

THE PERFECT IN SWEDISH

The PERFECT in narrative and dialogue compared in Swedish, English, German, and Dutch, for improvement of Natural Language Processing

Marthe Visser

First Supervisor

dr. Jos Tellings

Second Supervisor

dr. Bjørn Jespersen

BSc Kunstmatige Intelligentie

7,5 ECTS

June 28, 2019



Utrecht University

Table of Contents

- 1. Introduction 3**
 - 1.1 Relationship with Artificial Intelligence 3*
 - 1.2 Academic relevance 4*
 - 1.3 Hypothesis 5*

- 2. Method 7**
 - 2.1 Arrangement of the Swedish text 7*
 - 2.2 TimeAlign 7*
 - 2.3 Tense assignment 8*
 - 2.4 Mapping the tenses 8*

- 3. Results 9**
 - 3.1 Narrative 10*
 - 3.2 Dialogue 15*

- 4. Discussion 19**
 - 4.1 Narrative 19*
 - 4.2 Dialogue 19*
 - 4.3 Implications for AI 21*
 - 4.4 Points of improvement 21*

- 5. Conclusion 23**

- 6. References 24**

1. Introduction

The Time in Translation (TinT) project by prof. dr. Henriëtte de Swart and dr. Bert Le Bruyn (started in 2017) focuses on the semantics of the PERFECT. There are two forms of the PERFECT in English: the Past Perfect and Present Perfect. The Present Perfect (now referred to as just PERFECT) will be the main focus of this paper. In English it is marked by the auxiliary HAVE, followed by a past participle. In other languages, German and Dutch in this project, the auxiliary BE can also be used. The PERFECT mostly conveys the meaning of a past event together with a present state. The aim of the TinT-project is to have more insight in the semantics and pragmatics of the PERFECT and its behaviour cross-linguistically. This will be achieved by corpus comparison, to bring together computational and theoretical linguistics.

Languages can differ in their use of grammatical tenses, while the meaning of a translated sentence remains the same throughout different languages. Likewise, the use of the PERFECT at the sentence and discourse level varies across languages. Moreover, the PERFECT is often translated to the past and present tenses¹. Currently, the project has mapped usage of the PERFECT in English, Dutch, German, French, Italian, and Spanish. The aim of this research is to further add to this list with the start of research on the PERFECT in Swedish in narrative and dialogue. The adding of Swedish to this project is relevant, because it will give more insight in the use of the PERFECT throughout. In this research I will compare the use of the PERFECT in four Germanic languages.

My research question is how the PERFECT is used in Swedish narrative and dialogue, compared to the use of the PERFECT in English, German, and Dutch in the narrative and dialogue.

To answer this research question, I will first look into the relevance of this research in the field of Artificial Intelligence and the already existing literature about this topic. This will help to formulate the hypothesis. Chapter two describes the method used for this research. Chapter three will discuss the results of the research, divided between results from the narrative and dialogue. The discussion will elaborate the results further, together the implications they make in the field of Artificial Intelligence. Moreover, the discussion will provide points of improvement for further research. Lastly, the conclusion will answer the research question.

1.1 Relationship with Artificial Intelligence

This study is relevant in the field of Artificial Intelligence (AI) since one of the goals of AI is to improve Natural Language Processing (NLP). Enhancement of NLP would ultimately give computers the ability to process and create human language (Jurafsky & Martin, 2009). Facets like human-machine communication, real-time speech-to-speech translations, and automated machine-translation (of, for example, news reports, and books) would all benefit from more semantic research.

Languages are infinite, that is, one can utter a sentence no one else has ever said or heard before (Saxton, 2010). Since languages are infinite, a translation machine cannot just learn all sentences and their translation to every other language. Jurafsky and Martin (2009) describe how the structural and lexical differences of languages make translation very difficult. Until recently, Statistical Machine Translation was the dominant translation method, Google Translate used it for example till 2016 (Wu et al., 2016). These statistical based translation models do not consider the meaning of a sentence or text, but only look at the statistical chances of a certain translation per word. Since also the verbs are statistically translated, I can imagine there is a high change for the translated verbs to be in the same (but maybe not preferred) tense as the original text. Neural Machine Translation (NMT) is an upcoming approach for improved automated translation. These NMT models translate parts of a text at once, for example sentences as a whole, rather than word by word, using a learning artificial neural network (Wu et al., 2016). This way, the model can review the translation of the sentences and its verb tenses properly. However, NMT systems are still computationally very expensive and therefore not yet suited for widespread usage (Wu et al., 2016).

Awaiting the development and affordability of these self-learning NMT models, I argue we need to give translation models more specific translation information. This way, they can carry out better translations than statistical models, but do not need to extract and learn the rules themselves as NMT models. With regard to translation of the PERFECT, the TinT-project can contribute to more insight in the use of the PERFECT throughout different languages. This can improve automated machine translation from a text in one language to another and real-time speech-to-speech translation. Moreover, more knowledge about the use of the PERFECT can help develop more sophisticated human-machine communication (Jurafsky & Martin, 2009).

¹ <https://time-in-translation.hum.uu.nl/project/>

1.2 Academic relevance

In this section relevant previous research will be discussed, to help formulate a hypothesis. The verb phrases that are considered will be displayed in bold.

In (1a), the event of Anna leaving is before the point of speech. In (1b), the importance of the speech time is highlighted because the sentence tells that Anna left with the result that she is not at the party at the moment (De Swart, 2007). In order to account for the relation between the Simple Past (1a) and the PERFECT (1b), Reichenbach (1947) formulates three points in time, the point of event (E), the point of reference (R), and the point of speech (S). In the Simple Past, E is simultaneously with R, indicated with a comma. E,R precedes the point of speech S, indicated with a hyphen. Thus, according to Reichenbach, the structure of Simple Past in English is E,R-S. Reichenbach proposes that the PERFECT indicates E-R,S, that is, E precedes R, and R and S coincides.

- (1) a. Anna **left** the party.
b. Anna **has left** the party.

There can be three forms of the PERFECT distinguished (Portner, 2003 and Van der Klis, Le Bruyn & De Swart, 2017). The experiential PERFECT, as in (2), says something that is relevant at the moment of speech, but took place before the time of reference. Moreover, in English and Swedish, the experiential PERFECT displays so called 'life time effects' (Rothstein, 2008). That is, the sentences in (2a) and (2b) strongly suggest that Donna is still alive, and would have been very odd if the subject was for example Einstein. Dutch and German do not have this life time effect.

The resultative PERFECT, as in (3), indicates the event precedes the reference time, but with a current relevance (Portner, 2003). This PERFECT is described as the core of the PERFECT. However, the indication of a current result of a past event can also be conveyed by a Past or Present tense (Van der Klis, Le Bruyn & De Swart, 2017), which highlights the problem to be tackled in the TinT-project.

The continuative PERFECT or universal PERFECT, as in (4), indicates the continuance of a past event into the present (Portner, 2003). For (4), it means that the event stretches from the time of the move till now. That is why Larsson (2009) proposes that the event time is an interval, rather than a point in time as Reichenbach (1947) proposed. Dutch frequently uses a Present tense or sometimes a Past tense to translate the continuative PERFECT (Van der Klis, Le Bruyn & De Swart, 2017). The *end of state dimension* of the Dutch PERFECT would imply that the state is over at the time of speech, contrary to English where the event stretches until point of speech. Usage of a PERFECT in (4d) would indicate that Donna does not live in Stockholm anymore.

- (2) a. Donna **has visited** Stockholm. (English)
b. Donna **har besökt** Stockholm. (Swedish)
c. Donna **hat** Stockholm **besucht**. (German)
d. Donna **heeft** Stockholm **bezocht**. (Dutch)
- (3) a. May **has moved** to Stockholm. (English)
b. May **har flyttat** till Stockholm. (Swedish)
c. May **ist** nach Stockholm **umgezogen**. (German)
d. May **is verhuisd** naar Stockholm. (Dutch)
- (4) a. Irene **has lived** in Stockholm for five years (now). (English)
b. Irene **har bott** i Stockholm i fem år (nu). (Swedish)
c. Irene **hat** schon fünf Jahre in Stockholm **gewohnt**. (German)
d. Irene **woont** al vijf jaar in Stockholm. (Dutch)

In a narrative context, the perspective is required to shift to the sequence of events, instead of staying at the time of speech (De Swart, 2007). This shift of perspective can occur through relations of the PERFECT with other times or events, as explained below. Thus, for English, the PERFECT with structure E-R,S is inappropriate to tell a story because the perspective R is not at E. The preferred tense in English for narrative is the Simple Past with structure E,R-S (De Swart, 2007).

In Dutch and German, the PERFECT is allowed to be used purely temporal, with past-time adverbials like *yesterday*, whereas English does not allow this preterite use of the PERFECT. Swedish resembles English, as it also does not have the preterite use (Rothstein, 2008). The relation of the PERFECT with other times is shown in (5) and (6), which are examples used in De Swart (2007) and Rothstein (2008), respectively. This phenomenon that some languages do allow the use of the PERFECT in

combination with adverbials expressing pastness, while other languages do not allow this, is called *the present perfect puzzle* by Klein (1992).

- (5) a. *Ada **has left** at six o'clock. (English)
 b. *Ada **har gått** klockan sex. (Swedish)
 c. Ada **ist** um sechs Uhr **abgefahren**. (German)
 d. Ada **is** om zes uur **vertrokken**. (Dutch)
- (6) a. *Marie **has come** yesterday. (English)
 b. *Marie **har kommit** igår. (Swedish)
 c. Marie **ist** gestern **angekommen**. (German)
 d. Marie **is** gister **gekomen**. (Dutch)

In German, the PERFECT can also be used to indicate relations with other events in narrative context (De Swart, 2007). That is, German, contrary to English, Swedish, and Dutch, can combine its PERFECT with other events introduced by *when*. For example, the event of Katherine seeing me, can be combined with the event of her getting frightened with use of the PERFECT. De Swart (2007) illustrates this as outlined in (7). In English, Swedish, and Dutch it is not possible to express this sentence with the PERFECT, not even if the second verb phrase would be a form of Simple Past.

- (7) a. *When Katherine **has seen** (PP) me, she **has got** (PP)/**got** (SP) frightened. (English)
 b. *När Katherine **har sett** (PP) mig, **har hon blivit** (PP) / **blev** hon (SP) rädd. (Swedish)
 c. Als Katherine mich **gesehen hat** (Perf), **hat** sie Angst **bekommen** (Perf). (German)
 d. *Toen Katherine me **heeft gezien** (VTT) **is** ze bang **geworden** (VTT)/ **werd** (OVT) zij bang. (Dutch)

Rothstein (2008) shows that in German and Swedish the PERFECT can be used as a Future PERFECT, while this is not possible in English, as can be seen in (8). I added the Dutch translation of this sentence, that shows Dutch also has the Future PERFECT.

- (8) a. *Tomorrow, the conference **has** already **ended**. (English)
 b. Imorgon **har** konferensen (redan) **slutat**. (Swedish)
 c. Morgen **hat** die Konferenz bereits **geendet**. (German)
 d. Morgen **is** de conference al **afgelopen**. (Dutch)
 'The conference will have ended by tomorrow.'

De Swart (2007) and Le Bruyn, Van der Klis, and De Swart (2019) state that English mostly uses the Simple Past in sentences where German and Dutch would prefer the PERFECT, as shown in the well-formedness of the use of the Simple Past in the translation in (9) of the sentences in (5), (6), and (7).

- (9) a. Ada **left** at six o'clock.
 b. Marie **came** yesterday.
 c. When Katherine **saw** (SP) me, she **got** (SP) frightened.

Likewise, Swedish also prefers the Imperfekt, the Swedish form of Past tense, to translate the sentences in (5), (6), and (7), as can be seen in (10).

- (10)a. Ada **gick** klockan sex.
 Ada went o'clock six
 b. Marie **kom** igår.
 Marie came yesterday
 c. När Katherine **såg** (Imperfekt) mig, **blev** (Imperfekt) hon rädd.
 When Katherine saw me, became she afraid

1.3 Hypothesis

Landman (2008) describes in his paper called '1066' how the auxiliary system of Old English resembled that of Dutch, while the modern auxiliary system that does not resemble Dutch was established by the time of Middle English. Landman argues that the only influence that could have caused this shift in English must have been the influence of the Vikings during their conquest of

England in the year 1066. This influence of North Germanic languages on English during the conquest of the Vikings is also described by Emonds and Faarlund (2014), who found around twenty grammatical changes from Old English to Middle English. They even argue that Old English did not get influenced by the language of the Nordic conquerors, but the language of the Nordic conquerors was influenced by Old English vocabulary and then adapted by every Englishman. Emonds and Faarlund (2014) argue that the grammatical system of Middle English developed from North Germanic with no changes.

This influence of Vikings and the Nordic languages they spoke on Old English, could have been the cause for a division between modern English and Swedish versus German and Dutch. For this research, it would imply that the use of the PERFECT in English and Swedish resembles each other.

Moreover, De Swart (2007) and Le Bruyn, Van der Klis, and De Swart (2019) state that out of English, Dutch, and German the use of the English PERFECT is the narrowest, since it does not allow relations with other events or times. The Dutch PERFECT does allow relations with other times, but not with other events. The German PERFECT is not subject to any constraints regarding connections with other events and states.

Swedish, German, and Dutch will use the PERFECT more than English, for example in sentences with the Future PERFECT. Moreover, German and Dutch have the preterite use of the PERFECT, contrary to Swedish and English.

Combining these previous researches, my hypothesis is that the use of the Swedish PERFECT will be broader than the use of the English PERFECT, but narrower than the use of the Dutch PERFECT. The use of the German PERFECT will be the broadest of the four Germanic languages discussed.

2. Method

To compare the use of verbs across languages, the TinT-project works with a parallel corpus approach. By comparing translations of a text, the meaning of it will remain the same, but the tense uses can differ between languages.

Capters from *Harry Potter and the Philosopher's Stone* written by J.K. Rowling in English, German, and Dutch were compared to their Swedish translation. We used *Harry Potter* rather than the other book investigated in the project, *L'Étranger* written by Albert Camus. While the latter uses a journal genre, *Harry Potter* has a more classic narrative structure (Le Bruyn, Van der Klis, and De Swart, 2019), which will thus give more insight in the PERFECT in narrative.

The sentences of the English chapter 1 and 17 had already been annotated and their verbs had already been marked. Our task was to annotate the Swedish sentences, so that we can align sentence-to-sentence translations. In addition, we assigned the tenses used in the Swedish verbs. With this information, the Swedish verbs could be mapped to the verbs from other languages.

2.1 Arrangement of the Swedish text

All pages of chapter 1 and 17 were scanned in a printer, with settings on maximum contrast and quality. The scans of chapter 1 were sent by the printer to my email address. The scans of chapter 17 were sent to Nikki Evers, with whom I worked together to include Swedish in the dataset. Using Adobe Acrobat Pro, the scans were edited so that the page numbers and stars on the page as decoration were cut out. Moreover, the image was corrected in order to present the texts in a horizontal manner. Lastly, each scan was cut in two, so each page was laid out separately.

The PDF-file containing the scans of the pages of chapter 1 was sent to Bert le Bruyn, who used the OCR-software ABBY FineReader to read the scans and convert it to text. He then sent me a Word file containing chapter 1 in text. This text was edited to align it with the chapters of the other languages.

All the diacritics had to be added by hand on the first page of the converted chapter 1. Moreover, whenever a hyphen was used to split a word in the original text, this hyphen needed to be deleted so that the word was one again.

After the text was corrected, the paragraphs could be aligned with the original text. That is, whenever there was a break in the original English text, there should be a break in the Swedish text as well. The Swedish text with aligned paragraphs was suitable for the TimeAlign software designed by Martijn van der Klis for the TinT-project.

2.2 TimeAlign

The TimeAlign software presented an English sentence from the *Harry Potter* chapters, with a marked verb. The same verb in the Swedish translation had to be marked. In addition, the check-boxes in the program needed to be marked correctly. The program asked if the translation of the original fragment is correct. This box was automatically checked and needed to be unchecked if the Swedish text was not a translation of the original English sentence. Moreover, the program asked if the selected words

The screenshot shows the TimeAlign interface. At the top, there is a navigation bar with links: Home, The project, Publications, Student Research, Meetings, Translation Mining, Contact, and a user profile 'Signed in as marthe'. The main heading is 'Annotation'. Below this, there are two columns. The left column is titled 'English (original)' and contains a text box with the sentence: 'Mrs Dursley was thin and blonde and had nearly twice the usual amount of neck , which came in very useful as she spent so much of her time craning over garden fences , spying on the neighbours .'. The right column is titled 'Swedish (translated)' and contains a text box with the Swedish translation: 'Mrs Dursley var smal och blond och hade nästan dubbelt så mycket hals som folk brukade , och den kom mycket väl till pass eftersom hon tillbringade så stor del av sin tid med att spana över trädgårdsstaketet och spionera på grannarna .'. Below the Swedish text, there are four checkboxes: 'The selected words in the original fragment do not form an instance of (a/an) simple past' (unchecked), 'This is a correct translation of the original fragment' (checked), 'The selected words in the original fragment are incorrectly marked as narration' (unchecked), and 'The translated fragment is not in the same structure (dialogue/narrative) as the original fragment' (unchecked). At the bottom, there is a 'Comments' section with a text input field.

Figure 1 Screenshot of the TimeAlign interface.

in the original fragment had their verb marked wrong, if the selected words in the original fragment were incorrectly marked as either narration or dialogue, and if the translated fragment was in a different structure (dialogue/narration) as the original fragment. These three boxes were automatically not checked and only needed to be checked if the statement in negation was true. A comment regarding the fragment could be made at the bottom. If everything was marked and checked correctly, it could be saved by clicking the blue Submit-button. Figure 1 shows a screenshot of the TimeAlign setup with the equivalent verb in Swedish marked.

2.3 Tense assignment

After every verb was marked, the verb tenses were to be assigned. Figure 2 shows an overview of the Excel-layout. In column B we added the Swedish verb tenses. The sheet also showed an overview of the Swedish verb phrase in column F, G, H, and I, the corresponding English verb phrase in column S and its tense in column T, and the English and Swedish sentences it appeared in, in column V and O respectively.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
	id	tense	otf	is c	is d	w1	w2	w3	w4	poi	poi	poi	poi	coi	full fragment	source i	source t	source s	source t	source t	soi	soi	ment
2	63922	imperfekt		yes	yes	overlevde				RG					Pojken som *overlevde	50214	1.xml	s1.1	Lived	simple past			The Boy Who *Lived*
3	64071	imperfekt		yes	yes	var				VB				Mr och mrs Dursley i ni	50563	1.xml	s2.1	were	simple past			Mr and Mrs Dursley , of	
4	63806	imperfekt		yes	yes	var				VB				Mr och mrs Dursley i ni	50564	1.xml	s2.1	were	simple past			Mr and Mrs Dursley , of	
5	63976	imperfekt		yes	yes	godtog				VB				De var de sista man kun	50279	1.xml	s2.2	didn't holc	simple past			They were the last peop	
6	63778	imperfekt		yes	yes	kunde tänka				VB	VB			De var de sista man *ku	50280	1.xml	s2.2	'd expect	future in the past			They were the last peop	
7	64052	imperfekt		yes	yes	var				VB				De *var* de sista man k	50281	1.xml	s2.2	were	simple past			They *were* the last pe	
8	63848	imperfekt		yes	yes	tillverkade				PC				Mr Dursley var chef för	50572	1.xml	s3.1	made	simple past			Mr Dursley was the dire	
9	63547	imperfekt		yes	yes	var				VB				Mr Dursley *var* chef fi	50573	1.xml	s3.1	was	simple past			Mr Dursley *was* the d	
10	63920	imperfekt		yes	yes	hade				VB				Han var en stor , fläskig	50353	1.xml	s3.2	did have	simple past			He was a big , beefy ma	
11	63908	imperfekt		yes	yes	var				VB				Han *var* en stor , fläsk	50354	1.xml	s3.2	was	simple past			He *was* a big , beefy m	
12	63532	imperfekt		yes	yes	tillbringade				VB				Mrs Dursley var smal oc	50472	1.xml	s3.3	spent	simple past			Mrs Dursley was thin ar	
13	63585	imperfekt		yes	yes	kom				VB				Mrs Dursley var smal oc	50473	1.xml	s3.3	came	simple past			Mrs Dursley was thin ar	
14	63571	imperfekt		yes	yes	hade				VB				Mrs Dursley var smal oc	50474	1.xml	s3.3	had	simple past			Mrs Dursley was thin ar	
15	63729	imperfekt		yes	yes	var				VB				Mrs Dursley *var* smal	50475	1.xml	s3.3	was	simple past			Mrs Dursley *was* thin	
16	63972	imperfekt		yes	yes	fanns				VB				Paret Dursley hade en li	50444	1.xml	s3.4	was	simple past			The Dursleys had a sma	
17	63769	imperfekt		yes	yes	hade				VB				Paret Dursley *hade* ei	50445	1.xml	s3.4	had	simple past			The Dursleys *had* a sn	
18	63792	Konditionalis	l	yes	yes	skul-le avslöja				NN	VB			Mr och mrs Dursley hac	50007	1.xml	s4.1	would disc	future in the past			The Dursleys had everyf	
19	63834	imperfekt		yes	yes	var				VB				Mr och mrs Dursley hac	50008	1.xml	s4.1	was	simple past			The Dursleys had everyf	
20	64060	imperfekt		yes	yes	hade				VB				Mr och mrs Dursley hac	50009	1.xml	s4.1	had	simple past			The Dursleys had everyf	

Figure 2 Screenshot of the Excel document used to assign the Swedish verb tenses.

A great deal of verb tenses was assigned with the built-in filter from Excel. Table 1 shows an overview of the filters used to automatically assign verb tenses. However, these rules did not mark every verb, so the remaining had to be assigned by hand. Several websites²³⁴ were used as information sources.

If Word 1...	If Word 2...	Verb tense =
Equals to "var"	Is empty	Imperfekt
Equals to "skulle"	Undefined	Konditionalis
Equals to "ska"	Is not empty	Futurum
Equals to "kommer"	Undefined	Futurum
Equals to "har"	Is not empty	Perfekt
Equals to "hade"	Is not empty	Pluskvamperfekt
Equals to "hade"	Is empty	Imperfekt
Ends with "r"	Undefined	Presens

Table 1 Filter rules used to assign part of the Swedish verbs automatically, note that they were admitted from top to bottom.

The verb tense that most appeared was Imperfect, followed by Presens. Other tenses that appeared, in order of number of appearances: Pluskvamperfekt, Perfekt, Imperativ, Konditionalis, Futurum, Supinum, and Infinitiv. We marked eleven verbs with a question mark, either because the Swedish translation of the English verb was not really a verb, but more an expression. For example, the English sentence '... which **were** the eyes of the cat watching him' translates to Swedish as '... **nämligen** ögonen på katten som betraktade honom', which uses the adverb 'namely' rather than a verb. Moreover, a few verbs were not assigned with the filters and could not be found by hand in the websites, so we left them question marked as well.

2.4 Mapping the tenses

The Excel-sheet with the tenses assigned to the Swedish verbs, was sent to Martijn van der Klis who performed the so-called 'TimeMapping'. This uses the output of the Excel document, to make a table

² <https://www.lysator.liu.se/language/Languages/Swedish/Grammar.html#verbs>

³ <https://werkwoorden.woxikon.nl/sv>

⁴ <https://nl.bab.la/werkwoorden/zweeds/>

in which every Swedish sentence is matched with the sentences and verb tenses used in the other languages investigated in this research. This leads to a four-tuple with aligned tense attributions in Swedish, English, German, and Dutch. An example of a tuple where all languages use their form of PERFECT is: <Perfekt, Present Perfect, vtt, Perfekt>. The order of the tuples is German, English, Dutch, Swedish.

The tuples are visually represented via multidimensional scaling in interactive ‘Semantic Maps’ by a distance function that calculates the distance between pairs of the tuples based on their variation in tense use. A distance function defines four-tuples to be similar if all the tenses match. If this is not the case, one is added for each mismatch. The total sum is divided by four (Van der Klis, Le Bruyn & De Swart, 2017).

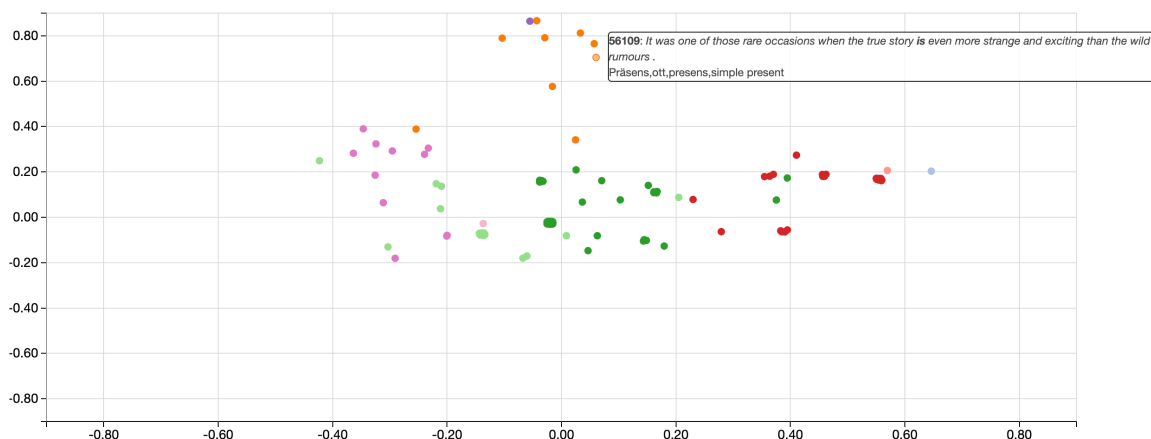
An example of a Semantic Map of English narrative can be found in Figure 3. The points are labelled by colour with the verb tenses of the selected language, English in this case. Single-clicking on a tense at the top will remove this tense from the visualisation. Double-clicking on a tense will remove all other tenses and only show this tense in the visualisation. It is also possible to show a combination of tenses in one language. Hovering the mouse over a point allows to inspect the English sentence with the concerning verb phrase in bold, together with the corresponding four-tuple. Clicking on a point will show all the verbs that have the same tuple. Inspecting one will give a fragment overview with all the translated sentences, their verbs highlighted and assigned tenses. Note that some points are so close together that they overlap each other.

The Semantic Maps are useful, because they show how tenses interact within a language. All the Semantic Maps are the same, the colours of the points differ per language, depending on the tense of the language inspected. This allow for comparison between languages, because the colour labelling per tense remains constant between languages (Van der Klis, Le Bruyn & De Swart, 2017). The dimensions used for the best overview are 1 on the x-axis and 2 on the y-axis. The generated Semantic Maps are discussed in the section Results.

MDS visualization (scenario *Harry Potter - Swedish - narration - Germanic*)

Stress: 55.411

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



Filters

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

Figure 3 Example of a Semantic Map.

3. Results

As described in the previous section, the tuples and their visualisation in the Semantic Maps will give insight in the use of the PERFECT throughout different languages. Each point on the Semantic Map represents a specific context, a four-tuple. The distance between two points is determined by variation in verb tenses used throughout languages. Thus, points closer together indicate little variation in verb tense used across the four languages. All the Semantic Maps for narrative and dialogue are the same, only the colours of the points differ based on the tense used in the language inspected. Clusters of points in the same colour indicate pattern in usage, and are circled in the corresponding colour of the verb tense. A large circle does not necessarily indicate that there are more occurrences of that tense. Since some points overlap because they are so close together, the dispersion of a tense can be limited.

The results are separately discussed in a narrative and a dialogue part. The dialogue part consisted of all the sentences between quotation marks in chapter 1 and 17 of *Harry Potter and the Philosopher's Stone*, the narrative part consisted the sentences outside the quotation marks. For the dialogue part, a subset with only indicative tenses, so without futures, was made. Including the future tenses would results in much greater variation, which would have made the Semantic Maps harder to read. Thus, to focus on PERFECT domain, tuples with future tenses were left out. In total, the narrative part contained 655 four-tuples, the dialogue part contained 388 four-tuples.

3.1 Narrative

Figures 4, 5, 6, and 7 show the Semantic Maps of the narrative text of English, Swedish, German, and Dutch, respectively.

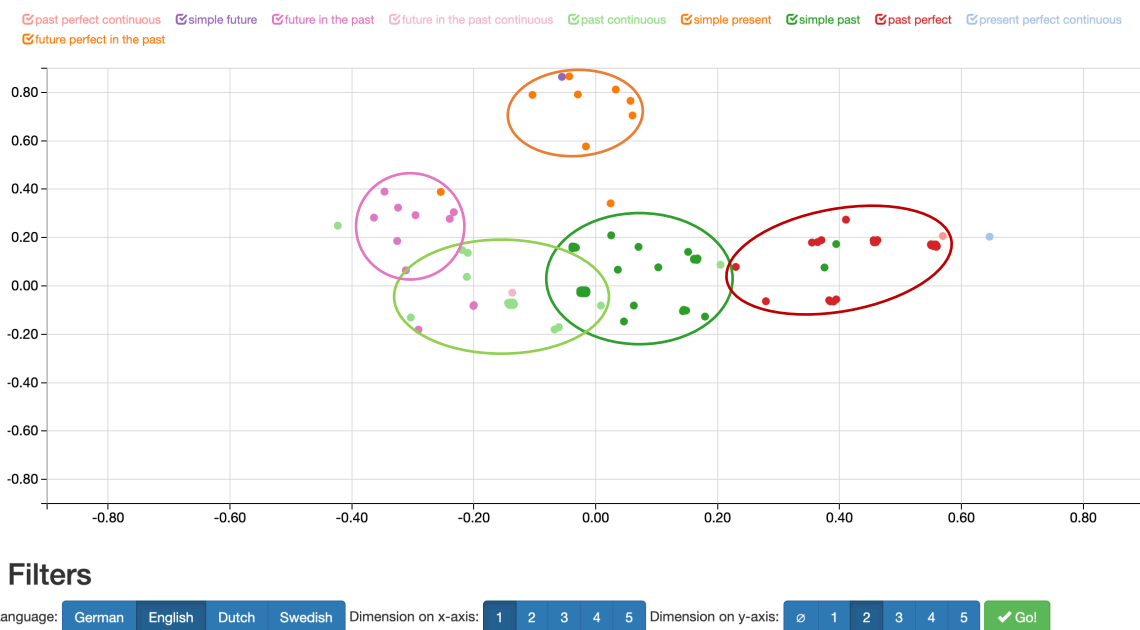
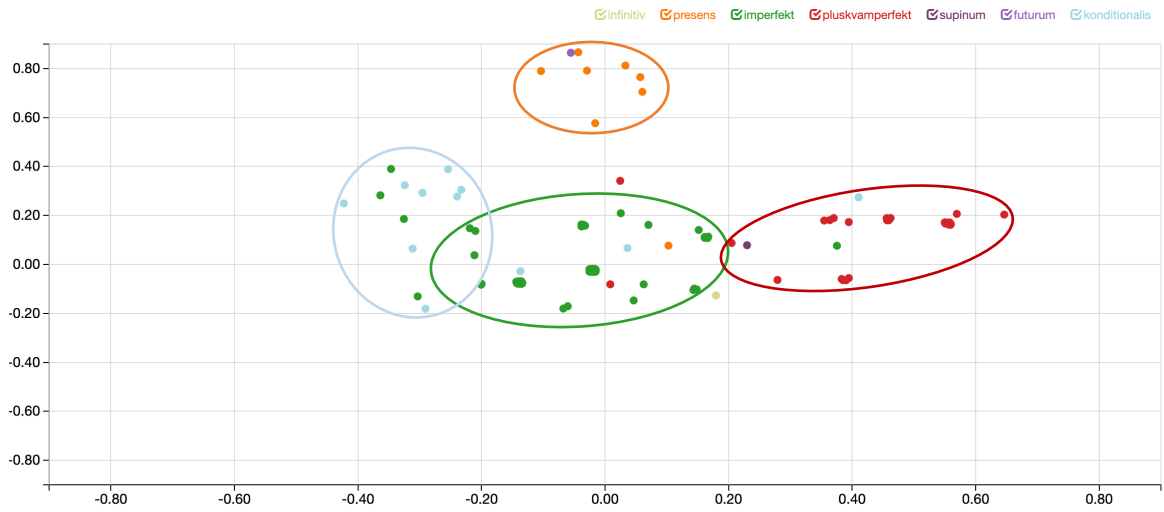


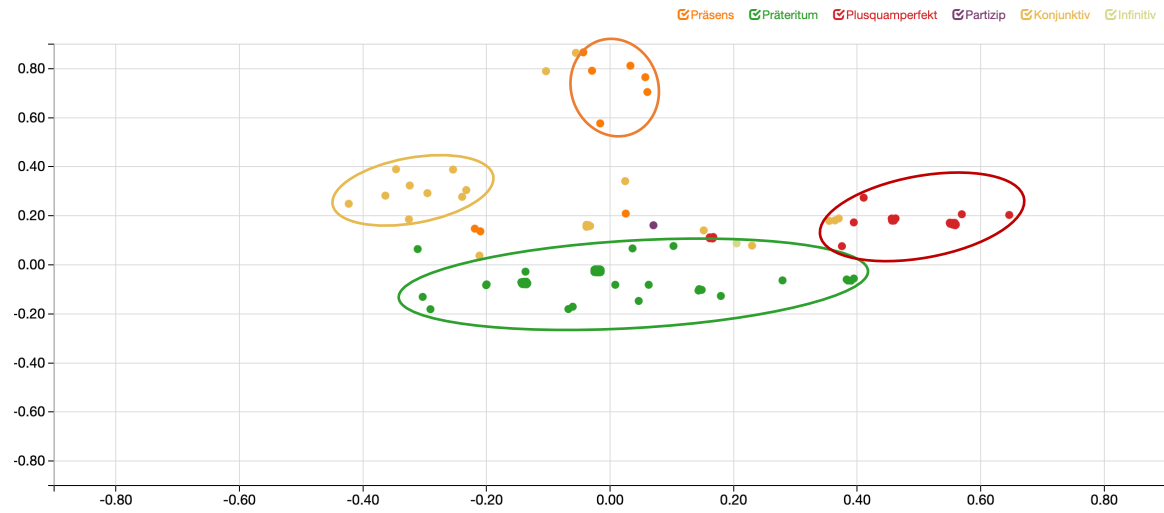
Figure 4 Semantic Map of English narrative.



Filters

Language: German English Dutch Swedish Dimension on x-axis: 1 2 3 4 5 Dimension on y-axis: ∅ 1 2 3 4 5 Go!

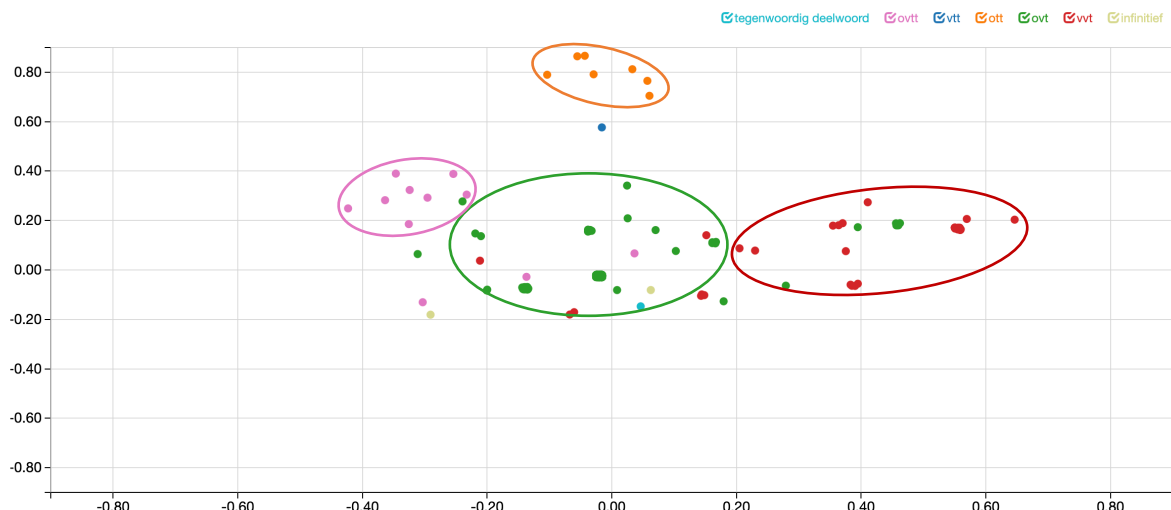
Figure 5 Semantic Map of Swedish narrative.



Filters

Language: German English Dutch Swedish Dimension on x-axis: 1 2 3 4 5 Dimension on y-axis: ∅ 1 2 3 4 5 Go!

Figure 6 Semantic Map of German narrative.



Filters

Language: German English Dutch Swedish Dimension on x-axis: 1 2 3 4 5 Dimension on y-axis: ∅ 1 2 3 4 5 Go!

Figure 7 Semantic Map of Dutch narrative.

Table 2 shows an overview of each Swedish tense in a contingency table matched with the English tenses used in the narrative part. The left column displays all the English tenses that occurred in the narrative part of the text. The top row displays all the Swedish tenses that occurred in the narrative part of the dataset. Only the tenses that are used in the narrative part of the text are displayed here. The numbers in the table represent how often the English verb tense is translated to the Swedish verb tense. For example, the English Simple Present is translated to the Swedish Presens seven times. In other words, Simple Present and Presens are in seven tuples together. The right column displays how often each English tense appeared in this dataset. For example, the English Simple Present is used a total of seven times in the narrative part of the text. The bottom row displays how often each Swedish tense is used. Note that the numbers in the middle, the right column, and the bottom row all add up to 655.

In table 3 the Swedish tenses used in the narrative part are compared in a contingency table to the German tenses used in the narrative part. In table 4 the Swedish tenses used in the narrative part are compared in a contingency table to the Dutch tenses used in the narrative part.

Swedish English	Presens	Imperfekt	Infinitiv	Supinum	Futurum	Pluskvamperfekt	Konditionalis	Total per English tense:
Simple Present	7							7
Simple Past	1	541	1			1	1	545
Past Perfect				1		34	1	36
Past Continuous		47				2	1	50
Present Perfect Continuous						1		1
Past Perfect Continuous						1		1
Simple Future					1			1
Future in the Past		5					6	11
Future in the Past Continuous							1	1
Future Perfect in the Past						1	1	2
Total per Swedish tense:	8	593	1	1	1	40	11	

Table 2 Contingency table of English and Swedish in narrative.

Swedish German	Presens	Imperfekt	Infinitiv	Supinum	Futurum	Pluskvamperfekt	Konditionalis	Total per German Tense:
Präsens	6	3						9
Präteritum	1	573	1			6	4	585
Konjunktiv	1	11		1	1	4	6	24
Plusquamperfekt		5				29	1	35
Infinitiv						1		1
Partizip		1						1
Total per Swedish tense:	8	593	1	1	1	40	11	

Table 3 Contingency table of German and Swedish in narrative.

Swedish Dutch	Presens	Imperfekt	Infinitiv	Supinum	Futurum	Pluskvamperfekt	Konditionalis	Total per Dutch tense:
onvoltooid tegenwoordige tijd (ott)	6				1			7
onvoltooid verleden tijd (ovt)	1	579	1			9	2	592
voltooid tegenwoordige tijd (vtt)	1							10
voltooid verleden tijd (vvt)		8		1		31	1	41
onvoltooid verleden toekomstige tijd (ovtt)		4					7	11
infinities		1					1	2
tegenwoordig deelwoord		1						1
Total per Swedish tense:	8	593	1	1	1	40	11	

Table 4 Contingency table of Dutch and Swedish in narrative.

Table 5 shows an overview of the six tuples that most appeared in the narrative part of this data set, and how often they appeared. For example, the tuple <Präteritum, Simple Past, ovt, Imperfekt > was formed 522 times. Thus, in 522 sentences, German used the Präteritum, English the Simple Past, Dutch onvoltooid verleden tijd, and Swedish the Imperfekt. Two tuples appeared 5 times each, sharing a place in the table.

Tuple	Frequency in narrative
<Präteritum, Simple Past, ovt, Imperfekt >	522
< Präteritum, Past Continuous, ovt, Imperfekt >	41
< Plusquamperfekt, Past Perfect, vvt, Pluskvamperfekt >	21
< Konjunktiv, Simple Past, ovt, Imperfekt >	6
< Konjunktiv, Simple Past, ovt, Imperfekt > < Plusquamperfekt, Past Perfect, ovt, Pluskvamperfekt >	5

Table 5 Six most frequent tuples in the narrative part.

3.2 Dialogue

Figures 8, 9, 10, and 11 show the Semantic Maps of the dialogue text of English, Swedish, German, and Dutch, respectively. Like in the narrative Semantic Maps, clusters of points from some tenses are circled in the corresponding colour of that tense.

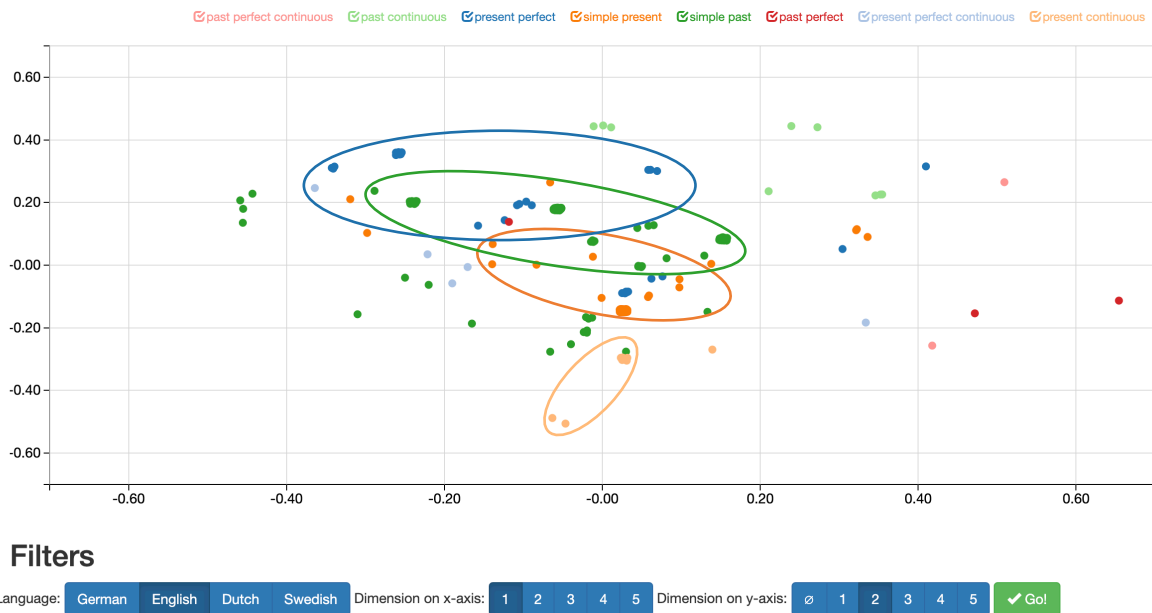


Figure 8 Semantic Map of English dialogue.

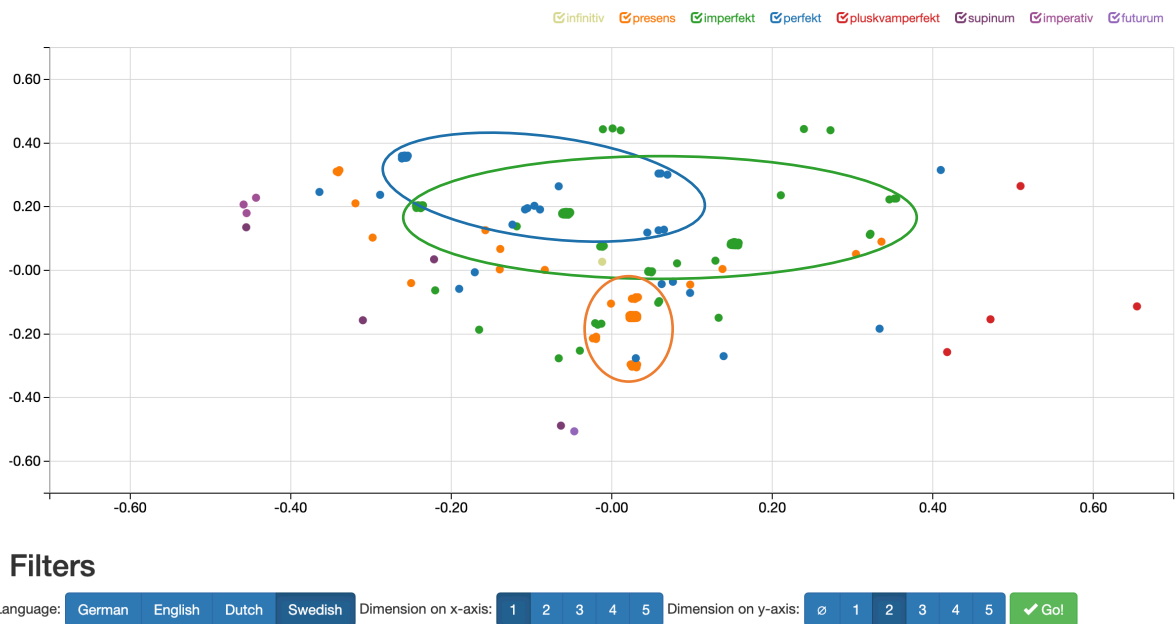
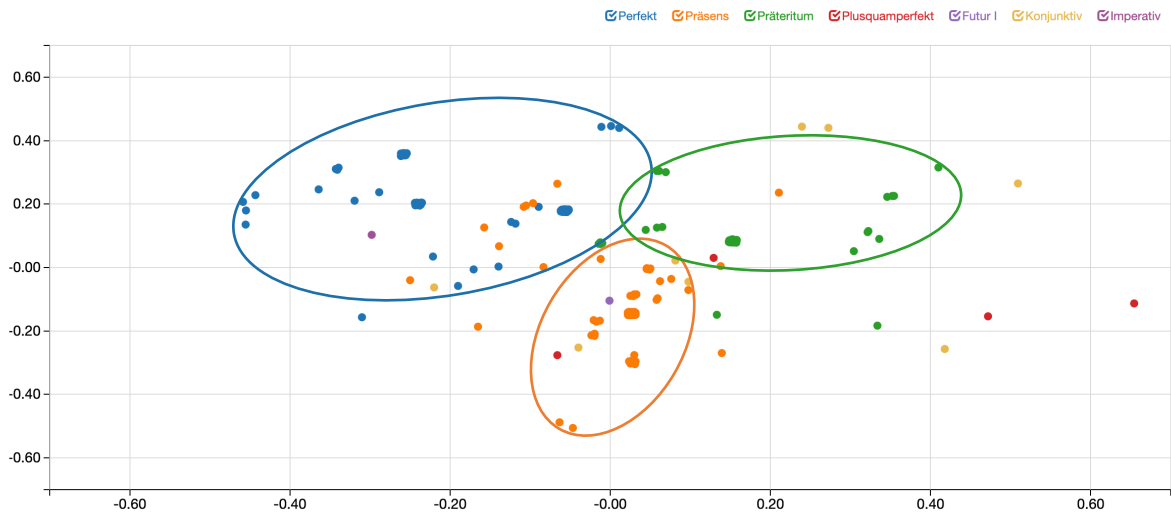


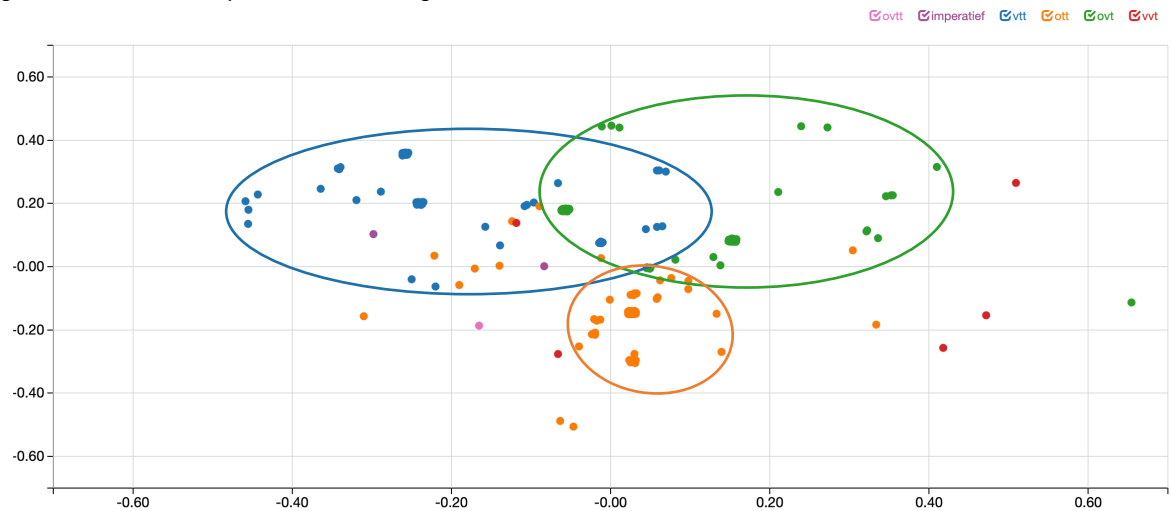
Figure 9 Semantic Map of Swedish dialogue.



Filters

Language: German English Dutch Swedish Dimension on x-axis: 1 2 3 4 5 Dimension on y-axis: ∅ 1 2 3 4 5 Go!

Figure 10 Semantic Map of German dialogue.



Filters

Language: German English Dutch Swedish Dimension on x-axis: 1 2 3 4 5 Dimension on y-axis: ∅ 1 2 3 4 5 Go!

Figure 11 Semantic Map of Dutch dialogue.

Tables 6, 7, and 8 show contingency tables of verb tenses used in the dialogue part of Swedish compared to the verb tenses used in English, German, and Dutch, respectively.

Swedish English	Presens	Imperfekt	Perfekt	Infinitiv	Imperativ	Supinum	Futurum	Pluskvam perfekt	Total per English tense:
Simple Present	158	4	2	1					165
Simple Past	5	135	5		3	2			150
Present Perfect	10		27						37
Present Continuous	14		1			1	1		17
Past Perfect		1						2	3
Past Continuous		9							9
Present Perfect Continuous			4			1			5
Past Perfect Continuous								2	2
Total per Swedish tense:	187	149	39	1	3	4	1	4	

Table 6 Contingency table of English and Swedish in dialogue.

Swedish German	Presens	Imperfekt	Perfekt	Infinitiv	Imperativ	Supinum	Futurum	Pluskvam perfekt	Total per German Tense:
Präsens	177	12	9	1		1	1		201
Perfekt	5	64	22		3	3			97
Präteritum	2	66	8						76
Konjunctiv	1	5						2	8
Plusquamperfekt		2						2	4
Imperativ	1								1
Futur I	1								1
Total per Swedish tense:	187	149	39	1	3	4	1	4	

Table 7 Contingency table of German and Swedish in dialogue.

Swedish Dutch	Presens	Imperfekt	Perfekt	Infinitiv	Imperativ	Supinum	Futurum	Pluskvam perfekt	Total per Dutch tense:
onvoltooid tegenwoordig e tijd (ott)	176	7	10			3	1		198
onvoltooid verleden tijd (ovt)	2	102	1					1	106
voltooid tegenwoordig e tijd (vtt)	7	37	28	1	3	1			76
voltooid verleden tijd (vvt)		2						3	2
imperatief	2								2
onvoltooid verleden toekomende tijd (ovtt)		1							1
Total per Swedish tense:	187	149	39	1	3	4	1	4	

Table 8 Contingency table of Dutch and Swedish in dialogue.

Table 9 shows an overview of the six tuples that most appeared in the dialogue part of this data set, and how often they appeared. There are 16 sentences in which English, Swedish, German, and Dutch all use their form of the PERFECT.

Tuple	Frequency in Dialogue
< Präsens, Simple Present, ott, Presens >	149
< Präteritum, Simple Past, ovt, Imperfekt >	56
< Perfekt, Simple Past, vtt, Imperfekt >	32
< Perfekt, Simple Past, ovt, Imperfekt >	28
< Perfekt, Present Perfect, vtt, Perfekt >	16
< Präsens, Present Continuous, ott, Presens >	14

Table 9 Six most frequent tuples in the dialogue part.

4. Discussion

In this section, the most interesting results described in the previous section will be highlighted and further elaborated. The results of the narrative are first discussed, followed by the results of the dialogue. Moreover, I will discuss some possible implementation rules that could help improve Natural Language Processing. Lastly, I will discuss some points of improvement that could improve further research.

4.1 Narrative

Like in English and German, Swedish also did not use the PERFECT in the narrative part of the text. Dutch used its PERFECT the *vtt* once, which translated to a form of Present tense in the other three languages, as can be seen in (14). However, I argue that also in Dutch a Present tense is used with the Present tense of verb “to be”, and the state of being hidden. Moreover, this sentence is part of the thoughts of the protagonists where he speaks to himself, described in first perspective. Thus, it resembles dialogue more, although it is outside the quotation marks.

- (11)a. – which means I 'll see where it's **hidden!** (English)
- b. – vilket betyder att jag kommer att se var den **är gömd!** (Swedish)
- c. Und das heißt, ich wüsste, wo er **versteckt ist!** (German)
- d. Betekent dat dan dat ik zie waar hij **verborgen is?** (Dutch)

All four language use a form of Past Perfect, and in 21 sentences the four Germanic language chose their form of Past Perfect simultaneously. Two of these sentences can be found in (12) and (13).

- (12)a. He'd never even **seen** the boy. (English)
 - b. Han **hade** aldrig ens **sett** pojken. (Swedish)
 - c. Er **hatte** den Jungen noch nicht einmal **gesehen**. (German)
 - d. Hij **had** die blaag per slot van rekening nooit **gezien**. (Dutch)
- (13)a. Most of them **had** never **seen** an owl even at nighttime. (English)
 - b. De flesta av dem **hade** aldrig **sett** en uggla ens nattetid. (Swedish)
 - c. Die meisten von ihnen **hatten** überhaupt noch nie eine **gesehen**, nicht einmal nachts. (German)
 - d. De meesten **hadden** zelfs 's nachts nog nooit een uil **gezien**. (Dutch)

German selects Präteritum while English, Swedish and Dutch use their Past Perfect in four sentences, of which two have the English verb phrase ‘had had’. Once can be found in (14).

- (14) a. Mrs Dursley **had had** a nice, normal day. (English)
- b. Mrs Dursley **hade haft** en trevlig, vanlig dag. (Swedish)
- c. Mrs Dursley **hatte** einen netten, gewöhnlichen Tag hinter sich. (German)
- d. Mevrouw Duffeling **had** een leuke, doodgewone dag **gehad**. (Dutch)

4.2 Dialogue

Out of the four Germanic languages studies here, the PERFECT is most used in German, followed by Dutch. In 32 tuples, German uses the Perfekt, while the other three languages use their Past tense. In 28 tuples, German and Dutch use Perfekt and *vvt*, while English and Swedish use the Simple Past and Imperfekt. The English Present Perfect (the PERFECT) is used almost as often as the Swedish Perfekt, although this is no indication that English and Swedish use the PERFECT at the same time.

In five sentences, the Swedish Perfekt translates to the English Simple Past. Two times the sentence contains the adverb ‘never’, and once the adverb ‘always’. (15) and (16) show the sentences with the adverb ‘never’. English and German use their form of Past tense, while Swedish and Dutch selects their PERFECT. (17) shows the sentences with the adverb ‘always’, where Swedish, German and Dutch select their PERFECT, while English selects the Simple Past.

- (15)a. “He never **had** much sense.” (English)
- b. “Han **har** aldrig **haft** nåt vidare förstånd.” (Swedish)
- c. “Der **war** noch nie besonders vernünftig.” (German)
- d. “Die **heeft** nooit veel hersens **gehad**.” (Dutch)

- (16) a. "But he never **wanted** you dead." (English)
 b. "Men han **har** aldrig **velat** att du ska dö." (Swedish)
 c. "Aber er **wollte** nie, dass Sie sterben." (German)
 d. "Maar hij **heeft** nooit je dood **gewild**." (Dutch)
- (17) a. "I always **said** he was off his rocker" (English)
 b. "Jag **har** alltid **sagt** att han är vrickad" (Swedish)
 c. "Ich **hab** ja immer **gesagt**, dass er völlig von der Rolle ist" (German)
 d. "Ik **heb** altijd al **gezegd** dat hij niet goed snik is" (Dutch)

However, in this data set it was not always the case that English does not use a PERFECT combined with the adverbs 'always' or 'never'. There are three sentences in this dataset where English uses a PERFECT with 'never' (twice) or 'always' (once). (18) shows a sentence where all four languages use their PERFECT, in combination with the adverb 'never'.

- (18) 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. "Mein lieber Professor, ich **habe** noch nie eine Katze so steif dasitzen **sehen**." (German)
 d. "M'n beste professor, ik **heb** nog nooit een kat zo stijfjes **zien** zitten." (Dutch)

English uses the Present Perfect Continuous five times in this data set, all occurrences are in the dialogue part. The Present Perfect Continuous formed by the PERFECT of the verb 'to be' (have/has been), and the present participle of the main verb (root + ing). Three out of these five sentences, the sentence contains the adverb 'today' or 'all day'. Four times, this Present Perfect Continuous translates to the Perfekt in Swedish. One sentence is displayed in (19), where German uses Perfekt and Dutch ott.

- (19) a. "I've **been watching** them all day." (English)
 b. "Jag **har iakttagit** dem hela dan." (Swedish)
 c. "Ich **habe** sie den ganzen Tag **beobachtet**." (German)
 d. "Ik **houd** dat huis al de hele dag **in de gaten**." (Dutch)

In six sentences, the English PERFECT translates to the Present tense in the other three languages. Here, English uses a form of 'have got' which translates to the Present tense of 'har' (Swedish), 'haben' (German), or 'hebben' (Dutch). Two of these sentences can be found in (20) and (21).

- (20) a. "Have a Chocolate Frog, I've **got** loads..." (English)
 b. "Ta en chokladgroda, jag **har** massor..." (Swedish)
 c. "Nimm einen Schokofrosch, ich **hab** ganze Wagenladungen davon..." (German)
 d. "Vooruit, neem een Chocokikker, ik **heb** dozen vol..." (Dutch)
- (21) a. "He's **got** the Stone!" (English)
 b. "Han **har** Stenen!" (Swedish)
 c. "Er **hat** den Stein!" (German)
 d. "Hij **heeft** de Steen!" (Dutch)

In three other sentences, English, German, and Dutch select their PERFECT, while Swedish selects the Present tense. All three sentences are a translation of 'is gone', which translates to German as 'ist verschwunden' or 'is gegangen' and to Dutch as 'is verdwenen'. In Swedish, the phrase 'är borta' is used, where 'är' is the Present form of the verb 'vara' which is to be, and 'borta' is an adverb meaning away. One sentence can be found in (22). Moreover, in (23), a sentence with the PERFECT form of 'have / is gone' in English and Dutch, that translates to a Present tense in German and Swedish.

- (22) a. "Rejoice, for You-Know-Who **has gone** at last!" (English)
 b. "Jubla och var glad, för Ni-Vet-Vem **är** äntligen borta!" (Swedish)
 c. "Freuen wir uns, denn Du-weißt-schon-wer **ist** endlich von uns **gegangen**!" (German)
 d. "Wees blij, want Jeweetwel **is verdwenen**!" (Dutch)
- (23) a. "As I say, even if You-Know-Who **has gone** –" (English)

- b. "Som jag sa, även om Ni-Vet-Vem **är** borta..." (Swedish)
- c. "Wie ich schon sagte, selbst wenn Du-weißt-schon- wer wirklich fort **ist** –" (German)
- d. "Zoals ik al zei, zelfs als Jeweetwel inderdaad **is verdwenen** –" (Dutch)

A tuple is left out of the research if there is no translation of the English verb to one of the other languages in the tuple, to prevent a negative effect on the analysis (Van der Klis, Le Bruyn & De Swart, 2017). The more languages are simultaneously discussed, the more chance there is no verb translation exists to at least one of the languages. Therefore, I also compared scenarios of just English and Swedish to each other. This set of dialogue verbs consists of 561 tuples, which is 173 more than the four-tuple. In this set, the PERFECT appears 43 times in English and 45 times in Swedish. They occur simultaneously 31 times. Moreover, the Present Perfect Continuous translates 5 times to the Perfekt, which is once more than in the scenario with the four Germanic languages. However, it seems like Dutch was wrongly not labelled with a tense, because it contains a form of vtt, as can be seen in (24).

- (24) a. "Perhaps people **have been celebrating** Bonfire Night early –" (English)
- b. "Folk kanske **har** firat valborgsmässoafton i förväg –" (Swedish)
- c. "Vielleicht **haben** die Leute zu früh Silvester **gefeiert** –" (German)
- d. "Misschien **hebben** sommige mensen hun vuurwerk wat te vroeg **afgestoken** –" (Dutch)

4.3 Implications for Artificial Intelligence

The results of this research can contribute to the improvement of Natural Language Processing (NLP). I argued that while waiting for the development and affordability of self-learning Neural Machine Translation models, specific translation rules about the PERFECT could improve NLP on short-term. A rule that could be implemented in translation machines is that if English uses the Present Perfect Continuous, the Swedish and German translation most likely will use the PERFECT and the Dutch translation will probably use ott. Moreover, another rule that could be implemented is that whenever English uses a PERFECT with 'have got', it translates to a Present tense in Swedish, German, and Dutch. Furthermore, a translation rule could be that whenever English, German, or Dutch use a PERFECT with 'gone', Swedish will use the Present form of 'to be' with the adverb 'borta'.

These preliminary rules are just the first outlines. For actual implementation, more research in a larger dataset needs to be done, in order to see if the abstracted rules will always apply.

4.4 Points of improvement

A fundamental idea of the TinT-project is the correctness of the *Harry Potter* translations. I want to reflect on this, since I want to investigate if we can consider these translations unquestionably correct. I doubt there is one perfect translation of a whole text. Different translators might make slightly distinct translation choices, while the conveyed message of the whole will stay the same. Van der Klis, Le Bruyn and De Swart (2017) argue that the indication of a current result of a past event can be conveyed by either a Past tense, a Present tense, or the PERFECT tense. For example, this sentence and its translations in (25) that uses a form of PERFECT in English, Swedish, and German. In Dutch the Simple Present tense ott is chosen.

- (25) a. '... even the Muggles **have noticed** something 's going on.' (English)
- b. '... till och med mugglarna **har märkt** att nånting är på gang.' (Swedish)
- c. '... selbst die Muggel **haben bemerkt**, dass etwas los ist.' (German)
- d. '... zelfs de Druzels **merken** dat er iets aan de hand is.' (Dutch)

I, as a mother tongue speaker of Dutch, would also approve of the use of a PERFECT in this Dutch sentence as follows: '... zelfs de Druzels **hebben gemerkt** dat er iets aan de hand is.' For me, these two sentences, with Simple Past or PERFECT tense, convey the same meaning.

It might be interesting to explore these different translation possibilities in further research, especially if this can also occur in Swedish. New research can ask native speakers of Dutch, German or Swedish how they would translate English sentences that might have ambiguous translations. This would give a more accurate representation of the language since more opinions on preference are weighted.

Furthermore, although the sentence in (25) shows the opposite, I consider translators to have a slight preference for translations that resemble the original fragment the most. If I follow this thought,

translators might have a preference to use the same verb tense used in the original fragment, if they have a choice when there are multiple correct options in the translated languages. If this is true, it would suggest that since the original chapters 1 and 17 of *Harry Potter* novel do not contain many forms of the PERFECT, the translations use fewer forms of PERFECT than one would expect based on normal verb tenses used in the language or in comparison with a book in the same genre written by a native speaker.

Thus, new research can make more use of native speakers of each language investigated. However, if it is true that translators prefer a translation that resembles the original fragment the most in terms of grammatical order or verb tense, comparing translations of text will never give an optimal view of verb tenses used in each language.

The Tint-project uses morphological taggers for English tense assignment. Because of that, sometimes a verb was wrongly tagged. Moreover, all the verb tenses assigned by hand could also contain mistakes. Some mistakes were noticed during the processing of the results. However, some mistakes may not have been noticed, and thus might have influenced the data. For example, in (25) Dutch was wrongly labelled. Moreover, (25) shows a sentence that uses the PERFECT in English and the Present tense in the other three languages, yet German was labelled as Präteritum rather than Präsens.

- | | |
|---|-----------|
| (26) a. "Hurry up, boy, we haven't got all day." | (English) |
| b. "Skynda dig på, pojk, vi har inte hela dan på oss." | (Swedish) |
| c. "Beeil dich, Junge, wir haben nicht den ganzen Tag Zeit." | (German) |
| d. "Haast je, jongen. We hebben niet de hele dag de tijd." | (Dutch) |

In further research, more advance morphological taggers could be used. Moreover, all the marked and assigned verb tenses need to be double-checked, preferably by a native speaker of the languages labelled.

Beforehand, I did not have any familiarity with Swedish. During the course of the research, I learned some words and I found it easier and easier to read the Swedish sentences as a whole, because I could trace back most Swedish words to either English or Dutch, languages I am familiar with. Yet, unfortunately, I still experienced some disadvantages of not being familiar with Swedish. I was not capable of judging the grammatical correctness of the Swedish sentences, or could think about example sentences to support my argument. This is the same for German, since I only have basic knowledge about those languages. Even for some English sentences my grammatical judgement was incorrect. I solved this proficiency problem by using the example sentences provided in the academic literature I referred to, as I could rely on them being grammatical correct. Moreover, whenever I wanted to add new example sentences, for instance when I wanted to add a language to sentences I extracted from other articles, I asked mother tongue speakers for the correct translation. I want to specially thank Dr. Josefin Lindgren, a native speaker of Swedish and researcher in the field of linguistics, for her input. However, not all native speakers I contacted had a background in linguistics or formal translation, so they may have misinterpreted the sentences. In further research, it would be ideal to have mother tongue speakers of all languages involved in the project, to prevent grammatically incorrect translations.

5. Conclusion

This research covered the inclusion of Swedish to the Time in Translation project on the semantics of the Present Perfect (PERFECT). Adding Swedish was relevant to gain more insight in the use of the PERFECT throughout. This knowledge can, among other things, contribute to development of automated machine translation. Based on previous literature about the English, Swedish, German, and Dutch PERFECT, the hypothesis was that the use of the Swedish PERFECT was broader than the use of the PERFECT in English, but narrower than the use of the PERFECT in German and Dutch.

On the basis of two chapters from the novel *Harry Potter* and their translations, verb tenses used across languages could be compared. Using translated text ensures the conveyed meaning of the sentences remained the same across languages, and thus gave specific insight in similarities and differences in used verb tenses between languages. Each tense used in English, Swedish, German, and Dutch formed a four-tuple with aligned tense attributions. With a distance function, this four-tuple could be plotted into an interactive visualisation called a Semantic Map. The tuples and their corresponding points on the Semantic Map constituted the results of this research. Notable results were that the PERFECT only occurred in the dialogue text and more in German and Dutch than in English and Swedish. In English and Swedish, the PERFECT occurred about as often, but not always in the same sentences. The English PERFECT and the Swedish Present tense were selected alongside the English verb phrase 'have got' or 'is gone'. The former translates to a Present tense in German and Dutch as well, while in the latter all but Swedish use their PERFECT. Sentences with the adverbs 'never' or 'always' used a PERFECT in Swedish and Dutch, in German either a PERFECT or Past tense, and in English half of the sentences the Simple Past and the other half the PERFECT. In 27 tuples in this research, English and Swedish select the PERFECT in the same sentences, this is two-thirds of the times English and Swedish select a PERFECT. In a little more than half of these 27 sentences, all four languages investigated use the PERFECT.

This research showed that German had the broadest use of the PERFECT. This is followed by Dutch, where the PERFECT is used less but still more than in English and Swedish. This is in line with the hypothesis that German and Dutch have a broader distribution of the PERFECT than English and Swedish. English and Swedish use the PERFECT about as often, although not always in the same sentences. This is not in line with the hypothesis that Swedish would use the PERFECT more than English. This shift between a broad and narrow use of the PERFECT could be a consequence of the Vikings influencing English language. These results could contribute to the improvement of Natural Language Processing. I proposed some rules in preliminary form that could eventually be implemented as translation rules. However, first, more research needs to be done with a larger dataset and with more native speakers of each language investigated involved for optimal insight.

- (27) a. Thank you for reading, I **enjoyed** writing it.
b. Tack för att du läste, jag **tyckte** om att skriva det.
c. Danke für das Lesen, ich **habe es genossen** ihn zu schreiben.
d. Bedankt voor het lezen, ik **heb genoten** van het schrijven.

6. References

- De Swart, H. (2007). A cross-linguistic discourse analysis of the Perfect. *Journal of Pragmatics*, 39(12), 2273–2307. <https://doi.org/10.1016/j.pragma.2006.11.006>
- Emonds, J. E., & Faarlund, J. T. (2014). *English: the Language of the Vikings* (1st ed.). Olomouc: Palacký University.
- Jurafsky, D., & Martin, J. H. (2009). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (Second Edition). Upper Saddle River, New Jersey, US: Prentice Hall.
- Klein, W. (1992). The present perfect puzzle. *Language*, 525-552.
- Landman, F. 2008. 1066. On the differences between the tense-perspective-aspect systems of English and Dutch. In S. Rothstein (ed.), *Theoretical and Crosslinguistic Approaches to the Semantics of Aspect*. Amsterdam: John Benjamins.
- Larsson, I. (2009). *Participles in time: the development of the perfect tense in Swedish ; (Svensk sammanfattning)*. Göteborg: Acta Universitatis Gothoburgensis.
- Le Bruyn, B., Van der Klis, M., De Swart, H. (2019, in press). *The Perfect in dialogue: evidence from Dutch*
- Reichenbach, H. (1947). The Tenses of Verbs. In *Elements of Symbolic Logic*. New York: The Macmillan Company.
- Rothstein, B. (2008). *The Perfect Time Span: On the present perfect in German, Swedish and English*. Universität Stuttgart, Germany: John Benjamins Publishing Company.
- Saxton, M. (2010). *Child Language: Acquisition and Development*. London, United Kingdom: SAGE Publications.
- Portner, P. (2003). THE (TEMPORAL) SEMANTICS AND (MODAL) PRAGMATICS OF THE PERFECT. *Linguistics and Philosophy*.
- Van der Klis, M., Le Bruyn, B., De Swart, H. (2017). Mapping the Perfect via Translation Mining. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers 2017*, 497-502.
- Wu, Y., Schuster, M., Chen, Z., Le, Q.V., Norouzi, M., Macherey, W., ... & Klingner, J., 2016. Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*.