Cross-linguistic research, parallel corpora, and replication in the *Translation Mining* tradition

Bert Le Bruyn & Henriëtte de Swart

[review copy April 19, 2023]

Abstract | Parallel corpus research is an empirical approach to the study of the scope and limits of variation across languages. We start from the observation that it is one of the few empirical branches in linguistics for which the relevance of replication has not been recognized. We then argue for its relevance in the *Translation Mining* parallel corpus research tradition and reflect on how replication studies in this tradition are best carried out. Throughout the paper, we draw illustrations from Van der Klis et al.'s (2022) and Le Bruyn et al.'s (2023) work on the *HAVE*-PERFECT in Germanic and Romance as well as on the design decisions and preliminary results of a replication study that we are currently carrying out (Fuchs et al., *in prep.*).

Key-words | Cross-linguistic research, Parallel corpora, Replication, *Translation Mining*, *HAVE*-PERFECT

1. Introduction

The relevance of replication has been recognized in all empirical branches of linguistics, including corpus linguistics (Sönning & Werner 2020). However, recent state-of-the-art articles on parallel corpus research make no mention of replication, despite the fact that they come with recommendations for the future development of the field (De Sutter & Lefer 2020; Hasselgård 2020; Granger & Lefer 2022). We assume that the absence of replication studies in parallel corpus research (see, however, Ebeling 2016) is inspired by the general consensus that the study of parallel corpora should always be done in conjunction with the study of comparable corpora (Enghels et al. 2020) and that the latter can be regarded as the counterpart of replication in parallel corpus research. In Section 2, we argue that there is a central role for genuine replication in at least one parallel corpus research tradition, *viz.* the *Translation Mining* tradition (Le Bruyn et al. 2022; Le Bruyn et al. 2023). What such studies should look like is the topic of Section 3, where we argue for a specific implementation of replication in *Translation Mining* research, drawing illustrations from a replication study of Van der Klis et al. (2022)/Le Bruyn et al. (2023) that we are currently carrying out (Fuchs et al., *in prep.*). Section 4 concludes the paper.

2. Cross-linguistic corpus research, the *Translation Mining* tradition and the *why* of replication

Cross-linguistic research aims to inform linguistic theory by identifying the scope and limits of variation across languages. Two types of corpora lend themselves to cross-linguistic research: (i) comparable corpora – collections of texts in different languages that are matched in a number of ways (e.g., genre, date of publication of the texts, etc.), and (ii) parallel corpora – collections of source texts with their translations.

State-of-the-art mainstream cross-linguistic corpus research explores comparable and parallel corpora based on the following quantitative procedure:

(1) Procedure for the exploration of variation across languages

   i. collect maximally comparable samples of a particular phenomenon in two or more languages;
   ii. annotate a high number of linguistic properties in the different samples;

iii. look for relevant dimensions of contrast through exploratory techniques of bivariate or multivariate nature;
iv. the most discriminating variables constitute the input of more detailed quantitative (and qualitative) analyses, which will measure, model and try to explain the impact of observed contrasts.

Enghels et al. (2020:10-11)

Despite its merits, this procedure presents the challenge of presupposing knowledge about the properties that distinguish the instantiations of the phenomenon under investigation. Indeed, even though Step (ii) intends to cast the net wide, the procedure presupposes that the relevant variation can be captured by a subset of the properties that are annotated for. This challenge naturally extends to Steps (iii) and (iv) but also to Step (i), as cross-linguistic comparability can only be properly assessed if one already knows which properties to control for. We conclude that the above procedure makes sense if the researcher has a fairly precise idea of where to look for variation but that it is not generalizable given that it presupposes at least part of the knowledge it aims to uncover.

Translation Mining is an alternative to the procedure in (1) and has the advantage of not presupposing knowledge about what it is that underlies variation. In its standard exploratory variant, the starting point of *Translation Mining* is the qualitative analysis of a particular phenomenon in a single text and its translations. In the remainder of this section, we defend the *Translation Mining* approach and argue that it requires a replication-based setup. We base our discussion of *Translation Mining* on the two studies that we are currently replicating (Fuchs et al., *in prep.*) and that underlie the methodological discussion of replication in Section 3: Van der Klis et al. (2022) (henceforth *VDK*) and Le Bruyn et al. (2023) (henceforth *LB*). A full discussion of the results of VDK and LB goes beyond the methodological aims of the current paper and we will consequently focus on the findings that we will come back to in Section 3.

Defending Translation Mining

VDK and LB use Camus' *L'Étranger* and its translations to look into the variation of the *HAVE*-PERFECT (Dahl & Velupillai 2013), a construction that combines a present tense auxiliary (*HAVE* or *BE*) with a past participle. The language-specific names of the construction that will return throughout the paper are the following: *Passé Composé* (French), *Perfekt* (German), *Voltooid Tegenwoordige Tijd* (henceforth, *VTT* | Dutch), *Pretérito Perfecto Compuesto* (Spanish), and *Present Perfect* (English).

There are several advantages of the one text and its translations strategy that is central to the standard exploratory variant of *Translation Mining*. First, given that this is a parallel and not a comparable corpus, it allows researchers to adopt cross-linguistic comparability as a working assumption and prevents them from having to depend on the *a priori* assumptions of comparability underlying comparable corpora and the sampling step in (1i).

Within the *Translation Mining* tradition, the working assumption of cross-linguistic comparability is key and rather than taking individual instances of the phenomenon under investigation as the relevant datapoints, *Translation Mining* researchers base their analysis on translation tuples. For the *HAVE*-PERFECT, examples of tuples drawn from VDK are given in (2) and (3). We restrict ourselves here to a subset of the languages under investigation in VDK, *viz.* French, German, Dutch, Spanish and English. For concreteness, we present both the tuples themselves and their underlying data.

(2) <Passé Composé, Perfekt, VTT, Pretérito Perfecto Compuesto, Present Perfect>

'Nous l'**avons transportée** dans notre petite morgue.'

French | *Passé Composé*

'Wir **haben** sie in unsere kleine Leichenhalle **gebracht**.'

German | *Perfekt*

'Wij **hebben** haar naar ons lijkenhuisje **gebracht**.'

Dutch | *VTT*

'La **hemos transportado** a nuestro pequeño depósito.'

Spanish | *Pretérito Perfecto Compuesto*

'We'**ve transferred** her to our little mortuary.'

English | *Present Perfect*

(3) <Passé Composé, Perfekt, VTT, Pretérito Perfecto Compuesto, Present Perfect>

Je les voyais comme je n'**ai** jamais **vu** personne

French | *Passé Composé*

Ich sah sie, wie ich nie jemanden **gesehen habe**

German | *Perfekt*

Ik zag hen zoals ik nog nooit iemand **heb gezien**

Dutch | *VTT*

Los veía como nunca **he visto** a nadie

Spanish | *Pretérito Perfecto Compuesto*

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

English | *Present Perfect*

Next to providing illustrations of what translation tuples look like, (2) and (3) further show that the *HAVE*-PERFECT of each of the languages can appear in resultative (2) and negative existential contexts (3). We get back to this generalization in Section 3.3.2.

With translation tuples as the primary unit of analysis, summaries of the *L'Étranger* data can be presented in tuple frequency tables like *Table 1* and scatter-map visualizations as in *Figure 1*.

French	German	Dutch	Spanish	English	#
Passé Composé	**Perfekt**	**VTT**	**Pretérito Perfecto Compuesto**	**Present Perfect**	10
Passé Composé	**Perfekt**	**VTT**	**Pretérito Perfecto Compuesto**	Simple Past	6
Passé Composé	**Perfekt**	**VTT**	Pretérito Indefinido	Simple Past	25
Passé Composé	**Perfekt**	OVT	Pretérito Indefinido	Simple Past	297
Passé Composé	Präteritum	OVT	Pretérito Indefinido	Simple Past	19

Table 1: Frequency table of translation tuples of the Passé Composé in Chapters 1 to 3 of Camus' L'Étranger *(abridged)*

Table 1 contains the five most frequent translation tuples involving the *Passé Composé* in French. Rather than ordering them based on their frequencies, we ordered them based on the subset pattern that emerges from the data: the *Present Perfect* appears in a (proper) subset of the contexts that the *Pretérito Perfecto Compuesto* appears in, the *Pretérito Perfecto Compuesto* appears in a subset of the contexts the *VTT* appears in, and so forth for the *VTT*, the *Perfekt* and the *Passé Composé*.

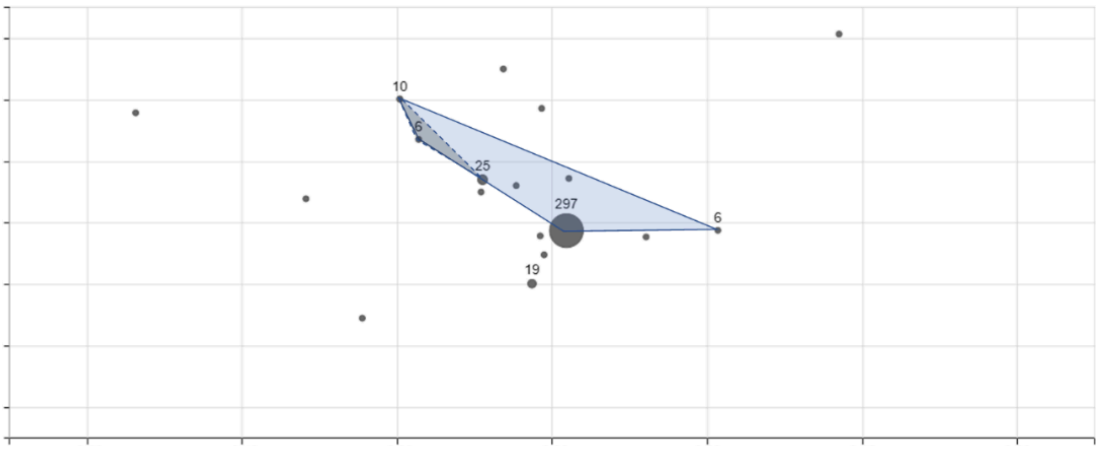

Figure 1: An MDS map of translation tuples of the Passé Composé in Camus' L'Étranger

The dots in *Figure 1* correspond to the translation tuples in the corpus. Tuples with the same tense combinations are merged into bigger dots with the accompanying number indicating how many tuples are merged. The positioning of the different tuples is generated automatically through Multidimensional Scaling (MDS) and optimizes the data for visual cross-linguistic comparison (see Van der Klis & Tellings (2022) for technical details). The markup we added shows how the Dutch *VTT* (dotted blue line) is used in a subset of the contexts that the German *Perfekt* (full blue line) appears in.

In *Translation Mining*, the comparisons of the contexts instantiating the different tuples feed the identification of the dimensions of variation. These dimensions are thus based on the variation that is attested in the data and are defined in a truly exploratory fashion with cross-linguistic comparability of parallel data as the crucial working assumption. In Section 3, we will argue that the dimensions of variation should be at the center of replication research and that the specific patterns they give rise to in a given dataset (e.g., the subset pattern in *Table 1* and *Figure 1*) are a powerful heuristic but ultimately subordinate to the dimensions themselves (see Sections 3.2, 3.3.2 and 3.3.3).

The second advantage of working with a single text and its translations is that researchers can focus on variants of the source and target languages written by a single author/translator in an overall consistent style. Variation is not only at play across languages but also within languages. For the *HAVE*-PERFECT, there is – among others – well-known geographical variation (see, e.g., Hennig 2000; Howe 2013; Azpiazu 2019; Valente 2021; Fuchs & Gonzalez 2022). We argue that taking all possible variation into account when doing an analysis of cross-linguistic differences for a given phenomenon is not a feasible enterprise, especially when the analysis is meant to be exploratory. Within the *Translation Mining* tradition, we consider restricting our attention to a single source text and its translations as a transparent way to control for the variation we take into account and deliver a high-quality analysis of the dimensions of variation that emerge from the data.

The third and final advantage of working with a single text and its translations is that it offers a unique microcosm in which researchers can explore the phenomenon they are interested in as part

of a bigger paradigm without running into the issue of variation we noted above. This is the advantage that Le Bruyn et al. (2023) exploit in an immediate extension of Van der Klis et al. (2022). Rather than only looking at translations of the French *Passé Composé*, Le Bruyn et al. (2023) look into translation tuples of all finite indicative tense forms. This allows them to explore how the *HAVE*-PERFECT relates to other tenses within and across the different languages and, e.g., to establish that there are interactions with the perfective past domain but not with the imperfective past domain and that the English *HAVE*-PERFECT is the only one to interact with the present domain. We illustrate these conclusions with the MDS maps in *Figures 2* to *4*. We note that these maps are based on data from Chapter 1 of *L'Étranger* (see also *Figure 4* in Le Bruyn et al. (2023)).

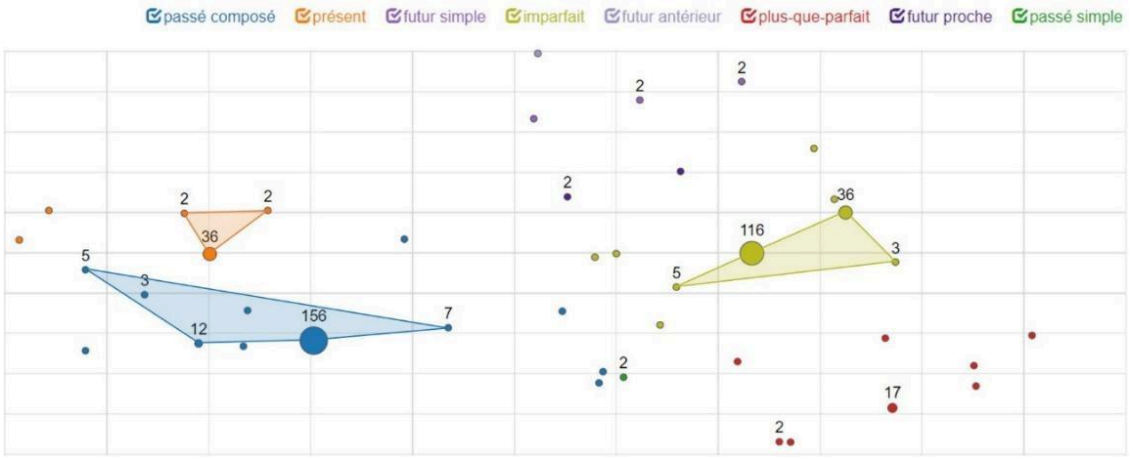

Figure 2: An MDS map of translation tuples of all finite indicative tense forms in Chapter 1 of Camus' L'Étranger | *Markup for French.*

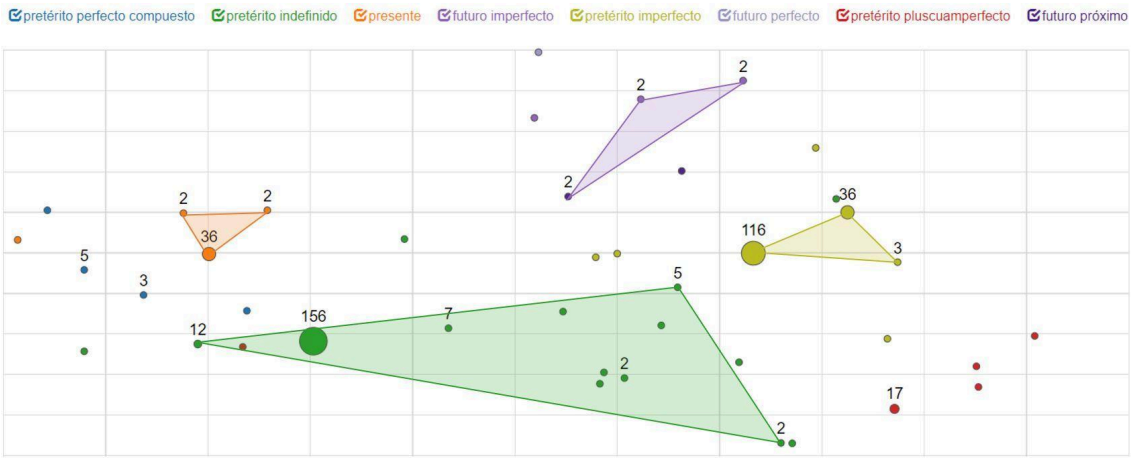

Figure 3: An MDS map of translation tuples of all finite indicative tense forms in Chapter 1 of Camus' L'Étranger | *Markup for Spanish.*

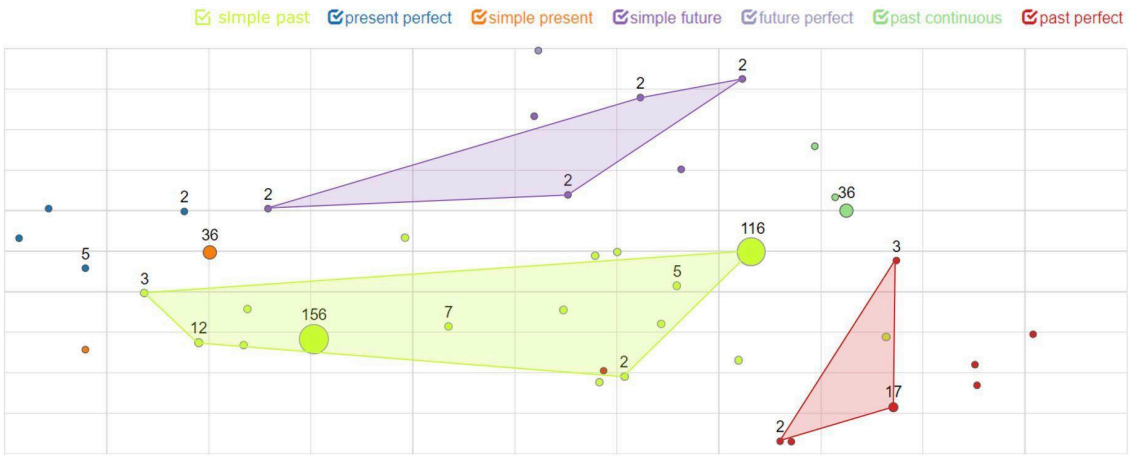

Figure 4: An MDS map of translation tuples of all finite indicative tense forms in Chapter 1 of Camus' L'Étranger | *Markup for English.*

From *Figure 2* to *Figure 3*, we see that the number of blue dots decreases and that this decrease is at the benefit of the *Pretérito Indefinido* (dark green) – the Spanish PERFECTIVE PAST – and not at the benefit of the *Pretérito Imperfecto* (olive green) – the Spanish IMPERFECTIVE PAST. This confirms the intuition from the literature that the *HAVE*-PERFECT encroaches on the perfective past domain (Harris 1982; Squartini & Bertinetto 2000). A relevant datapoint is given in (4):

(4) Il **a perdu** son oncle, il y a quelques mois.

French | *Passé Composé*

Er **hat** vor ein paar Monaten seinen Onkel **verloren**.

German | *Perfekt*

Hij **heeft** zijn oom een paar maanden geleden **verloren**.

Dutch | *VTT*

Perdió a su tío hace algunos meses.

Spanish | *Pretérito Indefinido*

He lost his uncle, a few months ago.

English | *Simple Past*

French, German, and Dutch use their *HAVE*-PERFECT in (4) but Spanish resorts to its PERFECTIVE PAST. In line with Van der Klis et al. (2022), Le Bruyn et al. (2023) hypothesize that one of the dimensions that distinguishes between the languages is that the Spanish *HAVE*-PERFECT is sensitive to hodiernality, imposing the use of its PERFECTIVE PAST in a context like (4) that refers to an event that took place before the day of speech (see, for a more fine-grained discussion, Van der Klis et al. (2022) and de Swart et al. (2022)). Hodiernality is one of the dimensions of variation that we return to in our discussion of replication (see Section 3.3.1).

From *Figures 2* and *3* to *Figure 4*, we see that two orange dots turn blue, showing an interaction between the *HAVE*-PERFECT and the present domain in English. (5) presents the underlying data of one of the translation tuples:

(5) 'Il y a longtemps que vous **êtes** là ?'

French | *Présent*

'**Sind** Sie schon lange hier?'

German | *Präsens*

'Hoe lang **bent** u al hier?'

Dutch | *OTT*

‘¿Hace mucho tiempo que **está** usted aquí ?’

Spanish | *Presente*

‘**Have** you **been** here long?’

English | *Present Perfect*

Datapoints like (5) show that English allows for its *HAVE*-PERFECT with state verbs in continuative contexts (Portner 2003) whereas the other languages opt for the PRESENT (see also de Swart et al. (2022) for discussion). Continuativity is the other dimension of variation that we return to in our discussion of replication (see Section 3.3.2).

The significance of the patterns in *Figures 2* to *4* depends on them stemming from the different linguistic choices of one author/translator per language. For the datapoint in (5), for instance, its linguistic relevance is supported by the fact that the English translator opts for a *HAVE*-PERFECT in this particular context but consistently follows Camus and the other translators in choosing a PRESENT in 36 other contexts (see *Figures 2*, *3* and *4*). If (5) had been part of a set of randomly chosen datapoints from a variety of texts, it would have been impossible to oppose the variation in it to the consistent choice for the PRESENT in the 36 other contexts. And unless we had annotated the data for continuative contexts, it would have proven hard to distinguish the cross-linguistic nature of the variation in (5) from – among others – language-internal geographical or idiosyncratic variation. We conclude that relying on a single text and its translations allows researchers to abstract away from various dimensions of variation and focus on generating data-driven hypotheses about the dimensions of cross-linguistic variation of the phenomenon they are interested in and the paradigm it belongs to.

The need for replication

Up till now, we have introduced *Translation Mining* as an alternative to mainstream exploratory cross-linguistic corpus research and we have argued for the advantages of the initial setup of a *Translation Mining* study. However, stopping the analysis after the study of one text and its translations would not bode well with the way we have argued for the advantages of using this setup as a starting point. Indeed, we defined cross-linguistic comparability as a working assumption and argued for the restriction to one text and its translations as a way to bring focus to an exploratory analysis. As working assumptions need to be assessed and an initial focus invites extensions, there are several steps to be taken after the first exploratory analysis. In the remainder of this section, we argue that replication should be one of them.

Within mainstream cross-linguistic corpus research, it is standard to follow up a parallel corpus study with the study of a comparable corpus. The rationale behind this move is that it allows researchers to control for the influence of translation. Even though we see the relevance of triangulation, we argue that – at least within the *Translation Mining* tradition – it should be preceded by replication. As we pointed out before, there is no *a priori* way to assess what it is that makes a context in one language comparable to a context in another language. It consequently only makes sense to try and triangulate results from parallel corpus research on a comparable corpus when we have sufficient control over the dimensions of variation that we hypothesize to be at play. Crucially, initial *Translation Mining* studies based on one text and its translations do not come with this guarantee and making the step from parallel to comparable corpora right away might turn out to be unproductive. Indeed, successful triangulation would confirm the relevance of the dimensions that were hypothesized on the basis of the parallel corpus, but their refinement would not be straightforward, as comparability would become an issue again. Conversely, unsuccessful triangulation might be due to the interaction with a previously unidentified dimension rather than

indicating problems with the dimensions that were hypothesized to be at play. For comparability reasons, exploring these new dimensions is however best done in a parallel corpus. What we find then is that, independently of the projected outcome of triangulation on a comparable corpus, replication imposes itself as the crucial way forward in the *Translation Mining* tradition, be it to identify new dimensions and the way they interact with previously hypothesized ones or to refine the latter. We conclude that – at least within the *Translation Mining* tradition – a replication-based approach is to be preferred over early triangulation. Replication allows researchers to check whether new data are in line with previously hypothesized dimensions of variation, building up critical mass, and – in the process – assess the working assumption of cross-linguistic variation. The output of replication studies then provides the input for triangulation, be it within a corpus-based or an experimental paradigm (see, e.g., Fuchs & Van der Klis 2022).

3. The *how* of replication in the *Translation Mining* tradition

In Section 2, we established the need for genuine replication studies within the *Translation Mining* tradition. In this section, we present a number of methodological reflections on what such studies should look like, taking illustrations from a replication study of Van der Klis et al. (2022)/Le Bruyn et al. (2023) that we are currently carrying out (Fuchs et al., *in prep.*). We follow the setup of an empirical research report leading us from the data and objectives (Section 3.1) over the methodology (Section 3.2) to the results and discussion (Section 3.3). We note that we focus on *how* to run a *Translation Mining* replication study. In 3.3, we will consequently focus on methodological reflections, a full interpretation of the results lying beyond the scope of the paper.

3.1 Data and objectives

For our replication study, we selected J.K. Rowling's *Harry Potter and the Philosopher's Stone* and its translations as our parallel corpus. In the spirit of replication, we thus adopt the one text and its translations strategy from Van der Klis et al. (2022)/Le Bruyn et al. (2023). While browsing through the English original and the translations, we noticed that there were significantly fewer *HAVE*-PERFECTS than in *L'Étranger*. To compensate for this, we extracted over four times more finite indicative tense forms from the English original than we did for *L'Étranger* in Le Bruyn et al. (2023), leading to a database spanning Chapters 1, 16 and 17 and including over 1700 contexts. After alignment, this led to 157 contexts that have a *HAVE*-PERFECT in the original or in at least one of the translations. The final dataset we selected consists of the corresponding translation tuples that are attested at least twice (N=131).

The aims we set for the replication study are to check whether the new data are in line with the dimensions of cross-linguistic variation that Van der Klis et al. (2022)/Le Bruyn et al. (2023) (henceforth, *VDK&LB*) hypothesized on the basis of the *L'Étranger* dataset, and to refine, extend or adjust them where relevant. By checking the status of previously hypothesized dimensions of variation, we also indirectly assess our working assumption of cross-linguistic comparability: if the new data are in line with the previously hypothesized dimensions of variation, this implies that original texts and their translations are comparable enough to lead to robust generalizations across different corpora.

3.2 Methodology

In replication studies, the methodology needs to be the same as in the original studies. In the case of a *Translation Mining* study, this means that we generate an overview of the different translation tuples and qualitatively compare the contexts that instantiate them, assessing the internal

homogeneity of the sets of contexts that fall under identical translation tuples and interpreting the heterogeneity of the sets of contexts that fall under different translation tuples.

Replicating the role of the dimensions of cross-linguistic variation hypothesized by VDK&LB crucially does not require the *HAVE*-PERFECTS in the Harry Potter dataset to display the same patterns as the ones we found for the *L'Étranger* dataset, *viz.* an interaction with the perfective past domain in the form of a subset relation, no interaction with the imperfective past domain and, for the present domain, an interaction that is limited to the English *HAVE*-PERFECT (cf. *Table 1*). Rather, what we expect to find is that the variation in the Harry Potter dataset is in line with the dimensions of variation hypothesized by VDK&LB. If these are the only dimensions at play, we expect to find the same patterns but if they interact with previously unidentified dimensions, we expect to see novel patterns.

3.3 Results and discussion

Table 2 lists all the translation tuples that are attested at least twice in the Harry Potter dataset with the last column indicating their frequencies. We added row numbers to facilitate reference to sets of translation tuples.

	French	German	Dutch	Spanish	English	#
1	Passé Composé	Perfekt	VTT	Pretérito Indefinido	Simple Past	39
2	Passé Composé	Perfekt	OVT	Pretérito Indefinido	Simple Past	24
3	Passé Composé	Präteritum	OVT	Pretérito Indefinido	Simple Past	13
4	Passé Composé	Perfekt	VTT	Pretérito Perfecto Compuesto	Present Perfect	12
5	Imparfait	Perfekt	OVT	Pretérito Imperfecto	Simple Past	6
6	Passé Composé	Präsens	OTT	Presente	Simple Present	6
7	Passé Composé	Perfekt	VTT	Pretérito Indefinido	Present Perfect	5
8	Passé Composé	Perfekt	VTT	Pretérito Perfecto Compuesto	Simple Past	4
9	Passé Composé	Präteritum	OVT	Pretérito Imperfecto	Simple Past	4
10	Présent	Prasens	OTT	Presente	Present Perfect	4
11	Passé Composé	Präsens	VTT	Pretérito Perfecto Compuesto	Present Perfect	3
12	Présent	Präsens	OTT	Pretérito Perfecto Compuesto	Present Perfect	3
13	Passé Composé	Perfekt	OTT	Pretérito Perfecto Compuesto	Present Perfect Continuous	2
14	Passé Composé	Präsens	OTT	Pretérito Indefinido	Present Perfect	2
15	Passé Composé	Präteritum	VTT	Pretérito Indefinido	Simple Past	2
16	Présent	Perfekt	VTT	Pretérito Perfecto Compuesto	Present Perfect	2

Table 3: Frequency table of the translation tuples that occur at least twice in Chapters 1, 16 and 17 of Harry Potter and the Philosopher's Stone *and contain a* HAVE-*PERFECT in at least one of the language*

Before zooming in on a selection of the data, we briefly reflect on the overall similarities and differences between the *L'Étranger* and the Harry Potter datasets. Focusing first on the top translation tuples in *Table 2*, we overall find the same subset relation of the *HAVE*-PERFECT in the perfective past domain (lines 1,2,3,4 and 8) and the same special position of the English *HAVE*-PERFECT in the present domain (line 10). These preliminary similarities cannot hide major differences, though. To bring these out, we generated *Table 3*, the exact counterpart of *Table 2* for Chapter 1 of Camus' *L'Étranger*.

	French	German	Dutch	Spanish	English	#
1	Passé Composé	Perfekt	OVT	Pretérito Indefinido	Simple Past	156
2	Passé Composé	Perfekt	VTT	Pretérito Indefinido	Simple Past	12

3	Passé Composé	Präteritum	OVT	Pretérito Indefinido	Simple Past	9
4	Passé Composé	Perfekt	VTT	Pretérito Perfecto Compuesto	Present Perfect	5
5	Passé Composé	Perfekt	VTT	Pretérito Perfecto Compuesto	Simple Past	3
6	Présent	Präsens	OTT	Presente	Present Perfect	2

Table 3: Frequency table of the translation tuples that occur at least twice in Chapter 1 of L'Étranger *and contain a* HAVE-*PERFECT in at least one of the language*

Even a cursory comparison of *Tables 2* and *3* shows that there are far more different translation tuples in the Harry Potter dataset than in the *L'Étranger* dataset (16 vs. 6). Closer scrutiny reveals that this increase is not due to the size of the Harry Potter dataset, as it contains fewer translation tuples than the *L'Étranger* dataset (131 vs. 187). In Le Bruyn et al. (2023), we developed a measure to express the variation that a given tense displays in a dataset, *viz.* the Overall Variation Value of a category (henceforth, *OVV*). For the *HAVE*-PERFECT, the OVV takes into account the number of different translation tuples it appears in as well as their frequencies, striking a balance between high frequency as an indicator of a relatively stable pattern and low frequency as an indicator of a relatively unstable or even accidental pattern. For this measure too, we find that the *HAVE*-PERFECT displays a lot more variation in the Harry Potter than in the *L'Étranger* dataset (3,28 vs. 0,99). All other things being equal, the conclusion that suggests itself is that the dimensions of variation identified by VDK&LB may play a role in accounting for the cross-linguistic variation in the Harry Potter dataset but that their coverage will be fairly limited. We will see that this conclusion is premature, but that the Harry Potter data do push the need for detailed analysis, both at the qualitative and at the quantitative level.

In what follows, we discuss three selections of sets of translation tuples. For each selection, we introduce the relevant dimensions of variation hypothesized by VDK&LB and we discuss how the Harry Potter data relate to these. Each selection allows us to present a reflection on the objectives or methodology of replication studies in the *Translation Mining* tradition. The selection allows us to touch upon the interactions of the *HAVE*-PERFECT with the perfective past domain (Section 3.3.1), the present domain (Section 3.3.2), and the imperfective past domain (Section 3.3.3).

3.3.1 Straightforward replication: hodiernality and the Spanish perfective past domain

In our discussion of VDK&LB in Section 2, we referred to their hypothesis that Spanish is different from French, German, and Dutch in that its *HAVE*-PERFECT is barred from reference to past events that occurred before the day of speech. Example (4) illustrated this for the *L'Étranger* dataset. Example (6) repeats the observation for the Harry Potter dataset:[1]

(6) a. Dumbledore **gave** me the day off yesterday ter fix it.

English | *Simple Past*

b. Dumbledore m'**a accordé** un jour de congé hier pour le préparer.

French | *Passé Composé*

c. Dumbledore **hat** mir gestern dafür **freigegeben**.

German | *Perfekt*

d. Perkamentus **heb** me gisteren vrijaf **gegeven** om 't te regelen.

Dutch | *VTT*

e. Dumbledore me **dio** libre el día de ayer para hacerlo.

Spanish | *Pretérito Indefinido*

[1] The speaker in the English original is intended to have a West Country accent. The use of *ter* instead of *to* is a reflection of this.

The speaker in (6) is Hagrid, who refers to the event of Dumbledore giving him a day off. Given that the day off was the day before the day of speech, permission must have been given before as well. The tense use in (6) shows that Spanish resorts to its PERFECTIVE PAST for the pre-hodiernal past domain whereas French, German, and Dutch rely on their *HAVE*-PERFECTs. This tense use is in line with the role of hodiernality VDK&LB hypothesized to account for the cross-linguistic variation of the *HAVE*-PERFECT.

The anecdotal replication of the role of hodiernality in (6) can also be confirmed in a more thorough way by looking at the opposition between the Spanish *HAVE*-PERFECT and one of its closest relatives, *viz.* the Dutch *HAVE*-PERFECT. The relevant sets of translation tuples are those in which Dutch uses a *VTT* and Spanish relies on a *Pretérito Perfecto Compuesto* or a *Pretérito Indefinido* (lines 1, 4, 7, 8, 11, 15 and 16 of *Table 2*, n=67). The prediction VDK&LB make is that the *VTT* can be used both for events that occurred on or before the day of speech whereas the *Pretérito Perfecto Compuesto* is restricted to the hodiernal past and the *Pretérito Indefinido* imposes itself for the pre-hodiernal past. To check this prediction, we annotated the data underlying the sets of translation tuples for whether they referred to an event on or before the day of speech. This led to unambiguous annotations for 42 contexts, 14 referring to events on the day of speech and 28 referring to events before the day of speech. These annotations show that the *VTT* occurs in hodiernal and pre-hodiernal contexts. As for the Spanish *Pretérito Perfecto Compuesto*, it turns out to be limited to events on the day of speech (n=7). These data are in line with the dimension of hodiernality that VDK&LB hypothesize to be at play in the cross-linguistic variation of the *HAVE*-PERFECT.

The discussion of hodiernality allows us to reflect on the objectives of replication studies. As we pointed out in Section 3.1, the primary objective is to check whether the new data are in line with previously hypothesized dimensions of variation. For hodiernality, this is clearly the case. We underline that establishing that new data are in line with previously hypothesized dimensions does not unequivocally establish their role in cross-linguistic variation. It does, however, allow us to argue in favor of their robustness, in particular in light of the fact that the dimension's role turns out to be unaffected by the differences between the corpus of the initial study and the corpus of the replication study. In the case of the *L'Étranger* and the Harry Potter corpus, these differences include author, translators, and source language.

3.3.2 Meaning composition as a factor of variation: interactions with the present domain

In Section 2, we pointed out that the *L'Étranger* data led VDK&LB to the hypothesis that the interaction between the *HAVE*-PERFECT and the present domain is limited to English and specifically to stative continuative contexts like the one in (5). In the Harry Potter dataset, we also find that stative continuative contexts rely on the *HAVE*-PERFECT in English and on the PRESENT in the other languages:

(7) a. 'How long **have** I **been** in here?'
English | *Present Perfect*

b. 'Ça fait combien de temps que je **suis** là ?'
French | *Présent*

c. 'Wie lange **bin** ich schon hier?'
German | *Präsens*

d. 'Hoe lang **lig** ik hier al?'
Dutch | *OTT*

e. '¿Cuánto tiempo hace que **estoy** aquí ?'
Spanish | *Presente*

The context in (7) is a continuative one in which the speaker asks how long he has been in the state of being somewhere and nothing indicates that this state will end any time soon. The choice of tenses is identical to the one we found in (5): a *HAVE*-PERFECT in English and a PRESENT in the other languages. Despite this initial confirmation of the relevance of stative continuative contexts in cross-linguistic variation, the data in Table 2 suggest a far more involved interaction between the *HAVE*-PERFECT and the present domain than VDK&LB would lead us to believe. Indeed, if the interaction were limited to stative continuative contexts in English, we would predict *HAVE*-PERFECTS appearing alongside PRESENTS in a single set of translation tuples, *viz.* the one in line 10. However, we find many more of these sets of tuples with *HAVE*-PERFECTS not being restricted to English (lines 6, 11, 12, 13 and 16 of *Table 2*). In what follows, we take a closer look at these sets but leave aside the set in line 13 as it includes a *Present Perfect Continuous* in the source text and deserves a separate treatment.

Close analysis of the underlying data of the translation tuples in lines 6, 11, 12, and 16 leads us to the hypothesis that the variation we find is not due to previously unidentified dimensions of variation of the *HAVE*-PERFECT but to different language-independent ways of composing the same meaning. We argue for this claim on the basis of examples (8) to (12). For (8) and (11), we list all the languages, for (9), (10) and (12), we select the ones we need for our argumentation. Given that the languages in (9), (10) and (12) rely on different strategies to render the same meaning, we exceptionally provide glosses and translations for the non-English versions.

(8) a. See what I **have become**?

English | *Present Perfect*

b. Tu vois ce que je **suis devenu** ?

French | *Passé Composé*

c. Siehst du was au mir **geworden ist**?

German | *Perfekt*

d. Zie je wat er van me **geworden is**?

Dutch | *VTT*

e. ¿Ves en lo que me **he convertido** ?

Spanish | *Pretérito Perfecto Compuesto*

(9) a. Quirrel **does not have** the Stone.

English| *Present*

b. Quirrel **n'** **a** **pas** **volé** la Pierre.
Quirrel NEG have NEG stolen the Stone
'Quirrel hasn't stolen the Stone.'

French | *Passé Composé*

(10) a. I mean, he **hasn't gone**, has he?

English | *Present Perfect*

b. Ich meine, er **ist** **nicht** für immer **auf** **und** **davon**, oder?
I mean he is not for ever off and away or
'I mean, he isn't gone forever, is he?'

German | *Präsens*

(11) a. you **have**n't **heard** from your sister lately, have you?

English | *Present Perfect*

b. tu n'**as** pas **eu** de nouvelles de ta sœur récemment ?

French | *Passé Composé*

c. du **hast** in letzter Zeit nichts von deiner Schwester **gehört**, oder?

German | *Perfekt*

d. je **hebt** de laatste tijd toch toevallig niet iets van je zus **gehoord**?

Dutch | *VTT*

e. ¿**has sabido** últimamente algo sobre tu hermana?

Spanish | *Pretérito Perfecto Compuesto*

(12) a. Anyway, we**'ve never had** any proof Snape found out how to get past Fluffy.

English | *Present Perfect*

b. De	todos	modos,	**nunca**	**hemos**	**tenido**	pruebas	de	que	Snape
of	all	ways	never	have	had	evidence	of	that	Snape
encontrara	la	forma	de	burlar	a	FLUFFY.			
discovered	the	way	to	deceive	A	Fluffy			

'Anyway, we have never had evidence of Snape discovering how to deceive Fluffy.'

Spanish | *Pretérito Perfecto Compuesto*

c. Bovendien	**hebben**	we	geen	enkel	bewijs	dat	Sneep	ooit
moreover	have	we	not	any	evidence	that	Sneep	ever
heeft	uitgevogeld	hoe	hij	langs	Pluisje	moet	komen.	
has	figured_out	how	he	past	Fluffy	must	come	

'We moreover have no proof that Snape ever found out how to get past Fluffy.'

Dutch | *OTT*

Example (8) allows us to make the baseline observation that all languages in the Harry Potter dataset allow their *HAVE*-PERFECTs in resultative contexts. This is fully in line with what we found in (2) for the *L'Étranger* dataset. Moving to (9) and (10), we find that the result of a past action (or the lack thereof) can be expressed both compositionally through the use of the *HAVE*-PERFECT on an eventive verb or lexically through the use of a stative verb in the PRESENT. In (9), this is exemplified by the pair *to steal* (9b) / *to have* (9a). In (10), we find the variation in the other direction with *to go* (10a) / *to be gone* (10b). On the basis of the comparison of (8) and (9)/(10), we conclude that all languages pattern alike in the ability of their *HAVE*-PERFECT to convey a result but that the same meaning can be expressed lexically with a PRESENT stative verb. This is an important conclusion as it points to a way for translation tuples to show interactions between the *HAVE*-PERFECT and the present domain that are independent of the cross-linguistic variation of the *HAVE*-PERFECT.

Moving to (11), we make the baseline observation that all languages allow their *HAVE*-PERFECTs in negative existential contexts, conveying that something did not happen at any point in a given time frame ('lately' in (11)). This observation is in line with what we found in (3) for the *L'Étranger* dataset, where the time frame included all time points preceding the moment of speech. (12) however shows that this same meaning can be expressed in at least two ways in more complex sentences with a stative verb in the main clause: if the time frame is explicitly mentioned in the main clause, the *HAVE*-PERFECT is required (12a,b) but if the time frame is added to the subordinate clause, the PRESENT is required (12c). We conclude that all languages pattern alike in allowing their *HAVE*-PERFECTs to appear in negative existential contexts but that the same meaning can be expressed with a slightly different compositional make-up involving PRESENT state verbs in the matrix clause with a time frame adverb in the subordinate clause. This conclusion points to yet another way in which interactions between the

HAVE-PERFECT and the present domain can manifest themselves without there being any cross-linguistic variation at the level of the *HAVE*-PERFECT itself.

In our discussion of examples (7) to (12), we have argued that interactions between the *HAVE*-PERFECT and the present domain in our data do not always arise because of cross-linguistic variation of the *HAVE*-PERFECT. We distinguished the following cases with the first being the only one to depend on cross-linguistic variation:

i. state verbs in continuative contexts require the *HAVE*-PERFECT in English but typically go with the PRESENT in the other languages (cf. VDK&LB and (7));
ii. resultative interpretations of eventive verbs require the *HAVE*-PERFECT but result-state interpretations of stative verbs require the PRESENT (cf. (8)-(10));
iii. negative existential interpretations of state verbs require the *HAVE*-PERFECT if there is an explicit time span adverbial but the PRESENT in its absence (cf. (11)-(12)).

Figure 5 brings together these different cases in a decision tree-like representation. The relevant factors in the tree are language, aspectual class, and adverbs. We hypothesize that – together – they account for whether the PRESENT or the *HAVE*-PERFECT are used in our data:

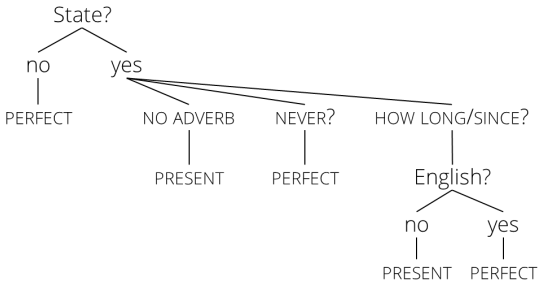

Figure 5: Decision tree summarizing the interactions between language, aspectual class and adverbs in tense choice and resulting interpretations in lines 6, 11, 12 and 16 of Table 2

To track the explanatory value of the hypotheses underlying the decision tree in *Figure 5* for all translation tuples in lines 6, 11, 12 and 16, we need a statistical model that is able to deal with small datasets and multiple interactions. We selected *Conditional Inference Trees* (CIT, Tagliamonte & Baayen 2012) as our model. The output of CIT organizes dimensions of variation as a decision tree, much like the representation in *Figure 5*. What the model does is to check whether the independent variables we select – the dimensions of variation – are significantly associated with the response variable in the dataset – the choice of tenses. If so, it evaluates which of them has the strongest association and uses the outcome of this evaluation to introduce a binary split in the dataset based on the values of the independent variable. These steps are repeated until no further significant associations are found ($\alpha = 0.05$). Given that the variation in hypotheses i. to iii. is not only cross-linguistic but also language internal, the input to the CIT is the individual language data underlying the different tuples. With lines 6, 11, 12 and 16 of *Table 2* covering 18 tuples with five languages, there are a total of 90 datapoints.

We annotated the datapoints for tense (PRESENT vs. *HAVE*-PERFECT), language (ENGLISH vs. REST), aspectual class (STATE vs. NON-STATE) and adverbs in the same clause as the verb (NONE, NEVER, SINCE/HOW LONG). *Figure 6* presents the CIT output:

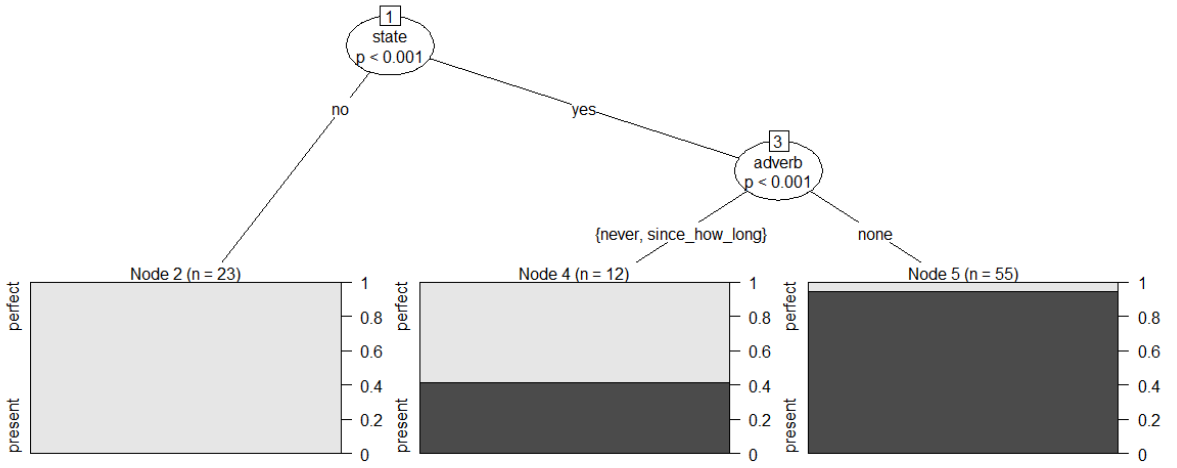

Figure 6: Conditional Inference Tree output for the data underlying the translation tuples in lines 6, 11, 12 and 16

We discuss *Figure 6* from top to bottom. Aspectual class has a significant effect and non-state verbs always appear in the *HAVE*-PERFECT. For state verbs, the tree shows an interaction with adverbs. In the absence of time frame adverbials like *never* and adverbial phrases indicating continuativity (*since* X / *how long*), the default choice is the PRESENT. In the presence of such adverbs, tense choice is mixed.

Even though *Figure 6* and *Figure 5* are different in a number of respects, the underlying generalizations are the same. With non-stative verbs, we only find the PERFECT. Relevant examples are (9b) and (10a), conveying a resultative interpretation. For stative verbs, the absence of adverbs leads to a resultative interpretation conveyed by the PRESENT (see (9a) and (10b)) or – with adverbs appearing in the subordinate clause – to a negative existential interpretation conveyed by the same tense (see (12c)). Once adverbs appear in the same clause as the verb, there is more variation, and we find both PRESENTS and PERFECTS conveying negative existential and continuative interpretations. For negative existential interpretations, we expected there to be no cross-linguistic variation but for continuative interpretations, we had expected there to be a role for language, opposing English to the other languages in the data. Even though this role of language does not come out of the model, we can find it back when zooming in on the datapoints that take a PRESENT in the presence of adverbs, all of them involving adverbs like *since* and *how long* and come from languages other than English (see, e.g., (7b) to (7e)). We conclude that the hypotheses we formulated on the basis of a qualitative analysis of the data are backed up by a quantitative analysis.

In this section, we looked into the interaction between the *HAVE*-PERFECT and the present domain. In line with VDK&LB, we found that stative continuative contexts set the English *Present Perfect* apart from the other *HAVE*-PERFECTS. However, our data show far more interactions with the present domain than VDK&LB lead us to believe. Throughout, we have argued that the interactions we found outside stative continuative contexts are independent of cross-linguistic variation and reflect different language-independent compositional ways of arriving at the same meaning.

For replication studies in the *Translation Mining* tradition, our discussion on cross-linguistic variation and compositionality invites a number of reflections on objectives and methodology. First, next to cross-linguistic variation at the level of the phenomenon under investigation, variation in the data can be due to language-independent compositional ways of deriving the same meanings. Compositionality of meaning is thus an important source of variation in the data that has to be controlled for. Second, properly controlling for the role of compositionality requires the use of

statistical models like CIT that are able to spell out the impact of compositionality on the data. Third, the input for these statistical models consists of individual language datapoints, highlighting the complementary value of analyses at the level of sets of translation tuples and at the level of their underlying data. Abstracting over these reflections, the general conclusion to be drawn is that replication studies within the *Translation Mining* tradition should aim for detailed analysis, both at the qualitative and at the quantitative level. We also note that the increased interactions we found between the *HAVE*-PERFECT and the present domain did not lead us to question any of the dimensions of variation identified by VDK&LB but do have a major impact on the overall patterns we find in our data. This confirms that data patterns are subordinate to dimensions of variation and that replication studies should focus on the latter.

3.3.3 States and the German *Perfekt*: a new dimension of variation in the imperfective past domain

To conclude our discussion of selections of translation tuples, we selected line 5 in *Table 2*, where we find *HAVE*-PERFECTs in German but (imperfective) past tenses in the other languages. (13) illustrates the type of contexts involved:

(13) a. He was at Hogwarts with your father, **did**n't you **know**?

English | *Simple Past*

b. Il était à Poudlard avec votre père, vous ne le **saviez** pas ?

French | *Imparfait*

c. Er und Ihr Vater waren zusammen in Hogwarts, **haben** Sie das nicht **gewusst**?

German | *Perfekt*

d. Hij heeft samen met je vader aan Zweinstein gestudeerd, **wist** je dat niet?

Dutch | *OVT*

e. Estaba en Hogwarts con tu padre, ¿ no lo **sabías** ?

Spanish | *Pretérito Imperfecto*

The use of the *HAVE*-PERFECT on state verbs like *to know* is known from the literature on German (e.g., Musan 2001) but has – to our knowledge – not been discussed in the literature on the other languages. VDK&LB do report on state verbs with a *HAVE*-PERFECT, but these are limited to narrative contexts in which the states are interpreted as episodes in narrative sequences:

(14) a. parce qu'il **a fallu** que je monte chez Emmanuel pour lui emprunter une cravate noire et un brassard

French | *Passé Composé*

b. weil ich noch zu Emmanuel hinauf **musste**, um mir einen schwarzen Schlips une eine Trauerbande von ihm zu borgen

German | *Präteritum*

c. omdat ik bij Emmanuel een zwarte das en een rouwband **moest** gaan lenen

Dutch | *OVT*

d. porque **fue** necesario que subiera a casa de Emanuel para que me prestase una corbata negra y un brazalete

Spanish | *Pretérito Indefinido*

e. because I **had** to go up to Emmanuel's place to borrow a black tie and armband

English | *Simple Past*

(14) is part of the beginning of *L'Étranger* and the narrator/protagonist reports on the events that preceded his leaving by bus to mourn his mother. We find the *HAVE*-PERFECT in French but not in the

other languages. (13) is different in three respects: the state of knowing is in no way part of a narrative sequence, German but not French opts for a *HAVE*-PERFECT, and Spanish relies on the IMPERFECTIVE PAST rather than on the PERFECTIVE PAST. For these reasons, we hypothesize that the use of the *Perfekt* in (13) requires the identification of a new dimension of variation.

Our discussion of the set of translation tuples exemplified by (13) allows us to add to our reflections on the objectives and methodology of replication studies within the *Translation Mining* tradition. Sections 3.3.1 and 3.3.2 showed how replication studies have to check whether new data are in line with previously hypothesized dimensions of variation and control for variation in the data that is independent of cross-linguistic variation of the phenomenon under investigation. The discussion in the current section reminds us of the fact that replication studies in the *Translation Mining* tradition should always be on the lookout for new dimensions of variation if the data call for these. As we pointed out in Section 2, the standard exploratory variant of *Translation Mining* relies on one text and its translations. This holds for initial studies as well as for replication studies. As a consequence, parallel corpus research in the *Translation Mining* tradition never comes with the assumption – let alone the guarantee – of covering all relevant data. Rather than trying to reduce replication studies to checking whether new data are in line with previously hypothesized dimensions of variation, they should consequently also come with an exploratory goal. The upshot of this is that replication in the *Translation Mining* tradition should be an iterative process, aiming for the accumulation of critical mass across studies. The final point we want to make is that the new dimension of variation we identified has no impact on previously identified dimensions but does have a serious impact on the patterns in our data: where the *L'Étranger* data only showed interactions with the present and the perfective past domain, the Harry Potter data add an interaction between the *HAVE*-PERFECT and the imperfective past domain. This is an important finding *per se* but also allows us to reiterate our methodological point that replication within the *Translation Mining* tradition should not be about replicating patterns in the data but about replicating the dimensions of variation that underlie them.

3.4 The *how* of replication in the *Translation Mining* tradition: conclusion

In this section, we presented a number of reflections on how to run a replication study in the *Translation Mining* tradition. For concreteness, we built our reflections on design decisions and preliminary findings of a replication study of Van der Klis et al. (2022)/Le Bruyn et al. (2023) (VDK&LB) that we are currently carrying out (Fuchs et al., *in prep.*). In line with the standard requirement of replication studies to be as close as possible to their original studies, we opted for the one text and its translations approach from VDK&LB as well as for a starting point in the qualitative analysis of differences and similarities between sets of different translation tuples. We further argued that the objectives of replication studies should be to test the robustness of the dimensions of variation against the background of the differences between the corpora (Section 3.3.1) as well as to hypothesize new dimensions of variation where required by the data (Section 3.3.3), leading to an iterative approach to replication and the buildup of critical mass across studies. Finally, we argued that compositionality of meaning can lead to variation in the data that is independent of cross-linguistic variation of the phenomenon under investigation and has to be controlled for, requiring detailed analysis both at the qualitative and at the quantitative level (3.3.2).

4. Concluding remarks

The starting point of this paper was the observation that the relevance of replication studies has been recognized in all empirical branches of linguistics except in parallel corpus research (Section 1). Rather than arguing for its relevance for parallel corpus research in general, we focused on its relevance for *Translation Mining*, highlighting the need in this tradition to build up critical mass

across different studies and – in the process – to assess the assumption of cross-linguistic comparability of original texts and their translations (Section 2). We concluded the paper with a number of reflections on *how* to run a replication study, arguing for the iterativity of replication and the need for detailed analysis, both at the qualitative and at the quantitative level (Section 3).

Throughout the paper, we hope to have illustrated the strengths of the *Translation Mining* tradition, especially in the replication-based setup we argued for. In Section 2, we submitted that the qualitative analysis of one text and its translations – the hallmark of standard exploratory *Translation Mining* – is a good starting point for exploratory research. The discussions on the role of compositionality in Section 3 show the complexity of natural language data, leading to surprising interactions even in the fairly well-studied domain of tense and aspect in Romance and Germanic. This complexity led us to embrace more quantitative methods. Crucially, though, we used these methods to model the patterns emerging from the data rather than as exploratory tools. In this sense, *Translation Mining* is a truly bottom-up approach that complements the more top-down strategy that is advocated in mainstream cross-linguistic corpus research (see the procedure in (1)).

References

Azpiazu, Susana (2019). *La composicionalidad temporal del perfecto compuesto en español. Estudio sincrónico y dialectal.* Berlin: De Gruyter.
https://doi.org/10.1515/9783110633658

Dahl, Östen & Velupillai, Viveka (2013). The perfect. In Dryer, Matthew & Haspelmath, Martin (eds.), *the world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
https://wals.info/chapter/68.

De Sutter, Gert, & Lefer, Marie-Aude (2020). On the need for a new research agenda for corpus-based translation studies: A multi-methodological, multifactorial and interdisciplinary approach. *Perspectives*, *28*(1), 1-23.
https://doi.org/10.1080/0907676X.2019.1611891

Ebeling, Signe Oksefjell (2016). Does corpus size matter? Revisiting ENPC case studies with an extended version of the corpus. *Nordic Journal of English Studies*, *15*(3), 33-54.

Enghels, Renata, Defrancq, Bart, & Jansegers, Marlies (2020). Reflections on the use of data and methods in contrastive linguistics. In Renata Enghels, Bart Defrancq & Marlies Jansegers (eds), *New Approaches to Contrastive Linguistics: Empirical and Methodological Challenges*, 2-20. Berlin, Boston: De Gruyter Mouton.
https://doi.org/10.1515/9783110682588-001

Fuchs, Martín & Gonzalez, Paz (2022). Perfect-Perfective variation across Spanish dialects: a parallel corpus study. *Languages* 7(3).
https://doi.org/10.3390/languages7030166

Fuchs, Martín & Klis, Martijn van der (2022). Crosslinguistic differences on the Present Perfect Puzzle: an experimental approach. *Proceedings of Experiments in Linguistic Meaning* (*ELM) 2.*

Fuchs, Martín, Le Bruyn, Bert, Klis, Martijn van der, Tellings, Jos, & Swart, Henriëtte de (in preparation). Replication and Triangulation in Parallel Corpus Research [working title].

Granger, Sylviane, & Lefer, Marie-Aude (2020). Introduction: A two-pronged approach to corpus-based crosslinguistic studies. *Languages in Contrast*, *20*(2), 167-183.
https://doi.org/10.1075/lic.00014.int

Harris, Martin B. (1982). The 'past simple' and the 'present perfect' in Romance. In Nigel Vincent & Martin B. Harris (eds), *Studies in the Romance Verb*, 42-70. London: Croon Helm.

Hasselgård, Hilde (2020). Corpus-based contrastive studies: Beginnings, developments and directions. *Languages in Contrast*, *20*(2), 184-208.
https://doi.org/10.1075/lic.00015.has

Hennig, Mathilde (2000). *Tempus und Temporalität in geschriebenen und gesprochenen Texten*. Tübingen: Niemeyer.
https://doi.org/10.1515/9783110929911

Howe, Lewis (2013). *The Spanish perfects: pathways of emergent meaning*, London: Palgrave McMillan.
https://doi.org/10.1057/9781137029812

Klis, Martijn van der, Le Bruyn, Bert, & Swart, Henriëtte de (2022). A multilingual corpus study of the competition between past and perfect in narrative discourse. *Journal of Linguistics* 58(2): 423-457.
https://doi.org/10.1017/S0022226721000244

Klis, Martijn van der, & Tellings, Jos (2022). Generating semantic maps through multidimensional scaling: linguistic applications and theory. *Corpus Linguistics and Linguistic Theory*, *18(3)*: 627-665.
https://doi.org/10.1515/cllt-2021-0018

Le Bruyn, Bert, Fuchs, Martín, Klis, Martijn van der, Liu, Jianan, Mo, Chou, Tellings, Jos, & Swart, Henriëtte de (2022). Parallel corpus research and target language representativeness: the contrastive, typological and *Translation Mining* traditions. *Languages* 7(3).
https://doi.org/10.3390/languages7030176

Le Bruyn, Bert, Klis, Martijn van der, & Swart, Henriëtte de (2023). Variation and stability: the *HAVE*-PERFECT and the tense-aspect grammar of western European languages. In Frank Brisard, Astrid De Wit, Carol Madden, Michael Meeuwis & Adeline Patard (eds), *Beyond Aspectual Semantics: Explorations in the pragmatic and cognitive realms of aspect*, edited volume to appear with Oxford University Press.

Musan, Renate (2001). The present perfect in German: outline of its semantic composition. *Natural Language & Linguistic Theory*, *19*, 355-401.
https://www.jstor.org/stable/4047972

Portner, Paul (2003). The (temporal) semantics and (modal) pragmatics of the perfect. *Linguistics and Philosophy*, *26*, 459–510.
https://doi.org/10.1023/A:1024697112760

Sönning, Lukas, & Werner, Valentin (2021). The replication crisis, scientific revolutions, and linguistics. *Linguistics*, *59*(5), 1179-1206.
https://doi.org/10.1515/ling-2019-0045

Squartini, Mario, & Bertinetto, Pier Marco (2000). The simple and compound past in Romance languages. In Dahl Östen (ed), *Tense and Aspect in the languages of Europe*, 403–440. Berlin: Mouton de Gruyter.
https://doi.org/10.1515/9783110197099.3.403

Swart, Henriëtte de, Le Bruyn, Bert, Grisot, Cristina, & Xiqués, Teresa M. (2022). Perfect variations in Romance. *Isogloss*, *8(1)*.

Tagliamonte, Sali A., & Baayen, R. Harald (2012). Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language variation and change*, *24*(2), 135-178. https://doi.org/10.1017/S0954394512000129

Valente, Chiara (2021). El reparto tempo-aspectual de los pretéritos perfectos simple y compuesto en italiano: un estudio empírico para delinear el estado de la cuestión, In Carmen Quijada van Den Berghe & Jose J. Gómez Asencio (eds), *Los pretéritos perfectos simple y compuesto en Español Peninsular y en otras lenguas.* Madrid: Arco Libros, 183-214.

Appendix A

English (original): Rowling, J.K. (1997). *Harry Potter and the Philosopher's Stone.* London: Bloomsbury.

Dutch: Rowling, J.K. (1997/1998). *Harry Potter en de Steen der Wijzen.* (Wiebe Buddingh', Trans.). Amsterdam: de Harmonie.

French: Rowling, J.K. (1997/1998). *Harry Potter à l'école des sorciers.* (Jean-François Ménard, Trans.). Paris: Éditions Gallimard Jeunesse.

German: Rowling, J.K. (1997/1998). *Harry Potter und der Stein der Weisen.* (Klaus Fritz, Trans.). Hamburg: Carlsen Verlag.

Spanish: Rowling, J.K. (1997/1999). *Harry Potter y la piedra filosofal.* (Alicia Dellepiane Rawson, Trans.). Barcelona: Ediciones Salamandra.