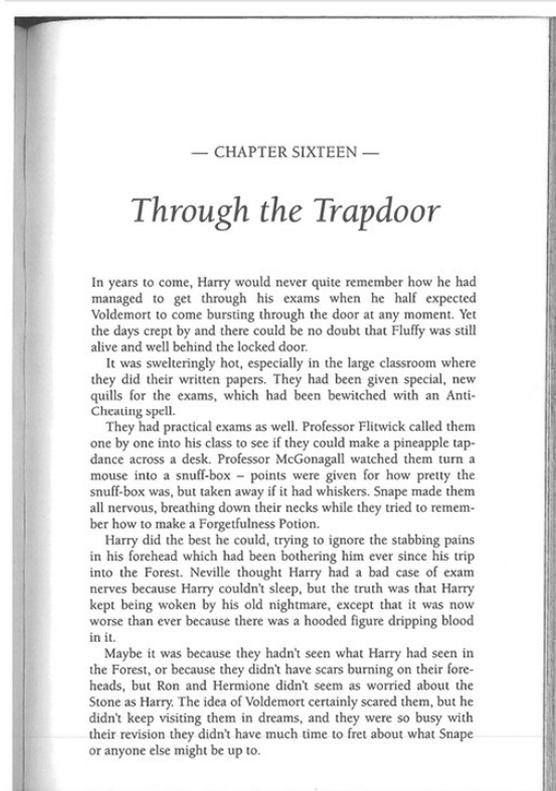
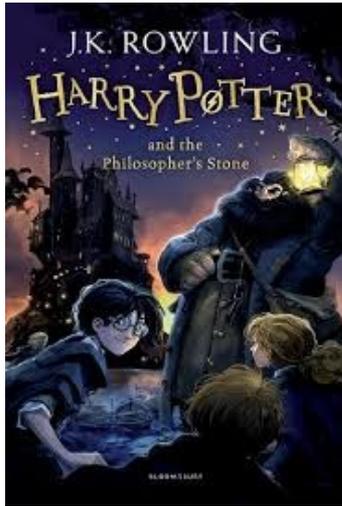


	en	de	es	fr	it	nl
PRESENT	6	6	4	3	6	5
PERFECT	0	0	0	0	8	0
PAST	430	461	336	313	316	465
IMPERFECT	0	0	126	130	112	0
PAST PERFECT	29	27	18	30	24	32

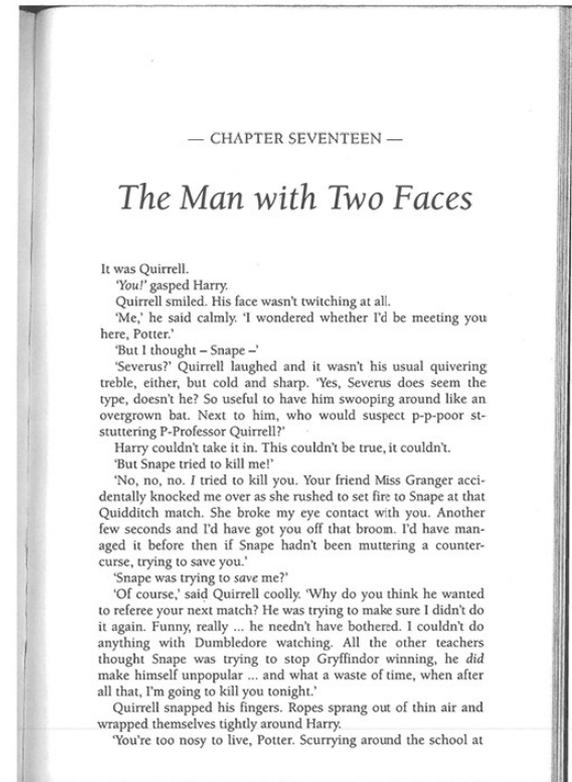
Tab. 1: Most frequently attributed tense categories in the **narrative** parts of the novel.

	en	de	es	fr	it	nl
PRESENT	126	154	122	129	129	162
PERFECT	27	78	23	99	100	60
PAST	120	64	88	0	2	84
IMPERFECT	0	0	38	50	39	0
FUTURE	20	15	13	11	23	9

Tab. 2: Most frequently attributed tense categories in the **dialogue** parts of the novel.



60% is present oriented



36% is present oriented

The Man with Two Faces

It was Quirrell.

'You!' gasped Harry.

Quirrell smiled. His face wasn't twitching at all.

'Me,' he said calmly. 'I wondered whether I'd be meeting you here, Potter.'

'But I thought – Snape –'

'Severus?' Quirrell laughed and it wasn't his usual quivering treble, either, but cold and sharp. 'Yes, Severus does seem the type, doesn't he? So useful to have him swooping around like an overgrown bat. Next to him, who would suspect p-p-poor st-stuttering P-Professor Quirrell?'

Harry couldn't take it in. This couldn't be true, it couldn't.

'But Snape tried to kill me!'

'No, no, no. *I* tried to kill you. 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. Another few seconds and I'd have got you off that broom. I'd have managed it before then if Snape hadn't been muttering a counter-curse, trying to save you.'

Storyline in dialogue

→ Splitting up dialogue further

141 verbs in 'storyline'

77 verbs in 'pure dialogue'

And what can we say about dialogue?

English 77 verbs 59 simple pasts

18 perfects

→ Same contexts as in Camus (= those known from the literature on sentence-level uses)

Dutch 77 verbs 12 simple presents

33 simple pasts

30 perfects

And what can we say about dialogue?

Dutch 12 simple presents

33 simple pasts

30 perfects

Hij wist alles al . Hij zei alleen : ' Harry *is* achter hem aan , hè ? ' en rende direct naar de derde verdieping . "

" Maar ik hoorde u een paar dagen geleden snikken - ik *dacht* dat Sneepe u bedreigde ... "

" *Hebt* u die trollen *binnengelaten* ? " You *let* the troll in ? '

He already knew - he just said , " Harry *'s* *gone* after him , hasn 't he ? " and hurtled off to the third floor .

' But I heard you a few days ago , sobbing - I *thought* Snape was threatening you ... '

Reflections

These are just the initial stages...

Relevant factors we're investigating (next to aspectual class): rhetorical relations, sequence of tense, modality, syntax, etc.

In parallel, we're looking into building tense/aspect extensions of conversation frameworks like inquisitive environments. **See Jos' SALT talk in May** □

The data already show that sentence, discourse and dialogue interact. Our final analysis will therefore – by necessity – incorporate all three levels.

Beyond

More languages
More areas of grammar
SLA

Beyond

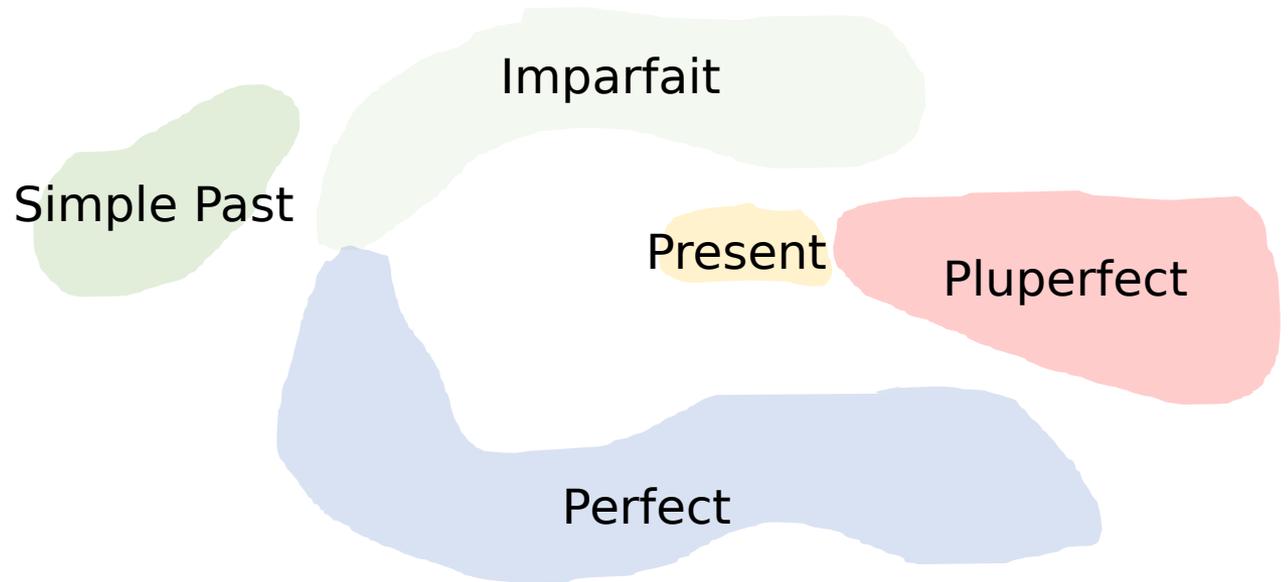
More languages

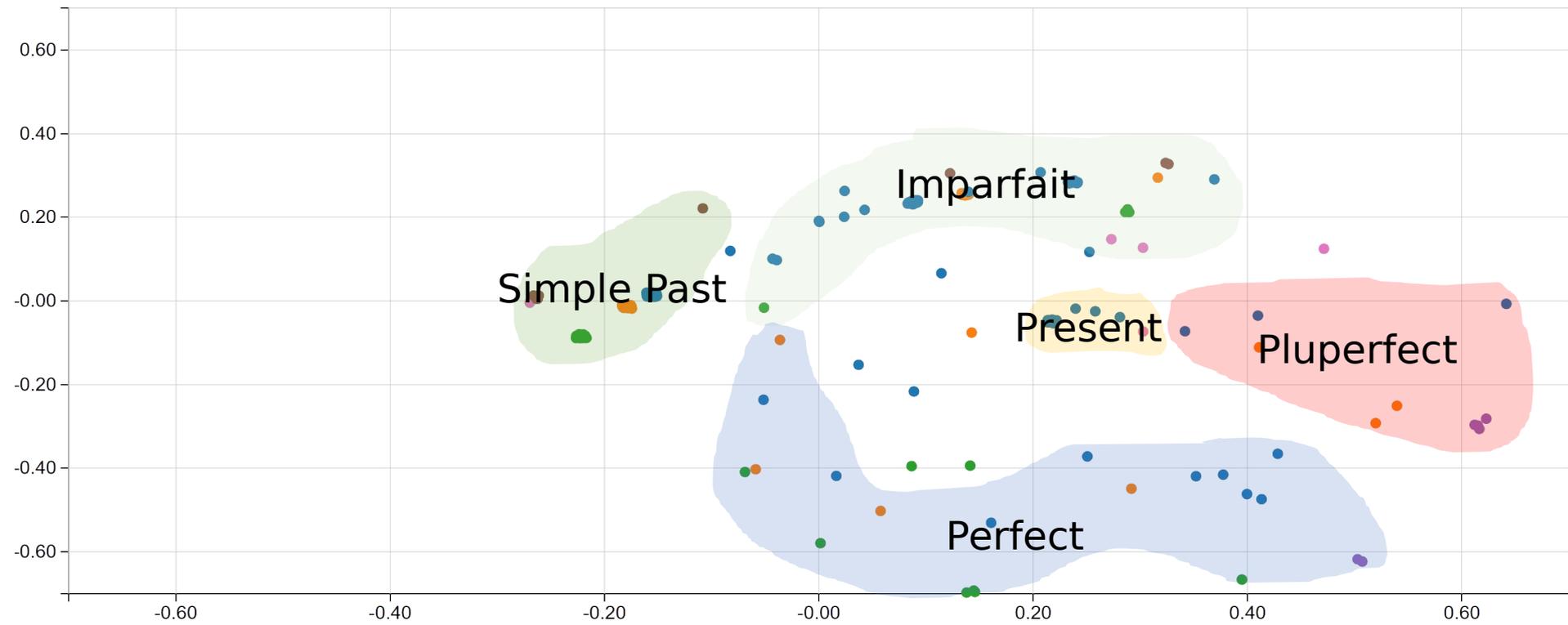
More areas of grammar

SLA

Example: translations of Harry Potter, Book 1, Chapter 1

FRENCH

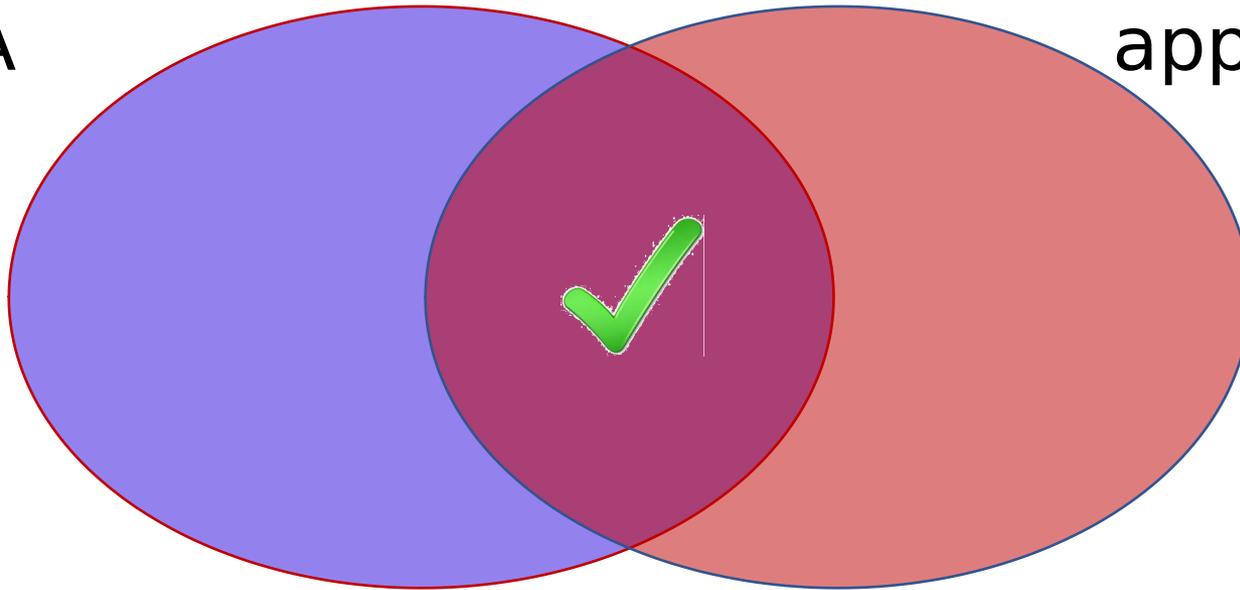




Translation Mining and SLA

GenSL
A

Usage-based
approaches



GenSLA

80s-90s

> Basic question: is there UG access in SLA?

> Role of Poverty of the Stimulus Argument and Transfer.

More recent developments

> Interface issues.

> Interpretable and uninterpretable features.

> Role of input and how it relates to languages that are already present in the learner's mind.

GenSLA

> Role of input and how it relates to languages that are already present in the learner's mind.

Learner starts with a coarse-grained hypothesis on how the new language compares to those that s/he already knows.

Based on what?

When?

Learner refines the hypothesis on the basis of input.

When?

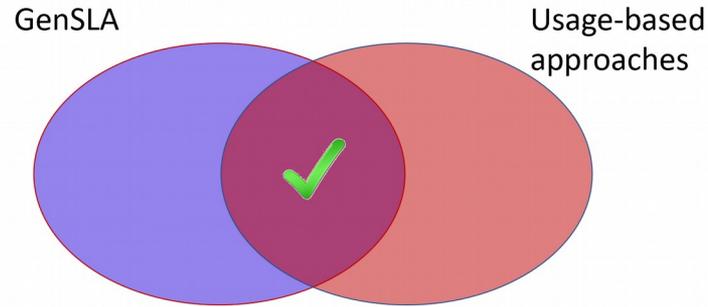
How?

Basic issue: what, when, why, how do we generalize (i.e. learn a language)?

Usage-based approaches

- > Frequency, contingency, prototypicality, ...
 - > Clear awareness of role of languages that are already present in the learner's mind but little hypothesizing about interaction with input.

SLA



Basic issue: what, when, why, how do we generalize (i.e. learn a language)?

Basic data we need:

Mapping data **similar** to those we get through Translation Mining

Envoi

Thanks for your attention!

And visit us at <https://time-in-translation.hum.uu.nl/>