How to Do Things with Translations

Cross-Lingual Phrase Embeddings for Translation Analysis

Jeff Jacobs

September 1, 2018

Abstract

The tumultuous history of translation – from Reformation-era Biblical translators being burned at the stake to the mass execution of Soviet translators (and destruction of their translations) during the "Great Purge"[1] – tells us that the act of translation has always been subject to acute political pressures. To what extent can we empirically study the effects of these political pressures on produced translations? In this paper we introduce a novel algorithm, the "Translation Mover's Distance", which leverages recent work in Natural Language Processing and Corpus-Based Translation Studies to measure the "semantic distance" between a text and its translation. Using an extension of the political ideology model of Jelveh, Kogut, and Naidu (2017), we provide an interpretation of our algorithm as measuring as the outcome of a utility maximization problem, wherein translators choose a point in "semantic space" as close to their own political-ideological ideal point as possible subject to external political pressures (including the pressure to have the translation semantically adhere to the original). To illustrate the algorithm and model, we perform a study of semantic deviations in two English translations of Karl Marx's *Capital* which were conducted under different levels of state repression. Utilizing a Cambridge School contextual historiographic approach, we examine what the translators were aiming to *do* with their respective translational speech-acts, in light of the ideological and historical dynamics of the Cold War. We conclude with a proposal for a larger multilingual study of the global dissemination of *Capital*.

1 Introduction

The oft-recited maxim *traduttore, traditore* – a play on the phonetic similarity of the Italian words for "translator" and "traitor" (López 2008) – encapsulates the tense relationship between translation and politics. The earliest Biblical translation, the 3rd-century BCE

[1] See Freedman (2016) and Rauhala (2017), respectively.

Greek-language "Septuagint", catalyzed the influence of Abrahamic thought on the Mediterranean world and sparked some of the earliest religious schisms, and was itself part of the broader political program of the Ptolemaic-Egyptian king Ptolemy II Philadelphus (Law 2013). Ottoman-era "dragomen", translators between Arabic, Persian, Turkish, and European languages, were explicitly viewed as diplomatic actors and were held in high esteem due to their ability to navigate the hazardous terrain of formal address, deference, and subtext (Ghobrial 2013, Ch. 4; Lewis 2004, Ch. 2). The decontextualized "We will bury you!" translation of Krushchev's 1956 remarks to Western ambassadors in Poland stoked the flames of Cold War mania, landing on the cover of *Time* magazine[2] and necessitating a toned-down clarification by Krushchev himself[3]. In all these cases we can glimpse the broader lesson: that translation has always been an inherently political and context-sensitive speech act. As renowned Machine Translation pioneer Martin Kay once put it, "In order to understand a sentence, your knowledge of linguistics is a relatively minor matter. Your knowledge of the world is what's important"[4].

In this paper we aim to take this notion seriously, arguing that the act of translation itself should be studied as a political "speech act" (Austin 1955) by scholars of intellectual history, History of Political Thought, and Contextual Political Theory, and ideally by anyone working with texts in translation more generally. Just as "Cambridge School" scholars of HPT have revolutionized our understanding of political texts by studying the broader politico-linguistic constraints acting upon their authors (Skinner 1969), we propose to bring a similar context-sensitive analysis to the study of translations as political acts. In a world where the vast majority of political texts are read more often in translation than in their original language[5],

[2] `http://www.bbc.com/culture/story/20150202-the-greatest-mistranslations-ever`. See also Birkby (2014).

[3] Krushchev's clarification ("Of course we will not bury you with a shovel. Your own working class will bury you") made it clear that his utterance was a reference to the *Communist Manifesto*'s famous "What the bourgeoisie therefore produces, above all, are its own grave-diggers" line (Marx and Engels 1988, 222).

[4] Quoted in `https://www.wired.com/2000/05/translation-2/`

[5] Even within the major English-speaking countries (US, UK, Canada, Australia, and New Zealand),

such a study can help shed light on the ideological transformations which occur when these texts are brought across cultures. To this end, our contribution is four-fold: (1) surveying the history of translation within and across socio-political contexts, (2) arguing that political pressures inevitably affect translation choices, and that the inevitable changes in meaning induced by these choices can themselves have important political consequences, (3) providing a tool with which scholars can explore and interrogate such changes, and (4) providing a case study of how the tool can be used to highlight notable shifts in meaning in a translation of Karl Marx's *Das Kapital*. Importantly, the tool developed in (3) will be made available as a web application (the "Translation Explorer") for researchers studying texts in translation[6], and its source code will be made available on GitHub to allow modification and extension[7]. The cross-lingual sentence and phrase alignments generated for the public-domain English translation of *Das Kapital* used in our case study[8] will also be provided in the project's GitHub.

The paper proceeds as follows. Section 2 motivates the aims of the study via a history of the politics of translation through a "Cambridge School" lens, including a survey of the tumultuous translation history of the paper's case study, Karl Marx's *Das Kapital*. Section 3 formalizes the insights from this historical survey via a mathematical model of translation choices made under political pressure. Moving into the computational portion of the study, Section 4 surveys prior work in Corpus-Based Translation Studies, a literature at the intersection of Natural Language Processing (NLP), corpus linguistics, and translation studies, noting that no study to this point has utilized word embeddings to study translation.

6 of the top 10 most-assigned texts in politics courses (and 9 of the top 10 in philosophy courses) are translations from Ancient Greek, French, Italian, or German, according to the Open Syllabus Project (`http://opensyllabusproject.org/`). Given that the vast majority of the world speaks none of these languages natively, and that less than 5% of translations produced today are from non-European languages (Bellos 2011), it stands to reason that the situation is even more extreme outside of the anglophone world.

[6] `http://textlab.econ.columbia.edu/translation`

[7] `https://github.com/jjacobs3/translation-embeddings`

[8] The Progress Publishers translation (Marx 1965) is in the public domain. The Fowkes translation (Marx 1976), is under copyright and thus we cannot provide the alignments for this work.

Section 5 continues with an introduction to the specific NLP methods utilized in the study: word embeddings, text alignment, and the "Word Mover's Distance" algorithm. Section 6 describes our "Translation Mover's Distance" algorithm, emphasizing the modifications made to the Word Mover's Distance which allow it to capture the specific textual features salient to translation analysis discussed in Section 4. Section 7 illustrates a use case of the algorithm on two English translations of *Das Kapital*, extending the "Cambridge School" analysis of Section 2 to pressures on these Cold War-era translations. Section 8 concludes and presents a proposal for a broader multilingual study of the dissemination of *Das Kapital* utilizing the Translation Mover's Distance.

2 Translations as Speech Acts

"If you were ever caught, you knew you would be killed in some savage way, because you would be considered a traitor", reads a recent testimonial from an Iraqi software engineer enlisted as an Arabic-English translator for the US military during its invasion of Iraq (Kukis 2011, 132). As this example shows, real-world instances of the seemingly straightforward task of translation are often perilously entwined within the broader dramas of politics. As scholars of intellectual history, therefore, it is important to move away from the notion that a translation is simply a type of "recoded" facsimile of a source, and towards an understanding of the political dynamics of translation itself. Such an approach will not only allow us to analyze important aspects of a translated text which simple linguistic analysis would not reveal, but also the ways in which political pressures bearing on the translation impacted the subsequent diffusion of the text's ideas into the translation's audience.

As mentioned in the Introduction, the earliest known act of translation – the "Septuagint" – was conducted as part of Ptolemaic-Egyptian king Ptolemy II Philadelphus' attempt to assimilate Jews (and other socio-cultural and religious groups) into his empire. It would be

harshly criticized by contemporaries and would go on to have significant and often violent cultural and political implications for centuries to come. Hence an influential work of Jewish history asserts that "[t]he day on which the seventy-two Sages completed the Greek translation of the Torah – the eighth of Teves – was a day of sorrow [...] as calamitous as the day on which the golden calf had been made" (Kitov 1970, 302). Nearly 2,000 years later, William Tyndale was burned at the stake for his "heretical" 1526 translation of the Bible, a translation which went on to serve as the primary basis for the predominant English version in use today, the King James Bible[9]. The entire Protestant Reformation can, in a sense, be understood as a struggle over whether the Bible should be restricted to Latin and read only by trained priests, or translated into vernaculars and read by laypeople. Controversies over the impact of translation have not subsided in the intervening centuries, nor are they exclusive to the Jewish or Christian religious traditions: in recent years, Qur'anic translators in Afghanistan have received harsh prison sentences for failing to retain the contents of the Arabic source[10]. In every case, the harsh responses came not because of some aesthetic opposition or aversion to translation as such, but because it was recognized that these translations were *doing something* in the social world.

It is not only in the sphere of religion where translation is fraught with political implications, however. In media reports of international affairs it is often the case that a single "headline-grabbing" translation choice, picked from among the myriad choices always available to a translator, can have severe global political implications. Moreover, by examining these cases, we can *directly* observe the political impacts of translations, whereas the examples in the previous paragraph arguably impacted politics only *indirectly*, by way of their effects on religious controversies. In addition to the "We will bury you!" example mentioned in the

[9]Nielson and Skouson (2015) estimate that 84% of Tyndale's words were retained in the King James' New Testament, and 76% in its Old Testament.

[10]`http://www.nbcnews.com/id/29049101/ns/world_news-south_and_central_asia/t/afghans-face-death-over-translation-quran/`

Introduction, a 2005 translation of a speech by Iranian President Mahmoud Ahmadinejad is instructive. While a number of studies by Farsi scholars have since argued via collocational analysis that "The regime occupying Jerusalem must fall!" would be a more faithful translation (Zanettin 2016; Wood 2010), the quote was reported in Western media as "Israel should be wiped off the map!" or "Israel should be wiped off the face of the earth!"[11]. This provocative "soundbite" led UN Secretary General Kofi Annan to cancel his planned trip to Iran and former/future Israeli Prime Minister Benjamin Netanyahu to pursue international prosecution against Ahmadinejad for incitement to genocide. Once again, we see that the choice of the speech act "Israel should be wiped off the face of the earth!" over "The regime occupying Jerusalem must fall!" – despite the fact that both represent feasible translations (from a linguistic standpoint) of the orignal Farsi – should be analyzed through the lens of political pressures and consequences.

While much ink has been spilled and will continue to be spilled over this and countless other translational controversies, for our purposes the key takeaway is the necessity for scholars of politics to study translation as a central facet of the practice of politics more generally. Turning to the subject of our case study, Karl Marx's *Das Kapital* Vol. I has a tumultuous translation history, perhaps second only to that of the Bible in terms of the dangers faced by its translators. First published in German in 1867, this tumult began almost immediately when the first translation was commissioned in 1868. This translation was not to be into the English of Marx's London, nor into the French of Paris where Marx was first radicalized, but into Russian, the language of a society Marx particularly disdained[12].

The task fell to none other than Mikhail Bakunin, the Russian anarchist who would

[11]For example, *The New York Times* stated that "Iran's conservative new president, Mahmoud Admadinejad, said Wednesday that Israel must be 'wiped off the map' and that attacks by Palestinians would destroy it" (Fathi 2005).

[12]In Edmund Wilson's words, Marx "growlingly put it down to the fact that the Russians 'always run after the most extreme ideas that the West has to offer': it was pure intellectual 'gourmandise'" (Wilson 1940, 348).

very soon become Marx's arch-rival in the First International – a feud so extreme it lead to the breakup of the organization into "Marxist" and "Bakuninist" camps (Eckhardt 2016). Bakunin loathed the book, subsequently writing to a friend: "You know I've been commissioned to translate that awful book by Marx – *Capital*. 784 pages of small print for 900 roubles" (*ibid.*, 325), and abandoned the project in 1870[13].

The translation finally saw the light of day in 1872, having been completed by a Russian member of the "Marxist" camp, Nikolai Danielson. Despite having banned all previous translations of Marx's writings, censors allowed Bakunin and Danielson's on the basis that it was a "colossal mass of abstruse, somewhat obscure politico-economic argumentation", concluding that "in Russia few will read it and even fewer will understand it" (Resis 1970). It was this translation that Vladimir Lenin would read in the decades leading up to the Bolshevik Revolution, and that would become the "official" translation read in the Soviet Union[14].

By 1972, translations into 22 languages of the USSR had been performed by Moscow's Progress Publishers under the supervision of Soviet officials, with varying degrees of censorship and repercussion for translators (Vygodsky 1979). The first Soviet Finnish translation of 1933, for example, was deemed seditious enough to warrant the execution of all involved between 1937 and 1938 and the destruction of all copies, as part of the larger destruction of over 1.5 million Finnish texts over this period (Rauhala 2017). As was the case centuries earlier during the Protestant Reformation, Soviet officials recognized and feared the potentially radical political effects an "inconvenient" translation could have on the status quo.

Moving to the world of English, Amini (2016) documents how, despite numerous partial translations being performed over the years, only two full English translations remain

[13]Marx and Engels would go on to conduct a harassment campaign against Bakunin demanding the return of his 300 rouble translation advance, an ordeal lasting several years (Eckhardt 2016, 323–330).

[14]A newer translation by Bolshevik founder Alexander Bogdanov was completed in 1909, but suppressed in the USSR on account of his opposition to the Bolsheviks, after being expelled the very year he completed the translation (Susiluoto 1982, 43).

widely-read today: the first English translation of 1887 by Edward Aveling and Samuel Moore (overseen by Engels himself shortly before his death) and an additional translation performed nearly a century later in 1976 by Ben Fowkes. Of note here is the fact that Progress Publishers, the Moscow-based publishing house which produced English-language material for Western consumption beginning in 1931, opted to publish a "corrected" version of the Aveling and Moore translation in 1965. In a sense, then, we can interpret the Progress Publishers edition as the translation most representative of the "official" Marxist doctrine of the Soviet Union, with the Fowkes translation (never published in the USSR) best representing popular non-Soviet Marxist doctrines like Trotskyism[15]. In Section 7 we thus employ our "Translation Mover's Distance" to probe the ways in which these English translations "moved" the semantics of the German original.

3 A Model of Political Translation

Having now analyzed several qualitative examples of how political pressures come to bear on translation, in this section we develop a set of simple mathematical models capturing the essential political dynamics behind the choices made by translators. The logic of the model, an extension of the political ideology model in Jelveh, Kogut, and Naidu (2017), is as follows: the primary agent, a translator T, is tasked by a principal P with producing a translation to be read by people within some broader society S, with both the principal and the society able to inflict costs upon (or, equivalently, provide rewards to) T. Here P could be interpreted as the state, who can (for example) censor it and penalize the translator, while S could be interpreted as "the public" who can (for example) reject it as a "bad" translation and not buy it. Importantly, the society S here acts as a "language-maker", in the sense that they

[15]In a similar vein, the first Chinese translation of *Capital*, published in the 1930s and widely read and discussed by communist students (Yoshihiro 2012 reports that the 20,000 copies of the first printing immediately sold out; see also Chiang 2001), can perhaps be interpreted as closer to the Chinese Communist Party's/Mao's Marxist doctrine.

are the general users of language who thus establish the rules (what Saussure (1916) termed the *langue*) of the "linguistic space" within which the translator writes[16].

3.1 The Translation Game

An act of translation proceeds in four stages:

1. The society S constructs the "rules" (*langue*) of the language, by establishing a shared semantic space representing (again following Saussure (1916)) the structure of interrelationships between signifiers within and between two languages X and Y. In addition to the pairwise distances between each signifier, they also establish a "feasible translation range" for each word in $w \in X$, encoding the set of terms that this public will find "reasonable" as translations of w.

2. The translator T encounters a linguistic utterance x (*e.g.*, a word, phrase, sentence, paragraph, chapter, or document – for the remainder of the paper we interpret it as a word) in the source language X and is tasked with choosing an utterance y in the target language Y to serve as the translation of x.

3. The principal P observes T's translation choice y and decides whether or not to punish the translator, paying a punishment cost p for each unit of punishment (negative utility) inflicted upon the translator.

4. The society, finally, observes T's translation choice y and inflicts an "interpretability" cost on T, which is 0 if y is within the "feasible translation range" created in step 1 but grows quadratically in the semantic distance otherwise.

The society creates the "rules" by establishing a multilingual "semantic distance" metric $D : (X \cup Y)^2 \to \mathbb{R}$, observable by all agents. This metric encodes their understanding that,

[16]Thus, the translator creates Saussure's complement to the abstract *langue*, namely, the concrete *parole*.

for example, the English "dog" is closer to the German "Hund" than it is to "Fisch", by taking values such that $D(dog_{EN}, Hund_{DE}) < S(dog_{EN}, Fisch_{DE})$. Note that this does not preclude the possibility that two separate words can be seen as "equally similar": if "einfach" and "simpel" are equally valid translations of the English "simple" in the eyes of the audience, then $D(simple_{EN}, einfach_{DE}) = D(simple_{EN}, simpel_{DE})$.

The role of the society in affecting the translation choice is simply that they choose a maximum semantic distance $m(x), m : X \to \mathbb{R}$ that they are willing to "tolerate" in a translation of x, beyond which they impose a quadratic penalty on the translator for making a "nonsensical" or "fraudulent" translation choice. In other words, they allow the translator leeway in selecting y from a *range* of terms feasibly "similar" to x, but require some basic level of adherence to the semantics of the original. For intuition, one can imagine small values of m for more "concrete" source-language words like "cat" or "coffee" and higher values for abstract (and therefore harder-to-translate) terms like "essence" or "value". The cost function C_S imposed on T by the society is thus

$$C_S(y) = \begin{cases} 0 & \text{if } D(x,y) \leq m(x), \\ (D(x,y) - m(x))^2 & \text{otherwise.} \end{cases}$$

Importantly, this constraint encapsulates what makes *translation* different from a "standard" speech act, namely, the need to choose words from a set "anchored" in the semantics of the original rather than from anywhere in the language. This is not meant to imply that non-translational speech acts are "unconstrained" – indeed, as Pocock (1985) convincingly argues, the range of linguistic innovations a given author is able to employ in service of a political argument is *always* constrained by the broader linguistic conventions of the era and society they're writing in. Rather, the difference in the case of translation is that these constraints on linguistic innovation exist *relative to* a particular utterance from the source language.

Unlike the society which specifies a range of feasible words about which they are equivocal, the translator and the principal have singular *ideal points* $\widehat{y_T}, \widehat{y_P} \in Y$ (respectively), translations which they prefer over all others. These ideal points are what "encode" the political preferences of the two agents. The principal wants the translation to be as close to $\widehat{y_P}$ as possible, while the translator wants it to be as close to $\widehat{y_T}$ as possible. In particular, agent i's utility from a translation choice y decreases proportional to the square of the distance between y and their ideal point: $S(\widehat{y_i}, y)^2$. Because punishing the translator is not costless (for example, because frequent crackdowns lower the regime's legitimacy or due to the cost of tracking down and destroying copies), the principal pays a cost c if they decide to punish, which inflicts n units of disutility on the translator. Their decision thus boils down to maxing the utility-maximizing punishment choice:

$$p^*(y_T) = \underset{p \in 0,1}{\operatorname{argmax}} U_P(p|\widehat{y_P}) = -(y_T - \widehat{y_P})^2 - pc,$$

where pc represents the cost incurred by the principal based on their punishment decision.

The translator, therefore, trades off the positive utility gained from a translation choice close to their ideal point with the negative utility they may experience due to (a) a translation choice far from the principal's ideal point and (b) a translation choice lying outside of the society's "feasible translation range". They choose y_T so as to solve:

$$y_T^*(p) = \underset{y_T \in Y}{\operatorname{argmax}} U_T(y_T|\widehat{y_T}) = -(y_T - \widehat{y_T})^2 - np - C_S(y_T),$$

where np is the cost the translator will incur due to the repression of the principal and $C_A(y_T)$ is the social cost for a "nonsensical" or "fraudulent" translation as discussed above.

However, because they cannot observe the principal's punishment choice in advance (since it's made in *response* to the translation decision), when actually making their choice of y_T the translator treats the principal's decision variable as a random variable and aims to maximize

expected utility:

$$y_T^* = \underset{y_T \in Y}{\operatorname{argmax}}\, \mathbb{E}_p[U_T(y_T|\widehat{y_T})] = -(y_T - \widehat{y_T})^2 - n\mathbb{E}[p] - C_S(y_T).$$

Note that the translator has full knowledge of the "socially-acceptable" range of translation distances and thus knows the value of $C_S(y_T)$ for any value of y_T, so C_S is not a random variable.

3.2 Interpretation

It is instructive to examine "reduced versions" of the model here. First, consider the case where the society chooses an infinite feasible translation range for every word ($m(x) = \infty \ \forall x \in X$), *i.e.*, there is no social pressure whatsoever for the translation to adhere to the original, and where the principal never punishes ($p = \infty$). In this case the model would represent the situation of *non-translational* private writing (perhaps in a personal diary): writers could choose any word in the entire semantic space without regard to the original text, and would always choose their ideal point $\widehat{y_T}$. In game-theoretic parlance, they always "truthfully report" their ideal point, and there are no social or political pressures affecting their writing.

If we introduce *just* the ability for the principal P to punish, we obtain the case of *non-translational* public writing (a book, newspaper article, etc.) with political pressure: an "original text" again plays no role, but the writer now has to strike a balance between their ideal point and (their perception of) the principal's ideal point, resulting in a choice which is some convex combination of the two ($y_T^* = \lambda\widehat{y_T} + (1-\lambda)\widehat{y_P}$). All else held constant, the more harsh the principal's punishments are (the higher the value of n is) the more the resulting translation will be closer to the principal's ideal point on the line connecting the two. In a situation of extreme repression, say if $n = \infty$ where the translator is killed instantly for a

translation not to the principal's liking, the translator will always choose (their perception of) the principal's ideal point. This reduced model could be interesting on its own for a study of political *writing*, but such an analysis lies outside the scope of our study.

Alternatively, if we introduce *just* the ability for the society S to punish, we obtain the case of private translation (perhaps translating a book for one's own understanding without publishing it) without political pressure: now the translator cannot choose any point within semantic space (for example, they may not be able to choose their ideal point), but instead must restrict their choice to points *within* the feasible translation range. The solution, therefore, will be for them to choose the point within this range closest to their ideal point. As in a personal non-censorable translation (a translation without external political pressures), the translator simply chooses their favorite term out of all terms which reasonably comport to the meaning of the original.

Having covered these three reduced cases, the dynamics captured by the full model become clear: the society's ability to punish introduces the need for the translator to choose terms such that the audience will understand it as a translation with respect to the "linguistic rules" (thus making it a translational *speech act*). The principal's ability to punish introduces the need for the translator to "tread carefully", moving their choices away from their own ideal point and towards the principal's out of fear of punishment (thus making it a *political* speech act).

Foreshadowing what is to come, a "word embedding space" is a widely-used computational method which will operationalize the notion of the "linguistic rules", or the *langue*, discussed in this section. By scanning over a large corpus of text in a given set of languages from a given set of time periods, multilingual word embedding methods allow construction of a "semantic space" capturing the ways in which words are used together and which translations are perceived as feasible. The "Translation Mover's Distance" we develop herein utilizes these semantic spaces to measure the distance that a linguistic utterance "travels" between

the original and its translation, or between two translations of the same original. Importantly, although "word" was used for simplicity in this section, one of the major strengths of the Translation Mover's Distance is its ability to aggregate or disaggregate linguistic units – *e.g.*, the distance computed between a paragraph and its translation can be decomposed into distances between its constituent sentences, phrases, or words. The next section moves us into the computational side of the study by introducing concepts from Corpus-Based Translation Studies which inform the construction of the Translation Mover's Distance algorithm.

4 Prior Work: Corpus-Based Translation Studies

To the best of our knowledge, no study to date has attempted to utilize word embeddings (and their associated distance metrics) to study translations, much less from a political-theoretic perspective[17]. Nonetheless, our study can be seen as a natural extension to the field of "Corpus-Based Translation Studies", which has fruitfully employed several "pre-neural" (see Section 5 for an explanation of the Neural Network underlying the word embeddings approach) NLP algorithms to uncover linguistic "laws" of translation. Baker (1993) proposes a dichotomy which has become standard in the field, between those laws which apply regardless of the target or source language ("universals") and those which characterize translations between a particular source-target language pair. Within the Corpus-Based Translation Studies literature which has emerged from Baker's ontology, the three most commonly-studied translation universals are:

- *Simplification*: Across numerous measures of linguistic complexity (lexical variety, average word/sentence length, etc.), researchers have found that translated text is consistently less complex than the corresponding source text. Much of Laviosa (2002), for

[17]A robust literature on qualitative studies of translation and politics, outside the scope of this study, does exist in anthropology and comparative literature. For a survey see Gal (2015) and for a qualitative work which specifically analyzes *Capital* see Liu (1999), Ch. 1.

example, is devoted to quantifying this effect across different language pairs via large translation corpora.

- *Explicitation*: Implicit references or implications made in passing in the source text are often "spelled out" explicitly in the translated version. In a canonical example from Baker (1993), the short sentence "The example of Truman was always present in my mind." is transformed into an entire paragraph explaining the reference in its Arabic translation: "In my mind there was always the example of the American President Harry Truman, who succeeded Franklin Roosevelt towards the end of World War II. At that time – and after Roosevelt – Truman seemed a rather nondescript and unknown character", and so on (243).

- *Standardization*: The least straightforward of the three universals, but involves a reduction in idiosyncratic language (say, replacing several uncommon idioms with the same common idiom) and use of contractions, resulting in more "formal"-sounding language. Olohan (2003), for example, finds a 76% not-contraction rate in the British National Corpus (which purposefully excludes texts in translation) compared to 52% in the Translational English Corpus.

As will be explained in Section 6, the "Translation Mover's Distance" employs a set of machine-learned textual features in order to capture the strength of these three effects within a particular instance of translation. In addition to these universal language-agnostic "laws", studies of regularities in translations between *particular* language pairs have posited various laws of "linguistic interference" (Toury 1982), whereby "fingerprints" of the source language remain evident in corpora of translated texts. The most basic examples of this phenomenon are loanwords, occurring when translators simply "carry over" a word in its original form from the source language and such transfers "stick" in the target language, with examples abound in English from "rendezvous" (French) to "ziggurat" (Akkadian). A less obvious example

is that of "cognate form" interference when translating English into Spanish. Experiments performed in Tercedor (2010), for example, find that the English bigram "industry sector" is translated into its "cognate form" – *sector industrial* – 50% more often than into the "more conceptually appropriate" *sector de la industria*, a finding which is robust across English-Spanish translation studies.

This "translational interference" can also work in the converse direction, in a sense (with the target language "interfering" with attempts at translation from various source languages). For example, Baroni and Bernardini (2006) are able to detect whether articles in an Italian geopolitics journal are translations or original Italian writings with 86.7% accuracy, 89.3% precision, and 83.3% recall[18], with these numbers dropping precipitously when non-clitic pronouns are removed from the feature set, a result *not* found when analyzing other languages in a similar manner. This result provides strong evidence for a "reverse interference" effect, whereby non-clitic pronouns are in a sense "hard" to translate into Italian in a "natural" manner. It is precisely this notion of "natural" that we aim to capture with a word embedding space – in essence, an encoding of "natural" language from a particular era which can thus be leveraged to detect out-of-the-ordinary linguistic innovations introduced by a given translation act. We thus introduce word embeddings, along with the other NLP methods underlying the Translation Mover's Distance, in the next section.

[18]Precision and recall measures are important supplements to accuracy which incorporate false positives and false negatives, respectively, when evaluating performance on classification tasks. Since a vast majority of the articles were in fact translations, a trivial classifier which always guessed "translation" would achieve high accuracy, for example, but low recall.

5 Prior Work: Natural Language Processing

5.1 Word Embeddings

In simplest possible terms, a word embedding is a numerical representation of the meaning of a word. This begs the question, however, of how signifiers come to have meaning in the first place – a question which far predates both word embeddings and computers. Although the sign-signifier-signified terminology comes from Saussure's 1916 *Course in General Linguistics*, the story of word embeddings begins in the 1950s. Summarizing the trajectory of pragmatics research to that point, John Firth (1957) famously asserted that "You shall know a word by the company it keeps", a proposition subsequently termed the "Firthian hypothesis". This hypothesis captures the following linguistic-cognitive logic: if one encounters a new word like "wampimuk" and doesn't know what it means, but observes that it was used in a sentence like "We found a little hairy wampimuk sleeping behind the tree" (McDonald and Ramscar 2001), they can use the words around it to make inferences regarding the meaning of "wampimuk" (that it's a mammal, that it's small enough to sleep "behind" a tree, etc.). Fifty years later, with the widespread availability of massive natural language corpora, the first word embedding algorithms put Firth's theory into practice on a large scale.

The most "naïve" word embedding method, but the method which serves as the basis for all others, is the construction of a co-occurrence matrix from a given text corpus. Given a "radius" r (or, equivalently, a window size $k = 2r$) and a corpus with v unique words, the co-occurrence matrix M is a $v \times v$ symmetric matrix where each element $M_{i,j}$ represents the number of times word j appears within the r words before and after word i in the corpus. With this matrix in hand the "naïve" embedding for a word is simply the column (or row) of the co-occurrence matrix corresponding to that word. As each element of this vector represents the number of times the word appeared "in the company of" some other word, the Firthian criteria is satisfied. The problem, however, is that modern corpora such

as the Common Crawl corpus[19] often have vocabulary sizes v exceeding 2 million, thus requiring over 4 trillion co-occurrence counts to be computed, stored, and utilized for distance calculations. For this reason, all popular word embedding algorithms – Google's word2vec (Mikolov et al. 2013), Stanford's GloVe (Pennington et al. 2014), Facebook's fastText (Bojanowski et al. 2016), and so on – represent various methods for reducing the dimensionality of this co-occurrence matrix while retaining maximal information about the word-to-word relationships.

In-depth explanation of these methods lies outside of the scope of this work[20]. For our purposes, however, what matters is that the rows or columns of the matrix (whether dimensionality-reduced or not) can be interpreted as points in a geometric space. Importantly, such an interpretation further coheres with the Firthian hypothesis, in that words which "keep similar company" will be closer together in this space. For example, the point corresponding to "meow" will be close to the point corresponding to "cat", as long as "meow" and "cat" frequently appear together in the corpus. As an additional detail, however, words themselves are often not the linguistic unit we are interested in analyzing. Rather, in a social science context we are often interested in phrases like "boots on the ground", "Israeli-Palestinian conflict", and so on. Thankfully, a recently-developed modification of word2vec called ngram2vec (Zhao et al. 2017) enables construction of embeddings for variable-length phrases.

With the plethora of efficient word embedding algorithms mentioned above (word2vec, GloVe, fastText, etc.), the real challenge for our study is not the construction of embedding spaces *as such* but rather the construction of *bilingual* embedding spaces enabling distance computations to be performed between words, phrases, and sentences in separate languages. Thus the Translation Mover's Distance depends crucially on more recently-developed cross-

[19] `http://commoncrawl.org/`

[20] For a comprehensive survey of the most popular embedding algorithms, see Goldberg (2017).

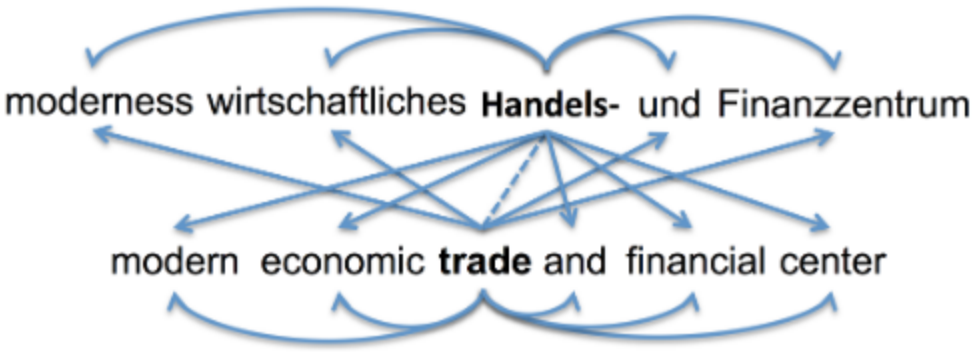

Figure 1: The intuition behind the bilingual embedding method of Luong et al. (2015). In the terminology of Section 5, A = English, B = German, i_A = "trade", and i_B = "Handels-". Figure from *ibid.*, pg. 3.

lingual embedding methods. As the details behind these methods are also outside the scope of this work, the following intuition (illustrated in Figure 1) will suffice: in addition to capturing intra-lingual co-occurrence counts for each word i_A in language A, cross-lingual embeddings utilize *aligned* bilingual translation pairs to ascertain the number of times i_B – the word aligned with i in another language B – co-occurs with each other word $j_B \neq i_B$ in B. This language-B co-occurrence information is then included in i_A's vector, which produces a space like the one illustrated in Figure 2 after being run through a word embedding dimensionality-reduction algorithm. The efficacy of these methods therefore depends crucially on our ability to properly align sentences and words between languages, an issue we turn to next.

5.2 Text Alignment

Though their difficulty may be masked by the ease with which humans are able to match portions of a translation with the original text, the independent tasks of statement and word alignment have been subject to decades of (still ongoing) algorithmic research. Though the inner-workings of state-of-the-art algorithms for these tasks are well outside the scope of this

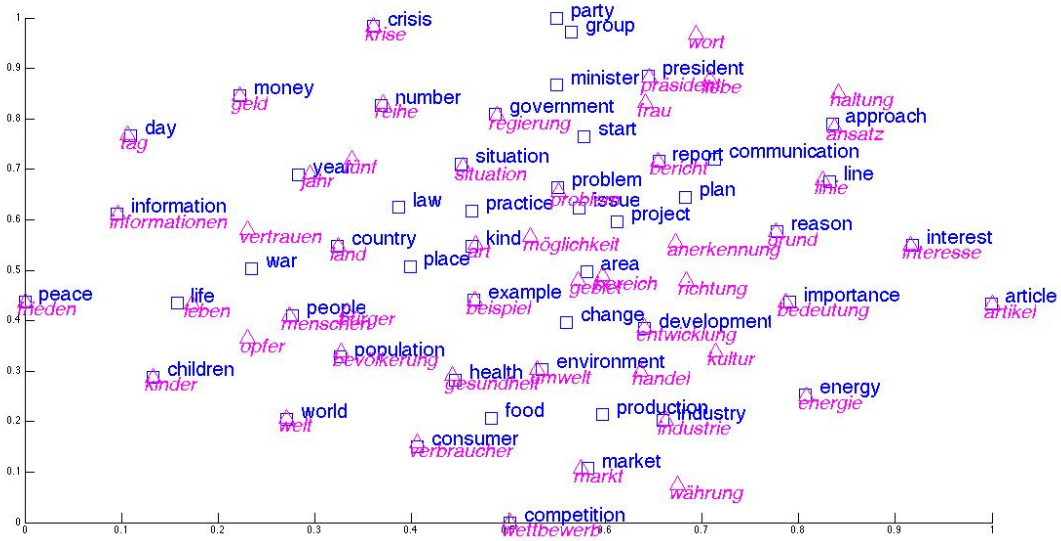

Figure 2: A two-dimensional (t-SNE) projection of the bilingual semantic space produced via the method of Luong et al. (2015). Figure from *ibid.*, pg. 7.

study[21], in this section we attempt to outline the basic problem structure.

The problem of *sentence alignment* is as follows: given a pair of documents, find the between-sentence "cutpoints" in each document such that the induced sentence pairings (*e.g.*, if one "cut" is made in each document, two pairings are induced) best capture the semantic correspondence between statements of the original and translated texts. Although in the easiest case the problem can be solved by placing cuts in between each sentence in both documents, many translations do not exhibit such a simple sentence-to-sentence correspondence. In our *Capital* case study, for example, single sentences in the original are often split into multiple sentences in the translation (or vice-versa, with multiple sentences in the original combined into a single sentence in the translation).

At first glance, the problem of *word alignment* seems similar: given a sentence and its translation, find a mapping between words in the sentence and words in the translation

[21]For technical details on state-of-the-art word alignment methods see Tamura et al. (2014) or Legrand et al. (2016). For a survey of both word and sentence alignment algorithms see Tiedemann (2011).

which best captures the semantic correspondence between the original and the translation. Working at the word level, however, introduces an additional level of complexity: whereas sentence alignment makes the explicit assumptions that (a) each sentence of the original will always correspond to a sentence in the translation and (b) textual ordering is maintained[22], translation at the word level is far too irregular for such simplifying assumptions to be admissible.

Although algorithms which are able to handle violations of both (a) and (b) are outside the scope of this overview[23], we can gain an intuition for word alignment by examining three cases solved by an algorithm able to handle violations of (a) (Brown et al. 1993). In the first case (Figure 3), a "one-to-many" translation, each word in the source is mapped to any number of words in the target. Notably, violations of (a) are handled by allowing words in the source to map to 0 words in the target, as is the case for the first word "And" in the English. In the second case (Figure 4), a "many-to-one" alignment, each word in the *target* is mapped into by any number of words in the source. The most general case (Figure 5), a "many-to-many" alignment, involves pairings where any number of words in the source can be mapped to any number of words in the target. Note that none of these cases violate (b), as can be seen by the fact that none of the lines from source to target ever cross one another. Examining English-to-Korean translation of even a simple sentence like "I see John." into "나는 존을 만난다." (*naneun jon-eul mannanda*, literally "I John see"), however, illustrates the necessity for word alignment algorithms to also handle violations of (b)[24].

[22]That is, if sentence A appears before sentence B in the source text, then the translation of sentence A will also appear before the translation of sentence B.

[23]See footnote 21.

[24]Tomlin (1986) estimates that 45% of languages use the Subject-Object-Verb (SOV) order of Korean, 42% use the Subject-Verb-Object (SVO) order of English, and 13% use an alternative standard ordering, making word order reversals a central challenge for multilingual translation algorithms.

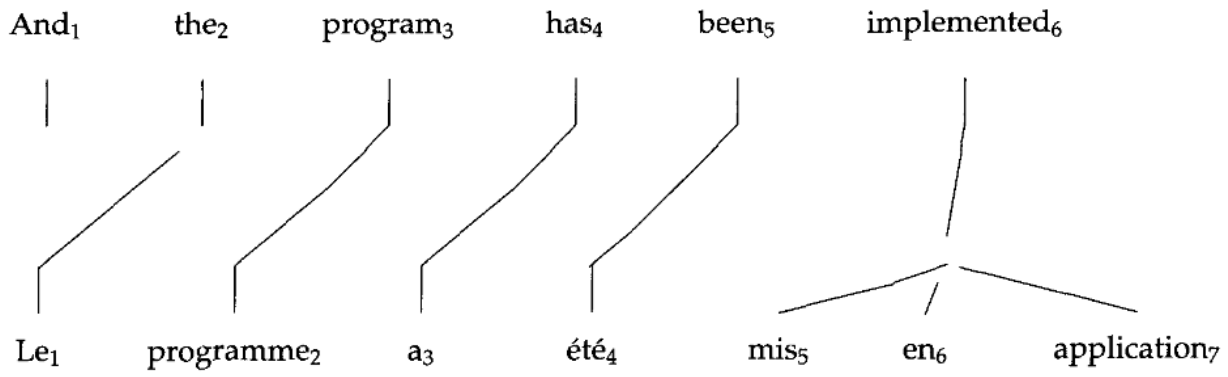

Figure 3: A "one-to-many" English-French word alignment. Figure from Brown et al. (1993).

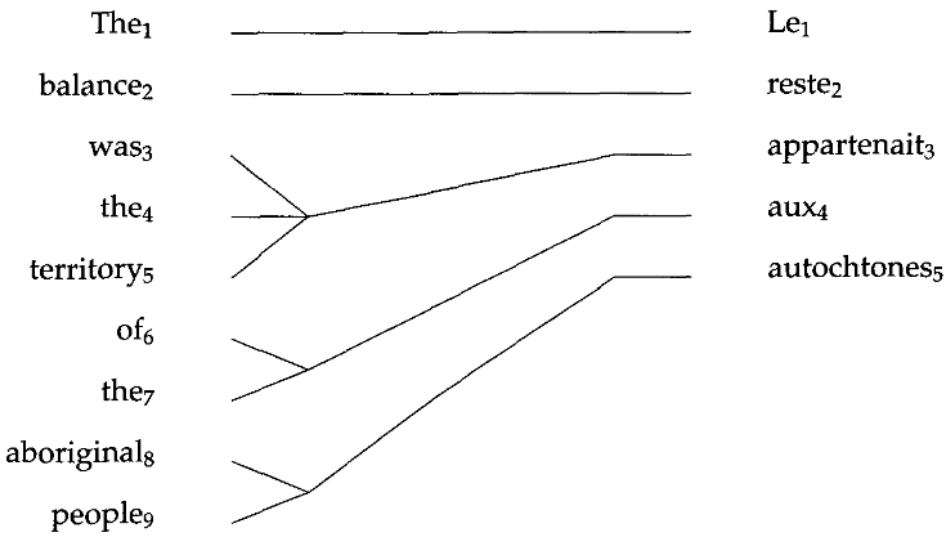

Figure 4: A "many-to-one" English-French word alignment. Figure from Brown et al. (1993).

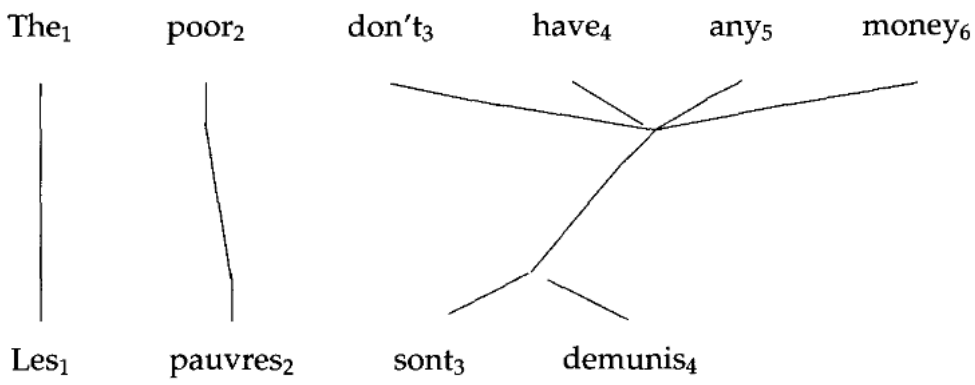

Figure 5: A "many-to-many" English-French word alignment. Figure from Brown et al. (1993).

5.3 The "Word Mover's Distance"

Even with a feasible set of text alignments and a trained word embedding space at hand, a key question remains regarding how best to aggregate individual word- or phrase-pair embedding distances into sentence-, chapter-, or even document-level distances. It is this question which the "Word Mover's Distance" (Kusner et al. 2015), an extension of the statistical "Earth Mover's Distance" (Rubner et al. 2000), provides a simple answer to.

Since we are working with distributions over discrete units (*e.g.*, words), we can gain an intuition for the Earth Mover's Distance through the analogy of rearranging piles of books. If books are piled throughout a room in a particular configuration A, and we want to rearrange the piles to achieve a different configuration B, the Earth Mover's Distance captures the minimal amount of work we would have to perform to accomplish this rearrangement, where "work" is defined by the number of books moved times the distances these moved books had to be carried. Thus two configurations are "similar" if little work was required to transform one into the other, *e.g.*, if only one book was moved from one pile to another, and "distant" otherwise, *e.g.*, if an arrangement with a single gigantic pile on one side of the room was transformed by moving the pile to the opposite side.

The "Word Mover's Distance", then, is an application of this measure to the word embed-

ding spaces described earlier. A sentence or document is an "arrangement" of words within the "room" of all possible sentences or documents that can be formed in a language. For example, if a language L consists of 7 words {Ana, Zainab, said, hello, goodbye, to, photovoltaic}, then the sentence "Zainab said hello to Ana" can be represented as the vector $[1, 1, 1, 1, 0, 1, 0]$, while "Ana said Zainab said hello" can be represented as $[1, 1, 2, 1, 0, 0, 0]$. This is the "bag-of-words" representation of the sentence, where each slot i represents the number of times word i appears in the sentence, analogous to the number of books on a pile in our prior example. Unlike in that example, however, the distance between two piles here is not spatial but instead based on "semantic distance" between the words in each slot. Specifically, this "semantic distance" is the Euclidean distance $d(\mathbf{w}_i, \mathbf{w}_j) = \|\mathbf{w}_i - \mathbf{w}_j\|_2$ between the word embeddings with respect to a pre-trained word embedding space. Since "Ana" will most likely be used in very similar contexts to "Zainab" if the word embeddings are trained on standard contemporary English, moving from the "Ana" pile to the "Zainab" pile will have a low cost, and similarly for "hello" and "goodbye". Moving from "Ana" to "photovoltaic", however, will be high-cost, since these words are not often used in similar contexts. As in the book-moving case, then, the distance between two sentences A and B is the minimum work it would take to "move" the vector for sentence A to the vector for sentence B through this semantic space.

Abstracting away from our examples, the reason the Word Mover's Distance is so helpful for our approach is that it is able to capture the fact that even two sentences without *any* common words can still be extremely semantically similar: consider the case where sentence A is "The Secretary-General criticized Trump's statement" and sentence B is "Guterres lambasted the President's remark". Despite the lack of common terms, the Word Mover's Distance will still capture the nearly identical semantics of A and B, since moving across the semantic "gap" between words with similar meanings – "Secretary-General" to "Guterres", "criticized" to "lambasted", or "statement" to "remark" – will require very little work. Indeed,

as many scholars of translation argue (Bellos 2011), this type of paraphrastic move can essentially be thought of as an "intralingual translation". Adapting this logic to *cross*-lingual translation, then, is precisely the goal of the Translation Mover's Distance, which we turn to in the next section.

6 The "Translation Mover's Distance" Algorithm

We now have enough background to describe the Translation Mover's Distance (TMD), the primary methodological contribution of our study. The TMD aims to adapt the Word Mover's Distance to the study of translations by leveraging insights regarding "translationese" from Corpus-Based Translation Studies (covered in Section 4). It is also, to our knowledge, the first method which incorporates bilingual word embeddings directly into a sentence or document distance measure. In prior work, such algorithms have typically utilized separately-trained *monolingual* word embedding spaces along with a transformation matrix constructed from corpora of pre-aligned translation pairs (see Ruder, Vulić, and Søgaard 2017, Section 6.1.1, for examples). To aid TMD-based analyses of translations, we also develop a visual Translation Explorer interface[25], described herein.

The algorithm consists of the following steps:

1. Before processing of the texts themselves, a cross-lingual jointly-aligned word embedding space is generated via the method of Luong, Pham, and Manning (2015) described in Section 5.1. For the *Capital* case study in the next section, for example, these embeddings were trained on the Europarl v7 English-German corpus (Koehn 2005)[26].

2. Raw text is now loaded and cleaned, chapter by chapter: footnote markers are re-

[25] Available at `http://textlab.econ.columbia.edu/translation`

[26] Bilingual vectors trained on this corpus are downloadable from `https://nlp.stanford.edu/~lmthang/bivec/`.

moved[27] and sentences are tokenized via the appropriate language-specific tokenizers (Kiss and Strunk 2006).

3. The first text alignment stage is performed, using the Hunalign multilingual sentence alignment algorithm (Varga et al. 2005) to produce "statement pairs" linking source and translation.

4. A weighted *statement-level* distance (the details of the weighting are given below) is now computed, thus giving the main *statement-level* Translation Mover's Distance between each aligned statement pair. The Translation Explorer interface allows the user to view these distances along with side-by-side comparisons of the aligned statements from the source text and translation, with the comparisons color-coded so that greater distances correspond to brighter backgrounds. The user can also explore these distance via a plot of statement distances over the course of a chapter.

5. Although it is technically possible to perform the Translation Mover's Distance at the chapter or even entire-document level, such a computation would be prohibitively expensive. Thus chapter- and document-level Translation Mover's Distances are computed simply as the sum of the statement-level distances in the chapter or document. The Translation Explorer presents users with the document distance as well as a plot of the chapter-by-chapter distances.

6. A second alignment stage now occurs, using the Berkeley Aligner multilingual word alignment algorithm (Liang et al. 2006) to produce pairs of linked *words* or *phrases* within each aligned statement pair.

7. With these aligned words/phrases we are now able to compute *phrase-level* Translation Mover's Distances, using the Euclidean distance between each word/phrase pair with

[27] That is, the numbers, letters, or symbols indicating footnotes are removed from the main body of the text. The footnotes themselves are preserved and moved to the end of the text for alignment.

respect to the bilingual embedding space computed in step 1. These distances are made examinable via the Translation Explorer interface, for researchers who want to "disaggregate" the statement-level TMD into distances between phrases within a statement.

As seen in the last section, the Word Mover's Distance algorithm (paired with a word embedding space trained on a sufficiently large corpus) is able to capture "how far" a sentence travels when it is paraphrased into another sentence, *even if* the paraphrase shares no terms in common with the original sentence. The distance computation in step 4 thus leverages this feature to compute *cross*-lingual distances between sentences with *mutually exclusive* tokens. To prevent cross-lingual cognates from being treated as the same token (*e.g.*, to avoid a case where the embedding for the English "gift" is used to represent the German "Gift", meaning "poison"), all words are prepended with a two-letter code in preprocessing (`en_` for English and `de_` for German) representing the language of the text they are drawn from.

As mentioned in step 4, the Translation Mover's Distance further augments the Word Mover's Distance by integrating a set of weighted machine-learned features representing insights from the semantics of "translationese" discussed in Section 4. These features are derived from the features in Wintner's (2017) state-of-the-art translation classifier, and thus we report the classification scores from the corresponding features in *ibid.*, Table 1. Categorized based on the translation effects they aim to capture and ordered by these classification scores, they are:

- *Interference*
  - *Token position* (classification score = .97): Partially countering the information loss resulting from the bag-of-words assumption, this is the absolute difference in the words' *normalized* position within their sentence. The normalization accounts for sentences with different lengths – *e.g.*, if we're computing the distance between

"John" and "존을" (*jon-eul*) in the sentences "I cannot see John." and "나는 존을 만날 수 없다." (*naneun jon-eul mannal su eobsda*, literally "I John see can not"), this token position feature will be $|4/4 - 2/5| = 3/5$.

  - *Parts of speech* (classification score = .90): A dummy variable which is 1 for pairs of words with different part-of-speech tags (generated by a part-of-speech tagger run on the full sentence which the word is drawn from) and 0 otherwise.

- *Simplification*

  - *Word rank* (classification score = .77): The absolute difference in the rank of the words with respect to their frequencies in the underlying corpora.
  - *Word length* (classification score = .66): The absolute difference in the lengths of the words.
  - *Syllable ratio* (classification score = .61): The absolute difference in the number of syllables in the words.

- *Explicitation*

  - *Explicit naming* (classification score = .58): The sum of *prop* and *pro*: *prop* is a dummy variable which is 1 for pairings between proper nouns (identified using a Named Entity Recognition algorithm) and words that are not proper nouns and 0 otherwise. *pro* is another dummy variable which is 1 for pairings between personal pronouns and words that are not personal pronouns and 0 otherwise.

- *Normalization*

  - *Contractions* (classification score = .50): A dummy variable which is 1 for pairings between contractions and words that are not contractions and 0 otherwise.

Since the underlying `pyemd` Python implementation of the Earth Mover's Distance[28] requires *all* pairwise distances between words as inputs, we can calibrate this weighting scheme to the range of distances at little additional cost, in a manner which ensures that the original TMD will still the predominate factor in the weighted outcome. To accomplish this latter goal, the algorithm takes a maximum scaling factor $\alpha \geq 1$ as input and sets the weights so that the final scaling factor on the distance lies within $[1, \alpha]$. This ensures, in other words, that the weighting process can never increase the distance by more than $100(\alpha - 1)\%$[29]. The sum of the classification scores, it turns out, is almost exactly 5.0, making it easy for us to take them into account in the scaling procedure:

1. Let S be the set of unique words in the source sentence and T be the set of unique words in the target sentence. Compute the $|S| \times |T|$ distance matrix D where each element $D_{i,j}$ is the Euclidean distance between the embedding for word $i \in S$ and the embedding for word $j \in T$.

2. For each feature $f_k,\ k \in \{1, 2, \ldots, 8\}$, compute the raw $|S| \times |T|$ feature matrix $M^{(k)}$ where each element $M^{(k)}_{i,j}$ is the value of the feature for the pair $(i, j),\ i \in S,\ j \in T$.

3. For each feature $f_k,\ k \in \{1, 2, \ldots, 8\}$, compute the *scaled* feature matrix $\widehat{M}^{(k)}$ where each element $\widehat{M}^{(k)}_{i,j}$ is scaled by $1/(\max\{M^{(k)}\})$, thus making the largest element in the matrix equal to 1, and scaling the remaining elements accordingly.

4. Construct an *overall* feature matrix $\widehat{M}$ by taking a weighted sum of the scaled feature matrices $M = \sum_{k \in \{1,2,\ldots,8} \beta_k \widehat{M}^{(k)}$. Each weight β_k on matrix $\widehat{M}^{(k)}$ is computed based on feature k's classification score σ_k as

$$\beta_k = \frac{100(\alpha - 1)}{5.0} \sigma_k$$

[28] `https://github.com/wmayner/pyemd`

[29] For the studies in Section 7, we set $\alpha = 1.1$, so that the "translationese" features never increase the distance by more than 10%.

For example, with a maximum scaling factor $\alpha = 1.1$ the weight β_1 on $\widehat{M}^{(1)}$ is $\frac{100(1.1-1)}{5.0}\sigma_1 = 2(.97) = 1.94$. As mentioned above, since the classification scores sum to 5.0 this ensures that no scaling factor ever exceeds α.

5. Finally, compute the weighted distance matrix D^* as the Hadamard product $D \circ \widehat{M}$, *i.e.*, simply weight each distance $D_{i,j}$ by the corresponding scaled overall feature weight $\widehat{M}_{i,j}$.

The Translation Mover's Distance then uses the weighted distance matrix D^* rather than the "default" distance matrix D used by the original Word Mover's Distance when calling the `pyemd` function.

In summary, the contribution of the Translation Mover's Distance is twofold: first, it uses a *bilingual* embedding space rather than the Word Mover's Distance's *monolingual* space to allow distance computations between sentences in different languages. Then, it further augments the WMD by using a weighted distance matrix which incorporates textual features informed by the Corpus-Based Translation Studies literature. In the next section we illustrate the use of the TMD via a case study of English translations of Karl Marx's *Das Kapital* and the controversies surrounding their translation choices.

7 Case Study: Marx's *Das Kapital*

In this section we put the "Translation Mover's Distance" to work, by using it to perform analyses of translations at the chapter, section, and document level. As discussed at the end of Section 2, the USSR's Progress Publishers released a "corrected" edition of the First English Edition (Marx 1887) of *Capital* in 1965, while British historian Ben Fowkes performed a second English translation in 1976. Since a large portion of post-1970s *Capital* translations have been made using one of these two versions as source texts, the global diffusion of the ideas in *Capital* is in a sense rooted in the translation choices made in these two editions.

Thus an extended study of post-colonial translations of *Capital* (as proposed in Section 8.1 below) should begin with an analysis of these choices.

Returning to our model, then, we make the following observations: first, because of the very different levels of political pressure from the state faced by the translators, our model implies that the resulting translations represent two distinct political-ideological ideal points. The translator (or "correcter") of the Progress Publishers edition, working under the direct supervision of the Soviet Politburo, should make translation choices very close to the latter's ideal point, which we term the "Soviet" ideal point. Fowkes on the other hand, working with less fear of state punishment, should make translation choices close to his own ideal point. Because of the choice to begin this latter edition with an interpretation of *Capital* by Ernest Mandel – a prominent leader of the Trotskyist Fourth International whose biography calls him "one of the most prominent anti-Stalinist Marxist intellectuals of his time" (Stutje 2009) – we refer to this ideal point as the "Trotskyist" ideal point[30].

With these assumptions established, we can now state the main questions of this study: (1) which of the two overall ideal points – "Trotskyist" or "Soviet" – is closer to the point represented by Marx's original German text? and (2) for which particular topics in *Capital* do these points deviate most from Marx's original text?

To answer these questions, we performed a full analysis of the Progress Publishers version and Fowkes version via the pipeline described in Section 6. The resulting chapter-level and book-level distances are reported in Table 1.

Several striking features stand out. First, we see that neither translation can be said to be straightforwardly "better" overall, where "better" is defined as more closely adhering to the semantic meaning of the original German text. Although the document-level TMD

[30]See also Mandel (1995). Though the personal political leanings of Fowkes himself are unknown, he is a regular contributor to the Trotskyist-aligned *New Left Review* (see Thompson 2007 for an overview of the historical and ideological links between NLR and the Fourth International) and has had works published by Haymarket Books, the publishing house of the Trotskyist International Socialist Organization.

Chapter			Soviet	Trotskyist	Ratio
I	1	Commodities	**1903.81**	1816.74	1.05
	2	Exchange	**342.40**	323.11	1.06
	3	Money, or the Circulation of Commodities	**1868.40**	1821.65	1.03
II	4	The General Formula for Capital	1081.37	**1133.89**	0.95
	5	Contradictions in the General Formula of Capital	**411.30**	402.12	1.02
	6	The Buying and Selling of Labour-Power	343.39	**346.82**	0.99
III	7	The Labour-Process and the Process of Producing Surplus-Value	857.56	**874.68**	0.98
	8	Constant Capital and Variable Capital	429.43	**441.03**	0.97
	9	The Rate of Surplus-Value	**663.09**	639.05	1.04
	10	The Working-Day	**2155.89**	2148.23	1.00
	11	Rate and Mass of Surplus-Value	**275.82**	274.36	1.01
IV	12	The Concept of Relative Surplus-Value	**312.41**	306.15	1.02
	13	Co-operation	**406.75**	403.13	1.01
	14	Division of Labour and Manufacture	**934.19**	915.82	1.02
	15	Machinery and Modern Industry	**4020.02**	3900.60	1.03
V	16	Absolute and Relative Surplus-Value	**382.43**	378.62	1.01
	17	Changes of Magnitude in the Price of Labour-Power	438.07	**439.06**	1.00
	18	Various Formula for the Rate of Surplus-Value	112.40	**112.51**	1.00
VI	19	The Transformation of the Value of Labour-Power	**268.35**	266.54	1.01
	20	Time-Wages	**242.61**	233.09	1.04
	21	Piece-Wages	196.22	**198.37**	0.99
	22	National Differences of Wages	114.95	**180.34**	**0.64**
VII	23	Simple Reproduction	**446.67**	438.85	1.02
	24	Conversion of Surplus-Value into Capital	**1100.09**	1059.43	1.04
	25	The General Law of Capitalist Accumulation	**3416.99**	3112.63	**1.10**
VIII	26	The Secret of Primitive Accumulation	141.09	**144.11**	0.98
	27	Expropriation of the Agricultural Population from the Land	**480.36**	443.73	1.08
	28	Bloody Legislation against the Expropriated	**298.29**	289.68	1.03
	29	Genesis of the Capitalist Farmer	70.85	**72.89**	0.97
	30	Reaction of the Agricultural Revolution on Industry	**145.93**	132.93	**1.10**
	31	Genesis of the Industrial Capitalist	**411.48**	399.94	1.03
	32	Historical Tendency of Capitalist Accumulation	**114.49**	109.37	1.05
	33	The Modern Theory of Colonisation	**338.21**	326.30	1.04

Table 1: Chapter-level Translation Mover's Distances. The point with the greater distance is bolded for each chapter, as are the highest (1.10) and lowest (0.64) distance ratios.

Section		Soviet	Trotskyist	Ratio
I	Commodities and Money	**4114.61**	3961.50	1.04
II	The Transformation of Money into Capital	1836.07	**1882.82**	0.98
III	The Production of Absolute Surplus-Value	**4381.78**	4377.35	1.00
IV	Production of Relative Surplus Value	**5673.38**	5525.71	1.03
V	The Production of Absolute and of Relative Surplus-Value	**932.89**	930.19	1.00
VI	Wages	822.14	**878.35**	**0.94**
VII	The Accumulation of Capital	**4963.75**	4610.91	**1.08**
VIII	Primitive Accumulation	**2000.69**	1918.95	1.04
Total		**24 725.31**	24 085.78	1.03

Table 2: Section-level Translation Mover's Distances. The point with the greater distance is bolded for each section, as are the highest (1.08) and lowest (0.94) distance ratios.

implies that the Soviet translation "moves" the semantic meaning 3% more than the Trotskyist translation does overall, disaggregating this document-level measure to the section and chapter level reveals a more subtle picture. For example, although Section III and Section V both exhibit a nearly even section-level distance ratio, the range of *chapter*-level distances in Section III (0.97 to 1.04) is 7 times greater than the corresponding range of distances in Section V (1.00 to 1.01). The general lesson from this is that for a "hierarchically organized" text like *Capital* (as opposed to, say, a collection of independent essays), in which chapters represent more fine-grained sub-topics of the topics indicated by section titles, the document- and section-level distances should not be analyzed in isolation.

We can also see that the low ratio in Section VI is almost entirely driven by the large discrepancy in distances between the Soviet and Trotskyist points for Chapter 22, "National Differences of Wages", with the Trostskyist point being significantly farther from the original semantics of Marx's text than the Soviet point. As this is the first chapter to focus on *cross-national* comparison of *Capital*'s economic categories (wages, productivity, and absolute surplus value), this observation has fascinating implications given the political and historical circumstances under which these two translations were performed. One of the most epochal events of 20th century socialism was the USSR's decision in November of 1956 to send

30,000 troops to crush the anti-Soviet Hungarian Uprising which had swept the country that October. Opposition to this move led to mass defections of Marxists around the globe away from Soviet-aligned Communist Parties and into anti-Soviet Trotskyist organizations[31]. Prominent leftist intellectuals and artists of the time also developed a new set of critiques of Soviet Communism. Albert Camus, who had been expelled from the Algerian Communist Party for being a "Trotskyist agitator" (Knorr 1996), published an open letter on the one-year anniversary of the uprising titled "The Blood of the Hungarians" (Camus 1957)[32], and Aimé Césaire penned an intensely critical letter of resignation from the French Communist Party (Césaire 2010).

In a sense, then, the Soviet Union was ideologically "on the defensive" by the time of the 1965 Progress translation. The Fowkes translation, on the other hand, was written in the aftermath of not just Hungary 1956 but also the 1968 Soviet invasion and occupation of Czechoslovakia in response to the "Prague Spring" liberalization movement. In fact, one of the main "Prague Spring" dissident groups, the Revolutionary Youth Movement, cited none other than Fowkes-introduction-writer Ernest Mandel as a primary influence on their platform and published translations of his works (along with Trotsky's) into Czech (Alexander 1991, 236; United Secretariat of the Fourth International 1969, 885).

Taking all of these observations together, we can surmise that one potential driver of the greater divergence in Fowkes' translation was the desire to introduce a more normative, critical perspective on the role of imperialism into Marx's descriptive discussion of cross-national differences. In particular, since Trotskyists denounced the Soviet Union as simply an imperialist "state capitalist" power in light of 1956 and 1968, such a critique would implicate the USSR alongside Western powers.

[31]For example, 7,000 members of the Soviet-aligned Communist Party of Great Britain – a quarter of its total membership – defected in 1956 (Beckett and Russell 2015; see also Brotherstone 2007).

[32]There is some evidence to suggest that Camus' death was the result of a KGB operation launched specifically in response to this letter. See `https://www.theguardian.com/books/2011/aug/07/albert-camus-killed-by-kgb`.

Indeed, Chapter 22 is often cited in discussions of the political-theoretic implications of *Capital* with respect to imperialism. Harry Cleaver's *Study Guide to* Capital *Volume I* (2005), for example, which draws on both the Progress Publishers and Fowkes translations (Cleaver 2000), analyzes the chapter through the lens of "economic imperialism", comparing Marx's discussion of British plantations in Jamaica to the workings of the World Trade Organization, NAFTA, and the "free trade movement". William Clare Roberts' *Marx's Inferno: The Political Theory of* Capital similarly situates the chapter in a discussion of outsourcing and the Global South (Roberts 2017, 180).

To summarize, then, we have been able to use the implications of our model alongside the results from our algorithm to draw inferences about what the translators were *doing* in issuing the translational speech act – namely, that the Fowkes translation's greater alteration of the semantics of Chapter 22 is indicative of a Mandel-esque critique of the "state capitalist" imperialism of the Soviet Union.

Returning to our findings, the final salient point for our discussion is that the final two sections, "The Accumulation of Capital" and "Primitive Accumulation", are the two sections with the greatest relative deviation of the Soviet point from the Trotskyist point, the reverse of the Chapter 22 case. These are the sections where Marx explicitly discusses topics of particular relevance to the political grievances of the "global south": colonialism, historical expropriation of land and the means of production, unequal accumulation of wealth via exploitation, and so on. To this day, for example, the notion of "primitive accumulation" introduced in Section VIII is explicitly cited in critiques of colonialism by a wide range of individuals, from Marx scholars (Harvey 2004) to indigenous activists (Coulthard 2014) to South Asian economists (Sanyal 2013).

Because of the wide availability of Progress Publishers editions (in English and otherwise) across the globe, especially in "subaltern" environments[33], Progress could be confident that

[33]Today, for example, the Fowkes edition is still under copyright and often quite expensive (and officially

the ideological "spin" they put on these two sections in particular would be absorbed by the individuals and movements engaging in the postcolonial struggles which swept the globe after World War II. Frantz Fanon for example, studying in France in the late 1940s, almost certainly read Progress Publishers editions of Marx's works.

As Prashad (2014) vividly documents, however, Cold War developments outside of the book-publishing world introduced a series of new challenges to Soviet influence. The USSR viewed the formation of the "non-aligned movement" after the Bandung Conference of 1955 as a grave threat to their ability to coordinate the trajectory of "world communism" via their material and ideological influence on anticolonial movements. The Sino-Soviet Split of 1956 further added to their concerns, introducing Communist China as another "challenger". Dannreuther (1998) and Chamberlin (2012) illustrate these tensions through case studies of the "tug-of-war" between the USSR, Communist China, and non-aligned Algeria for influence on the operations of the Palestine Liberation Organization, a dynamic repeated (with Yugoslavia in the place of Algeria) in North Korea (Cha 2012) and Vietnam (Khoo 2010).

We therefore make the inference that the greater "liberties" taken in Progress' translational speech acts of Sections VII and VIII represent an instance of them *doing something* with translation – namely, attempting to steer these postcolonial struggles in the direction of alignment with the Soviet Union. We note, however, that this argument is embedded within a broader argument regarding Progress' translation and publication of *Capital* in countries all across the "global south". At the end of the next section, therefore, we propose a broader study which will enable us to test our translational-speech-act hypotheses in contexts outside of the English-speaking world.

published only in a few Western industrialized democracies), while the Progress edition is in the public domain and freely available online (`https://www.marxists.org/archive/marx/works/1867-c1/`), as are numerous Progress translations based on this edition.

8 Conclusion and Future Directions

After surveying the history of translation through a political lens, we introduced a novel "Translation Mover's Distance" algorithm leveraging cross-lingual word embeddings to highlight portions of a translation which stand out as introducing large shifts in semantic meaning. By incorporating insights from the field of Corpus-Based Translation Studies, we argued that the algorithm improves upon the previous "Word Mover's Distance" algorithm in capturing the linguistic innovations associated with "translationese" – innovations which often "do things" in the Austinian/Cambridge School sense. We then operationalized this notion by using the Translation Mover's Distance to analyze the two major English translations of Karl Marx's *Capital* in light of the political and historical contexts in which they were produced.

In the remainder of this section, we propose an extension to the case study of Section 7, applying the algorithm to a study of the worldwide dissemination of *Capital* and how its latter chapters and sections were differentially translated into societies with varying levels of industrialization.

8.1 The Political Economy of the Dissemination of *Capital*

Broadly speaking, an analysis of the dissemination of *Capital* across the globe (illustrated in Figure 6[34]) shows a pattern organized around the "First, Second, and Third World" trichotomy:

1. Translations into the languages of the "First World" – English, French, Italian – occurred and were published mostly before 1900. Most of these translations were performed either by Marx and Engels themselves or under the direct guidance of Engels.

[34]Dates of publication were culled from a wide range of sources, but the primary "starting points" for the data gathering project were Hobsbawm (2012), Hoff (2017), Musto (2012), Amini (2016), Rauhala (2017), Guillermo (2018), Ertürk and Serin (2016), Afary (2012), Yoshihiro (2012), and Merlassino (2011).

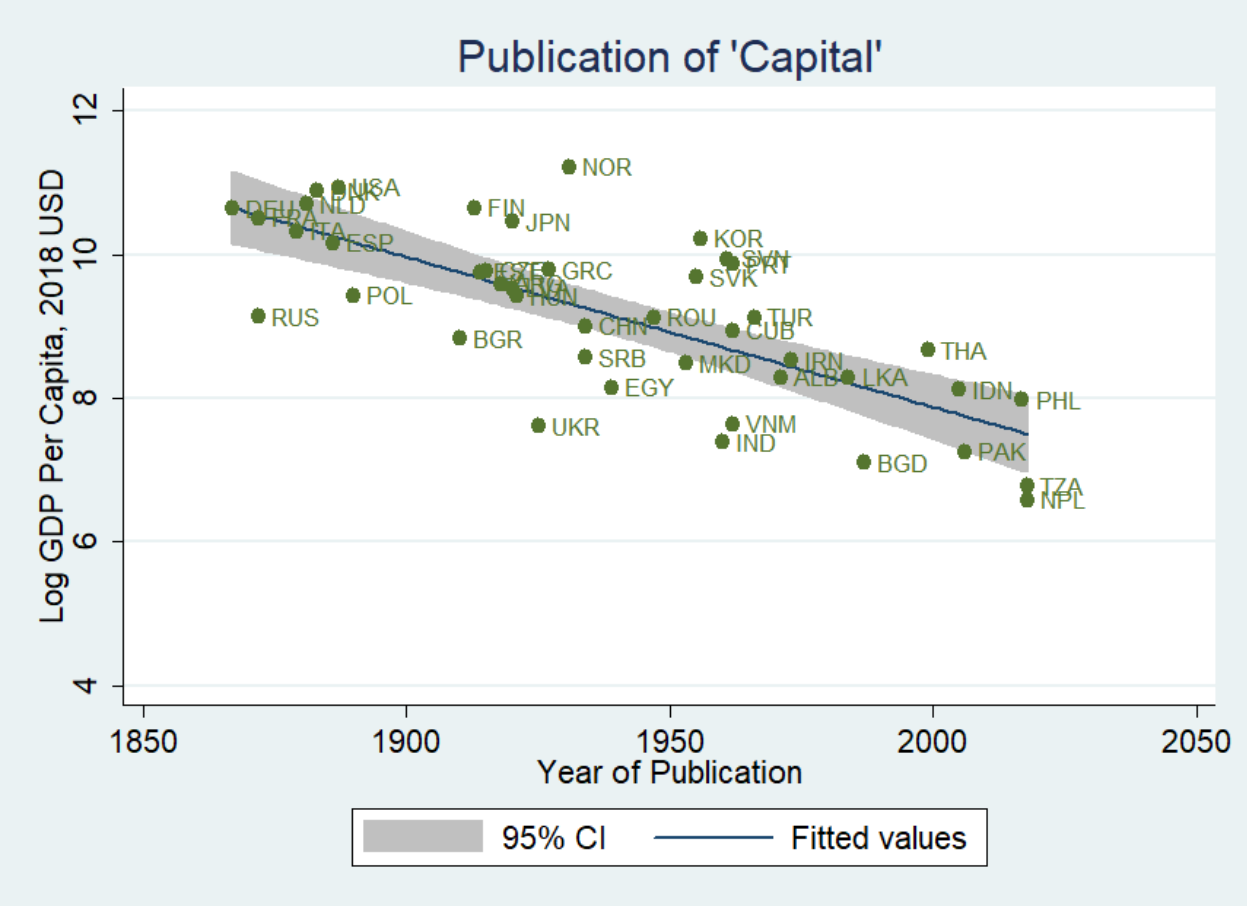

Figure 6: The dissemination of *Capital* from its 1867 German publication to the present.

For example, Engels essentially served as one of the uncredited translators of the first English edition, as revealed through his correspondence with Aveling, though he was credited as editor[35].

2. Translations into the languages of the Soviet-aligned "Second World" – Ukrainian, Serbian, Albanian, Korean[36] – then occurred and were published between 1910 and 1970.

[35]The translation was written mostly by a biology professor, Edward Aveling, who had never translated a non-biology-related text. Aveling was also assisted in a major way by his wife, Marx's daughter Eleanor Marx, who was not credited as an editor nor as a translator.

[36]Korean is listed under the "Second World" here since the first full Korean translation of *Capital* was published in North Korea in 1959 under the supervision of the Kim Il-Sung regime. In South Korea, on the other hand, from 1948 to 1987 possession of any book by Karl Marx carried a penalty of two years in prison, and several translators of Marx remained incarcerated beyond 1987 (Musto 2012, 222).

As discussed in the prior section, several of the pre-1938 translations were destroyed during Stalin's "Great Purge" and new translations were subsequently issued.

3. Finally, translations into the languages of the "Third World" – Thai, Urdu, Bengali, Swahili – occurred and were published from the 1990s onward, if they occurred at all. Latin America and the Middle East/North Africa seem to be exceptions to this trend, due to the widespread availability of Spanish translations (though, importantly, not Portuguese translations) after 1886 (Merlassino 2011) and the 1939 publication of an influential Arabic translation in Egypt (Hobsbawm 2011). However, the pattern holds robustly for most major languages of Sub-Saharan Africa (Swahili, Amharic), South Asia (Bengali, Nepali, Urdu, Sinhala), and Southeast Asia (Thai, Bahasa Indonesia, Tagalog).

As discussed in Section 7, Progress Publishers exerted control over most of the translations read during the final two "eras", in contrast to the Marx- or Engels-directed translations of the first. In the vein of Franco Moretti's (2013) study of the diffusion of world literature from "core" to "periphery"[37], the broad questions our study will ask are as follows:

1. As a key "Ideological State Apparatus" of the Soviet Union, in what ways did they construct their translational speech-acts at the statement, chapter, or section level so as to affect the trajectory of various "Third World" liberation movements? For example, were Sections VII and VIII "moved" in different semantic directions based on the class composition of the target societies? We hypothesize that indeed there are two separate "types" of translational speech-acts:

   (a) A "peasant-focused translation" of *Capital*, *i.e.*, a translational speech-act which aimed to influence the peasantry in countries where agricultural output formed

[37]Moretti purposefully adapted these terms from the "world-systems" model of Immanuel Wallerstein (2004), but Moretti's model works in reverse: literary output (and, by extension, ideology) flows *from* the core to the periphery.

a large share of GDP. During the Cold War era we discussed in Section 7, for example, the countries with agricultural-sector concentration above 50% were Lesotho, Somalia, Nepal, Oman, Bangladesh, and Uganda, according to World Bank data. We would therefore surmise that translations of *Capital* published in these countries (we already have full texts of the Progress-edition-based Nepali and Bengali translations) would be closest to the "peasant-focused" ideal point.

(b) A "proleteriat-focused translation" of *Capital*, *i.e.*, a translational speech-act which aimed to influence the working class in countries where industrial output formed a large share of GDP. In this case, the countries with industrial-sector concentration above 40% were Zambia, Argentina, Suriname, Chile, and South Africa.

A forthcoming handbook, the *Routledge Handbook of Marx's Capital: A Global History of Translation, Dissemination, and Reception*, will be of immense help in tracking down texts of Progress' translations. It will contain chapters detailing the translation history of *Capital* in (for example) Iran, North Africa, South Asia, Francophone West Africa, Angola and Mozambique, and Tanzania, thus encompassing most countries for which our hypotheses would apply.

2. Relatedly, what was the effect of these speech-acts on subsequent developments in the liberation movements they reached? This portion of the study will draw on insights from the field of "computational stylometry"[38] to trace the ideological impact of Progress' versions of *Capital* on subsequent texts issued by these movements. In essence, these methods allow one to infer a latent "stylistic fingerprint" from one text A and then measure the degree to which this "fingerprint" was left on another text B. Such methods have been used, for example, to study ideological influences on Livy by identifying "material that [he] loosely appropriated from earlier sources" (Dexter et al. 2017),

[38]See Stamatatos (2008) for a survey. A stylometric "corollary" to Corpus-Based Translation Studies is "translator stylometry", surveyed in Rybicki (2012) and El-Fiqi (2011).

and to demarcate "epistemological ruptures" by identifying temporal clusters within which ideological influence is strong but between which there is little observed influence (Hughes et al. 2012)[39]. In our case, then, the aim will be to apply these methods in a similar way to the texts of both rural and urban liberation groups, to measure the degree to which the fingerprints of Progress' "peasant" and "proletarian" translations are present. In other words, we aim to see whether the *parole* of their translational speech-acts indeed affected the *langue* of Marxist Third World liberation.

References

[1] Afary, Frieda. 2012. "Persian 'Capital': Hassan Mortazavi's New Translation of Marx's Classic Work". *The International Marxist-Humanist*. `https://www.internationalmarxisthumanist.org/wp-content/uploads/article-afary-20120331.pdf`

[2] Alexander, Robert J. 1991. *International Trotskyism, 1929–1985: A Documented Analysis of the Movement*. Durham, NC: Duke University Press.

[3] Amini, Babak. 2016. "A Brief History of the Dissemination and Receiption of Karl Marx's *Capital* in the United States and Britain". *World Review of Political Economy*, 7(3): 334–349. `http://www.jstor.org/stable/10.13169/worlrevipoliecon.7.3.0334`

[4] Arthur, Christopher J. 1996. "Engels as Interpreter of Marx's Economics". In Arthur, Christopher J. (ed.), *Engels Today: A Centenary Appreciation*, 173–210. New York, NY: St. Martin's Press.

[5] Ascher, Ivan. 2005. "All the Wor(l)d's a Stage: Persons and Personification in Karl Marx's *Capital*". *ECPR Workshop on Metaphor in Political Science*. `http://eis.bris.ac.uk/~potfc/Granada/Papers/Ascher.pdf`

[6] Austin, J. L. 1955. *How to Do Things with Words*. Cambridge: Harvard University Press.

[7] Baker, Mona. 1993. "Corpus Linguistics and Translation Studies: Implications and Applications". In Baker, Mona et al. (eds.), *Text and Technology: In Honour of John Sinclair*, 233–250. Amsterdam: John Benjamins.

[39] "Stylometric" methods were first developed to de-anonymize the Federalist Papers based on "fingerprints" from the known writings of Hamilton, Madison, and Jay (Mosteller and Wallace 1964). Most applications today still focus on authorship identification, for example to classify which portions of Supreme Court decisions were written by justices and which were written by their clerks (Rosenthal and Yoon 2010).

[8] Baroni, Marco and Silvia Bernardini. 2006. "A New Approach to the Study of Translationese: Machine-Learning the Difference between Original and Translated Text". *Literary and Linguistic Computing*, 21(3): 259–274. https://doi.org/10.1093/llc/fqi039

[9] Beckett, Francis and Tony Russell. 2015. *1956: The Year That Changed Britain*. London: Biteback Publishing.

[10] Beebee, Thomas O. 2014. "Shoot the Transtraitor!: The Translator as *Homo Sacer* in Fiction and Reality". *The Translator*, 2: 295–313. https://doi.org/10.1080/13556509.2010.10799473

[11] Bellos, David. 2011. *Is That a Fish in Your Ear?: Translation and the Meaning of Everything*. New York, NY: Farrar, Strauss, and Giroux.

[12] Berlin, Isaiah. 2000 [1946]. "The Arts in Russia Under Stalin". *The New York Review of Books*, October 19. http://www.nybooks.com/articles/2000/10/19/the-arts-in-russia-under-stalin/

[13] Birkby, Stuart J. 2014. "'We Will Bury You': How Mistranslation Heightened Cold War Tensions". *Galaxy*, 6(2): 89–102. http://repository.au.edu/bitstream/handle/6623004553/13797/galaxy-iele-n2-6-2013.pdf

[14] Bojanowski, Piotr, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2016. "Enriching Word Vectors with Subword Information". *ArXiv Preprint*. https://arxiv.org/abs/1607.04606

[15] Bonefeld, Werner. 2001. "Social Form, Critique, and Human Dignity". *Zeitschrift für Kritische Theorie*, 13: 97–112. https://libcom.org/library/social-form-critique-and-human-dignity

[16] Brotherstone, Terry. 2007. "1956 and the Crisis in the Communist Party of Great Britain: Four Witnesses". *Critique*, 35(2): 189–209. https://doi.org/10.1080/03017600701446173

[17] Brown, Peter F., Stephen A. Della Pietra, Vincent J. Della Pietra, and Robert L. Mercer. 1993. "The Mathematics of Statistical Machine Translation: Parameter Estimation". *Computational Linguistics*, 19(2): 263–311. https://dl.acm.org/citation.cfm?id=972474

[18] Buick, Adam. 1980. "A New Translation of Marx's 'Capital'". *Socialist Standard*, 908: 1–3. https://www.worldsocialism.org/spgb/socialist-standard/1980s/1980/no-908-april-1980/new-translation-marxs-capital

[19] Camus, Albert. 1957. "The Blood of the Hungarians". Available at https://bookhaven.stanford.edu/2011/08/albert-camus-the-words-that-may-have-cost-his-life/.

[20] Césaire, Aimé. 2010 [1956]. "Letter to Maurice Thorez". *Social Text*, 28(2): 145–152. https://doi.org/10.1215/01642472-2009-072

[21] Cha, Victor. 2012. *The Impossible State: North Korea, Past and Future*. New York, NY: HarperCollins.

[22] Chamberlin, Paul Thomas. 2012. *The Global Offensive: The United States, the Palestine Liberation Organization, and the Making of the Post-Cold War Order*. New York, NY: Oxford University Press.

[23] Chiang, Yung-Chen. 2001. *Social Engineering and the Social Sciences in China, 1919–1949*. New York, NY: Cambridge University Press.

[24] Cleaver, Harry. 2000. *Reading* Capital *Politically*. Oakland, CA: AK Press.

[25] Cleaver, Harry. 2005. *Study Guide to* Capital *Volume I*. Available at *https://la.utexas.edu/users/hcleaver/357k/357ksg.html*

[26] Colla, Elliott. 2015. "Dragomen and Checkpoints". *The Translator*, 21(2): 132–153. http://www.tandfonline.com/doi/full/10.1080/13556509.2015.1071523

[27] Coulthard, Glen Sean. 2014. *Red Skin, White Masks: Rejecting the Colonial Politics of Recognition*. Minneapolis, MN: University of Minnesota Press.

[28] Dannreuther, Roland. 1998. *The Soviet Union and the PLO*. New York: Palgrave Macmillan.

[29] Dexter, Joseph P., Theodore Katz, Nilesh Tripuraneni, Tathagata Dasgupta, Ajay Kannan, James A. Brofos, Jorge A. Bonilla Lopez, Lea A. Schroeder, Adriana Casarez, Maxim Rabinovich, Ayelet Haimson Lushkov, and Pramit Chaudhuri. 2017. "Quantitative Criticism of Literary Relationships". *Proceedings of the National Academy of Sciences*, 114(16): E3195–E3204. http://www.pnas.org/content/114/16/E3195.short

[30] Dunayevskaya, Raya. 1978. *Marx's* Capital *and Today's Global Crisis*. Detroit, MI: News and Letters.

[31] Eaton, Henry. 1980. "Marx and the Russians". *Journal of the History of Ideas*, 41(1): 89–112. https://www.jstor.org/stable/2709104

[32] Eckhardt, Wolfgang. 2016. *The First Socialist Schism: Bakunin vs. Marx in the International Working Men's Association*. Oakland: PM Press.

[33] Ehrbar, Hans. 2010a. "Annotations to Karl Marx's *Capital*". *Marx Annotations* Project. http://content.csbs.utah.edu/~ehrbar/akmc.pdf

[34] Ehrbar, Hans. 2010b. "Glossary to Marx's *Capital* and Other Economic Writings". *Marx Annotations* Project. http://content.csbs.utah.edu/~ehrbar/glossary.pdf

[35] El-Fiqi, Heba, Eleni Petraki, and Hussein A. Abbass. 2011. "A Computational-Linguistic Approach for the Identification of Translator Stylometry using Arabic-English Text". *IEEE International Conference on Fuzzy Systems*, 2039–2045. http://ieeexplore.ieee.org/document/6007535/

[36] Ertürk, Nergis and Özge Serin. 2016. "Marxism, Communism, and Translation: An Introduction". *boundary 2*, 43(3): 1–26. https://doi.org/10.1215/01903659-3572526

[37] Fathi, Nazila. 2005. "Wipe Israel 'Off The Map', Iranian Says". *The New York Times*, October 27. http://www.nytimes.com/2005/10/27/world/africa/wipe-israel-off-the-map-iranian-says.html

[38] Firth, John. 1957. *Papers in Linguistics, 1934–1951*. New York, NY: Oxford University Press.

[39] Gal, Susan. 2015. "Politics of Translation". *Annual Review of Anthropology*, 44: 225–240. http://www.annualreviews.org/doi/abs/10.1146/annurev-anthro-102214-013806

[40] George, Edward. 2013. "Marxism and Translation". *Reading Marx* Blog. https://edgeorgesotherblog.files.wordpress.com/2012/12/marxism-and-translation1.pdf

[41] Ghobrial, John-Paul. 2013. *The Whispers of Cities: Information Flows in Istanbul, London, and Paris in the Age of William Trumbull*. New York, NY: Oxford University Press.

[42] Goldberg, Yoav. 2017. *Neural Network Methods for Natural Language Processing*. Williston, VT: Morgan & Claypool. https://doi.org/10.2200/S00762ED1V01Y201703HLT037

[43] Guillermo, Ramom. 2018. "The Philippines". In Musto, Marcello and Babak Amini (eds.), *The Routledge Handbook of Marx's* Capital*: A Global History of Translation, Dissemination, and Reception* (Forthcoming). New York, NY: Taylor and Francis.

[44] Hanzel, Igor. 2010. "Mistranslations of 'Schein' and 'Erscheinung': The Structure of Chapter 1 of *Capital*, Volume I". *Science and Society*, 74(4): 509–537. https://doi.org/10.1521/siso.2010.74.4.509

[45] Harvey, David. 2004. "The 'New' Imperialism: Accumulation by Dispossession". *Socialist Register*, 40: 63–87. http://www.socialistregister.com/index.php/srv/article/view/5811

[46] Harvey, David. 2010. *A Companion to Marx's* Capital. New York, NY: Verso.

[47] Haug, Wolfgang Fritz. 2017. "On the Need for a New English Translation of Marx's *Capital*". *Socialism and Democracy*, 31(1): 60–86. http://www.tandfonline.com/doi/full/10.1080/08854300.2016.1274165

[48] Hill, Clive. 2015. *Marx's Phantasmagorical Metaphor*. Unpublished Manuscript. https://cliveehillacademic.wordpress.com/2015/08/19/karl-marxs-phantasmagorical-metaphor-an-introduction/

[49] Hobsbawm, Eric. 2012. *How to Change the World: Reflections on Marx and Marxism*. New Haven, CT: Yale University Press.

[50] Hoff, Jan. 2017. *Marx Worldwide: On the Development of the International Discourse on Marx since 1965*. Leiden: Brill.

[51] Hryn, Halyna. 2005. "The Executed Renaissance Paradigm Revisited". *Harvard Ukrainian Studies*, 27(4): 67–96. https://www.jstor.org/stable/41036862

[52] Hughes, James M., Nicholas J. Foti, David C. Krakauer, and Daniel N. Rockmore. 2012. "Quantitative Patterns of Stylistic Influence in the Evolution of Literature". *Proceedings of the National Academy of Sciences*, 109(20): 7682–7686. http://www.pnas.org/content/109/20/7682.abstract

[53] Jelveh, Zubin, Bruce Kogut, and Suresh Naidu. 2017. "Political Language in Economics". *Columbia Business School Research Paper No. 14-57*. http://dx.doi.org/10.2139/ssrn.2535453

[54] Khoo, Nicholas. 2010. "Breaking the Ring of Encirclement: The Sino-Soviet Rift and Chinese Policy toward Vietnam, 1964–1968". *Journal of Cold War Studies*, 12(1): 3–42. https://doi.org/10.1162/jcws.2010.12.1.3

[55] Kiss, Tibor and Jan Strunk. 2006. "Unsupervised Multilingual Sentence Boundary Detection". *Computational Linguistics*, 32(4): 485–525. https://dl.acm.org/citation.cfm?id=1245122

[56] Kitov, Eliyahu. 1970. *The Book of Our Heritage*. New York, NY: Feldheim Publishers.

[57] Knorr, Katherine. 1996. "Albert Camus: Being Right About the Left". *The New York Times*, February 23. http://www.nytimes.com/1996/02/23/style/albert-camusbeing-right-about-the-left.html

[58] Koehn, Philipp. 2005. "Europarl: A Parallel Corpus for Statistical Machine Translation". *MT Summit*. http://www.statmt.org/europarl/

[59] Kukis, Mark. 2011. *Voices From Iraq: A People's History, 2003–2009*. New York, NY: Columbia University Press.

[60] Kusner, Matt J., Yu Sun, Nicholas I. Kolkin, and Kilian Q. Weinberger. 2015. "From Word Embeddings to Document Distances". *Proceedings of the 32nd International Conference on Machine Learning*, 37: 957–966. https://dl.acm.org/citation.cfm?id=3045221

[61] Laviosa, Sara. 2002. *Corpus-based Translation Studies: Theory, Findings, Applications*. Amsterdam: Rodopi.

[62] Law, Timothy Michael. 2013. *When God Spoke Greek: The Septuagint and the Making of the Christian Bible*. New York, NY: Oxford University Press.

[63] Legrand, Joel, Michael Auli, and Ronan Collobert. 2016. "Neural Network-Based Word Alignment through Score Aggregation". *ArXiv Preprint*. https://arxiv.org/abs/1606.09560

[64] Lewis, Bernard. 2004. *From Babel to Dragomans: Interpreting the Middle East*. New York, NY: Oxford University Press.

[65] Liang, Percy, Ben Taskar, and Dan Klein. 2006. "Alignment by Agreement". *Proceedings of the Human Language Technology Conference of the North American Chapter of the ACL (NAACL)*, 104–111. http://www.aclweb.org/anthology/N06-1014

[66] Liu, Lydia. 1999. *Tokens of Exchange: The Problem of Translation in Global Circulations*. Durham, NC: Duke University Press.

[67] López, Antonio. 2008. "Traduttore, Traditore – Translator, Traitor". *Counterpoints*, 343: 141–147. https://www.jstor.org/stable/42980184

[68] Luong, Minh-Thang, Hieu Pham, and Christopher D. Manning. 2015. "Bilingual Word Representations with Monolingual Quality in Mind". *Proceedings of NAACL-HLT*, 151-–159. http://www.aclweb.org/anthology/W15-1521

[69] Mandel, Ernest. 1995. *Trotsky as Alternative*. New York, NY: Verso.

[70] Marx, Karl. 1867. *Das Kapital: Kritik der Politischen Ökonomie Buch I, Der Produktionsprozess des Kapitals*. Marx-Engels-Werke (MEW) Vol. 23. Berlin: Dietz.

[71] Marx, Karl. 1887 [1867]. *Capital, Volume I: A Critical Analysis of Capitalist Production*, trans. Samuel Moore and Edward Aveling. London: Sonnenschein, Lowrey & Co.

[72] Marx, Karl. 1965 [1867]. *Capital, Volume I: A Critical Analysis of Capitalist Production*, trans. Samuel Moore and Edward Aveling. Moscow: Progress Publishers.

[73] Marx, Karl. 1976 [1867]. *Capital, Volume I: A Critique of Political Economy*, trans. Ben Fowkes. New York, NY: Penguin.

[74] Marx, Karl. 2010 [1867]. *Capital, Volume I: A Critique of Political Economy*, trans. Hans Ehrbar. *Capital Annotations* Project. http://content.csbs.utah.edu/~ehrbar/cap1.pdf

[75] Marx, Karl and Frederick Engels. 1988 [1848]. *Economic and Philosophic Manuscripts of 1844 and the Communist Manifesto*. Amherst, NY: Prometheus Books.

[76] McDonald, Scott and Michael Ramscar. 2001. "Testing the Distributional Hypothesis: The Influence of Context on Judgements of Semantic Similarity". *Proceedings of the 23rd Annual Meeting of the Cognitive Science Society*, 611–616. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.104.7535

[77] Merlassino, Victoria Correa. 2011. "Guide to the Analysis of the Spanish Translation of Karl Marx's *The Capital: Critique of Political Economy*". *The EE-T Project*. https://eet.pixel-online.org/files/guides/pdf/Spanish%20Translation%20of%20Karl%20Marx's%20The%20Capital.pdf

[78] Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. "Efficient Estimation of Word Representations in Vector Space". *ArXiv Preprint*. https://arxiv.org/abs/1301.3781

[79] Moretti, Franco. 2013. *Distant Reading*. New York, NY: Verso.

[80] Morse, David. 2001. "'Enemies of the People': Poetry and Politics in the Time of Stalin". *Social Education*, 65(4): 198–207. https://eric.ed.gov/?id=EJ667127

[81] Mosteller, Frederick and David L. Wallace. 1964. *Inference and Disputed Authorship: The Federalist Papers*. Reading, MA: Addison-Wesley.

[82] Musto, Marcello. 2012. *Marx for Today*. New York: Routledge.

[83] Nielson, Jon and Royal Skousen. 2015. "How Much of the King James Bible Is William Tyndale's?: An Estimation Based on Sampling". *Reformation*, 3(1): 49–74. https://doi.org/10.1179/ref_1998_3_1_004

[84] Nitsch, Thomas O. 1992. "Marx on Man's Sociality by Nature: An Inexplicable Omission?". *International Journal of Social Economics*, 19(9): 108–120. https://doi.org/10.1108/EUM0000000000490

[85] O'Kane, Chris. 2013. "Notes on Translation in Part One of *Capital*". *Reification of Persons and Personification of Things* Blog. https://reificationofpersonsandpersonificationofthings.wordpress.com/2013/02/23/notes-on-translation-in-part-one-of-capital/

[86] Olohan, Maeve. 2003. "How Frequent are the Contractions?: A Study of Contracted Forms in the Translational English Corpus". *Target*, 15(1): 59–89. https://benjamins.com/#catalog/journals/target.15.1.04olo/details

[87] Pennington, Jeffrey, Richard Socher, and Christopher Manning. 2014. "GloVe: Global Vectors for Word Representation". *ArXiv Preprint*. https://nlp.stanford.edu/pubs/glove.pdf

[88] Pocock, J. G. A. 1985. *Virtue, Commerce, and History: Essays on Political Thought, Chiefly in the Eighteenth Century*. New York, NY: Cambridge University Press.

[89] Polanyi, Michael. 1958. *Personal Knowledge: Towards a Post-Critical Philosophy*. Chicago, IL: University of Chicago Press.

[90] Prashad, Vijay. 2014. *The Poorer Nations: A Possible History of the Global South*. New York, NY: Verso.

[91] Prieto, José Manuel. 2013. "On Translating a Poem by Osip Mandelstam". In Allen, Esther and Susan Bernofsky (eds.), *In Translation: Translators on Their Work and What It Means*, 127–142. New York, NY: Columbia University Press.

[92] Rauhala, Paula. 2017. "The Dissemination of *Capital* in Nordic Countries". Lecture presented at *Marx's* Capital *After 150 Years: Critique and Alternative to Capitalism* conference. Toronto, Ontario. https://www.youtube.com/watch?v=Wk_E-a70BIo

[93] Resis, Albert. 1970. "*Das Kapital* Comes to Russia". *Slavic Review*, 29(2): 219–237. https://www.jstor.org/stable/2493377

[94] Roberts, William Clare. 2017. *Marx's Inferno: The Political Theory of* Capital. Princeton, NJ: Princeton University Press.

[95] Rosenthal, Jeffrey S. and Albert H. Yoon. 2010. "Judicial Ghostwriting: Authorship on the Supreme Court". *Cornell Law Review*, 96: 1307–1343. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1668735

[96] Rubner, Yossi, Carlo Tomasi, and Leonidas J. Guibas. 2000. "The Earth Mover's Distance as a Metric for Image Retrieval". *International Journal of Computer Vision*, 40(2): 99–121. http://robotics.stanford.edu/~rubner/papers/rubnerIjcv00.pdf

[97] Ruder, Sebastian, Ivan Vulić, and Anders Søgaard. 2017. "A Survey of Cross-Lingual Word Embedding Models". *ArXiv Preprint*. https://arxiv.org/abs/1706.04902

[98] Rybicki, Jan. 2012. "The Great Mystery of the (Almost) Invisible Translator: Stylometry in Translation". *Studies in Corpus Linguistics*, 51: 231–248. http://dx.doi.org/10.1075/scl.51.09ryb

[99] Sanyal, Kalyan. 2007. *Rethinking Capitalist Development: Primitive Accumulation, Governmentality, and Post-Colonial Capitalism*. New York, NY: Routledge.

[100] Saussure, Ferdinand de. 1916. *Course in General Linguistics*. Chicago, IL: Open Court.

[101] Skinner, Quentin. 1969. "Meaning and Understanding in the History of Ideas". *History and Theory*, 8(1): 3–53. https://www.jstor.org/stable/2504188

[102] Stamatatos, Efstathios. 2008. "A Survey of Modern Authorship Attribution Methods". *Journal of the Association for Information Science and Technology*, 60(3): 538–556. http://onlinelibrary.wiley.com/doi/10.1002/asi.21001/full

[103] Stutje, Jan Willem. 2009. *Ernest Mandel: A Rebel's Dream Deferred*. New York, NY: Verso.

[104] Susiluoto, Ilmari. 1982. *The Origins and Development of Systems Thinking in the Soviet Union: Political and Philosophical Controversies from Bogdanov and Bukharin to Present-Day Reevaluations*. Helsinki: Suomalainen Tiedeakatemia.

[105] Tamura, Akihiro, Taro Watanabe, and Eiichiro Sumita. 2014. "Recurrent Neural Networks for Word Alignment Model". *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 1470–1480. http://aclweb.org/anthology/P14-1138

[106] Tercedor, Maribel. 2010. "Cognates as Lexical Choices in Translation: Interference in Space-Constrained Environments". *Target*, 22(2): 177–193. https://benjamins.com/#catalog/journals/target.22.2.01ter/details

[107] Thompson, Duncan. 2007. *Pessimism of the Intellect?: A History of the New Left Review*. London: Merlin Press.

[108] Tiedemann, Jörg. 2011. *Bitext Alignment*. Williston, VT: Morgan & Claypool. https://doi.org/10.2200/S00367ED1V01Y201106HLT014

[109] Tomlin, Russell. 1986. *Basic Word Order: Functional Principles*. London: Croom Helm.

[110] Toury, Gideon. 1982. "The Communication Situation and the Production of Interference Forms by L2 Learners". *RELC*, 13(2): 62–77. http://journals.sagepub.com/doi/abs/10.1177/003368828201300205

[111] United Secretariat of the Fourth International. 1969. "New Revolutionary Party in Czechoslovakia". *Intercontinental Press*, October 6, 885–888. http://www.themilitant.com/Intercontinental_Press/1969/IP0732.pdf#page=21

[112] Varga, Dániel, Péter Halácsy, András Kornai, Viktor Nagy, László Németh, and Viktor Trón. 2005. "Parallel Corpora for Medium-Density Languages". *Proceedings of the RANLP*, 590–596. https://catalog.ldc.upenn.edu/docs/LDC2008T01/ranlp05.pdf

[113] Vygodsky, V. S. 1979. "Das Kapital". *The Great Soviet Encyclopedia*. London: Macmillan Publishers.

[114] Wallerstein, Immanuel. 2004. *World-Systems Analysis: An Introduction*. Durham, NC: Duke University Press.

[115] Wilson, Edmund. 1940. *To the Finland Station: A Study in the Writing and Acting of History*. Garden City, NY: Doubleday.

[116] Wintner, Shuly. 2017. "Computational Approaches to Translation Studies". In Marcel, Patrick and Esteban Zimanyi (eds.), *Business Intelligence*, 38–58. Berlin: Springer.

[117] Wood, Denis. 2010. *Rethinking the Power of Maps*. New York, NY: Guilford Press.

[118] Yoshihiro, Ishikawa. 2012. *The Formation of the Chinese Communist Party*. New York, NY: Columbia University Press.

[119] Zanettin, Federico. 2016. "'The Deadliest Error': Translation, International Relations and the News Media". *The Translator*, 22(3): 303–318. https://doi.org/10.1080/13556509.2016.1149754

[120] Zhao, Zhe, Tao Liu, Shen Li, Bofang Li, and Xiaoyang Du. 2017. "Ngram2vec: Learning Improved Word Representations from Ngram Co-Occurrence Statistics". *Empirical Methods in Natural Language Processing*. http://aclweb.org/anthology/D17-1023