In search of directionality effects in the translation process and in the end product

Bogusława Whyatt
Adam Mickiewicz University

This article tackles directionality as one of the most contentious issues in translation studies, still without solid empirical footing. The research presented here shows that, to understand directionality effects on the process of translation and its end product, performance in L2 → L1 and L1 → L2 translation needs to be compared in a specific setting in which more factors than directionality are considered—especially text type. For 26 professional translators who participated in an experimental study, L1 → L2 translation did not take significantly more time than L2 → L1 translation and the end products of both needed improvement from proofreaders who are native speakers of the target language. A close analysis of corrections made by the proofreaders shows that different aspects of translation quality are affected by directionality. A case study of two translators who produced high quality L1 → L2 translations reveals that their performance was affected more by text type than by directionality.

Keywords: directionality, L2 translation, TPR, quality assessment, proofreader's comments, Polish-English language pair

1.   Introduction

The issue of directionality—whether translators work into their first or "native" language (L1) or out of their L1 and produce translations in their "first foreign" language (L2)—has been widely discussed in the literature. From the very beginning, it also gained an attitudinal prescriptive overtone according to which, the general consensus is that professional translators should work into their native language and avoid translating into their foreign language (Newmark 1981). Gile (2005, 10) attests that the opposing attitudes are motivated by "a mix of personal experience, ideology and tradition." The consequences of the inferiority label attached to L1 →

L2 translation are far reaching. Whyatt & Kościuczuk (2013) pointed out that they involve academia and major players on the translation market. Institutions and companies employing translators advertise their services by making L2 → L1 translation as their major asset and "a proof of quality" (Beeby 2011, 85). Yet, most scholars would agree that there is very little empirical research to understand how L1 → L2 translation differs from L2 → L1 translation in terms of cognitive load and the quality of the end product (overview in Ferreira & Schwieter 2017).

Before delving into the effects directionality has on the process and product of translation, several problems need to be addressed. First, terminology is still an unsettled matter. Some scholars use *direct translation* for L2 → L1 translation and *inverse translation,* or even *reverse translation* for L1 → L2 translation (Talaván & Rodríguez 2014; Ferreira & Schwieter 2017). Other scholars see the term *inverse* or *reverse translation* as a negative branding of something, intrinsically opposite to what it should be (Beeby 2011, 84). In interpreting, Gumul (2017) opts for *into A language/native language* (L1) and *into B language* (L2) or *retour interpreting.* To complicate matters even further, the labels used in bilingualism research and psycholinguistic studies are intuitively opposite to the translation studies terminology, with L2 → L1 translation called *backward translation* and L1 → L2 translation referred to as *forward translation* (Traxler 2012). The labels chosen in this article with reference to directionality are L2 → L1 and L1 → L2 translation (cf. Pavlović 2007; Pavlović & Jensen 2009), with the reservation that language dominance is not a stable value but always a product of experience in using a particular language. That is, in certain contexts and knowledge domains a translator can be more proficient in L2 than in L1.

The second fundamental problem to study directionality effects is the lack of a theoretical framework to accommodate insights from translation process research (TPR), neurolinguistc and psycholinguistic studies into bilingual language processing. This problem is much more difficult to overcome, mostly because of different research objectives and different methods of investigating how language is processed in the mind. TPR has provided a body of research to suggest that translation is an interplay between automatic fluent processing and analytical effortful problem solving. While translators tap into their cognitive resources, including language, they have to manoeuvre within meaning constraints created by their own interpretation of the source text and the hypothesized expectations of the target text readers. Psycholinguistic studies have demonstrated that the amount of cognitive effort needed to tap into language resources differs for L1 and L2, and is modulated by other factors, the most important being how asymmetrical the two proficiencies are. Most neurolinguistic and psycholinguistic studies use isolated words (mostly, concrete nouns) or single sentences as their research stimuli. For TPR experiments texts are used as stimuli. Still, the findings from both disciplines

can be pooled to explain how directionality affects the translation process and the product (e.g., Chmiel's 2018 study of semantic priming in interpreters).

Paradis (2009, 5) expounds that the language resources in L1 and L2 are not only unequal in terms of volume (usually, smaller vocabulary in L2, fewer stylistic variants, less nuanced knowledge of different shades of meaning) but also that they are stored in different manners: L1 is stored in the procedural memory and L2 (depending on the current proficiency level) may be partially stored as implicit linguistic competence and partially as explicit metalinguistic knowledge (Paradis 2009, 5). These procedural and declarative determinants of L1 and L2 use (e.g., the assumed L1 comprehension advantage and the L2 production disadvantage) must have a bearing on the translation process, which is most likely mitigated by translation expertise and results in error-preventing strategies used by translators. For example, Wu & Liao (2018) argue that interpreting trainees should be explicitly taught rule-of-thumb strategies to overcome the L2 cognitive disadvantage.

The third fundamental problem is the power-relations around L1 → L2 translation. Beeby (2011, 84) points out that attitudes towards directionality depend on contextual factors, mostly market demands and political issues. Most scholars agree that translated language, in whatever direction, is usually affected by the translation process. Toury (1995, 278) notes that tolerance of interference from the ST tends to increase "when translation is carried out from a 'major' or highly prestigious language/culture, especially if the target language/culture is 'minor' or 'weak' in any sense." Most L1 → L2 translation is produced from languages of low or limited diffusion (LLDs) into the world's major languages, mainly English as a *lingua franca*—a global language without native speakers. Many LLDs are also low resourced languages where access to translation technologies, including bilingual corpora, is limited and cannot assist L1 → L2 translators.

This article is an attempt to assess what we think we know about L1 → L2 translation and interpreting, and what we do not know about L2 → L1 translation. Section 1.1 reviews some major findings on the effects of directionality within the cognitively oriented TPR. Section 1.2 looks at the contentious issue of L1 → L2 translation quality, which has rarely been dealt with in TPR studies. Sections 3–5 report on research on the effects of directionality in the process and its ripple effect on the quality of translation as an end product.

1.1 What we think we know about L1 → L2 translation and interpreting

Most TPR studies have focused on investigating translators who work from their learned language—L2, mostly English, into their native language—L1, with Danish, Spanish and German as probably the most researched target languages. The number of studies comparing the process of L2 → L1 and L1 → L2 translation and

interpreting is limited, but their findings are valuable. In most studies on directionality effects, English is the target language.

In interpreting studies, Bartłomiejczyk (2006) found that the range and choice of strategies used by interpreting trainees differed depending on directionality. Gumul (2017, 320) compared the impact of directionality on explicitation patterns in trainee interpreters and reported that trainees tended to explicitate more when they worked into their L2 language—most likely "due to the interpreting constraints while rendering the source-text into a B language" (325). Donovan (2003) confirms that interpreting into L2 is very demanding in terms of cognitive processing for trainee interpreters. Interestingly, Gile (2005, 11) provides a hypothetical calculation of processing effort depending on directionality and concludes that the answer to the question of which direction presents a larger cognitive load depends on how much effort is needed for comprehension and production in a specific setting.

Similar to the body of research on L1 → L2 interpreting, most findings on directionality effects in translation come from TPR studies on students and the overall consensus is that translating into one's foreign language is cognitively more demanding than translating into a native language (Fonseca 2015, 123). For example, Buchweitz & Alves (2006) investigated the process of 10 trainee translators working from English (L2) into Brazilian Portuguese (L1) and vice versa. They triangulated thinking aloud, keylogging and screen capture data and showed that translating into English (L2) "required more time and, possibly, more effort, from the participants" (251) than translating into Portuguese (L1). The difference in terms of time was 30% more for the L1 → L2 direction. The analysis of the retrospective protocols revealed that, in the revision phase, the participants were mostly concerned with lexical choices (254). Further analysis revealed that the trainee translators processed smaller chunks of text when working into their L2 and that the process of writing the TT was more recursive—i.e., included more on-line revision. Ferreira (2012) replicated the study with 10 professional translators and corroborated the increased segmentation and longer translation time when working into L2.

Fonseca (2015) compared editing procedures in L1 (Brazilian Portuguese) and L2 (English) translation performed by 8 professional translators. The data included keylogging, eyetracking, questionnaires and verbal protocols. She concludes that "professional translators tend to improve their text by monitoring their translation processes during both the drafting and revision phases, regardless of the direction in which they are working" (123). Fonseca explains that the observed differences in L1 → L2 translation are not necessarily related to increased cognitive difficulty but may reflect strategic behaviour.

Other studies also challenge the assumed increased cognitive demand when translating into L2. Pavlović (2007) reported that her trainee translators (L1, Croatian; L2, English) came up with similar arguments when making decisions in both L2 → L1 and L1 → L2 translation. Her results also questioned the tacit assumption that it is easier to comprehend the meaning of STs in the L1, casting doubt on the L1 comprehension advantage for translation. Pavlović & Jensen (2009) analysed eyetracking data from four trainee translators and four professional translators working into L1 (Danish) and L2 (English) and also concluded that processing the ST can be equally demanding, irrespective of the translation direction (107). Ferreira et al (2016) reported the same findings, following eyetracking measures from a case study of four professional translators with Spanish and English. They conclude, "It seems that translators were more concerned with understanding the ST, regardless of the language, in order to be able then to convey the information in the target text" (76).

More recently, da Silva et al (2017) studied directionality effects on post-editing and translation in 18 Chinese translators with at least one year of professional experience working into and out of Portuguese (L2). They did not find significant differences in total task time between L2 → L1 and L1 → L2 translation and post-editing, and could not confirm "straightforwardly" the hypothesized impact of directionality on the eyetracking measures of cognitive effort (127). Ferreira et al (2018) investigated the metacognitive aspects of decision-making in 8 professional translators working into L1 (Spanish) and L2 (English) using retrospective protocols, and noted that the participants had more difficulty with L1 → L2 translation, especially concerning lexical decisions, pointing that "bilingual skills might play a fundamental role in translation directionality" (112).

Although the studies dealing with directionality reported interesting, albeit inconsistent, findings, most of them did not relate their findings to the quality of L2 → L1 and L1 → L2 translations (Buchweitz & Alves 2006, 264; Ferreira et al 2018).

1.2   Directionality and the problem of translation quality

The strong opposition towards professional translation being done into the translators non-native language has been fuelled by the apparent low quality of L1 → L2 translations (Ladmiral 1979, 40; Beeby 2011, 85). PACTE (2009, 227) admits that directionality may affect translation quality. Yet, translation quality is a fuzzy concept. Koby et al (2014, 414) explain this fuzziness in the following way:

> All translations carry an expectation of some degree of accuracy and fluency. Accuracy is a bilingual notion referring to the correspondence between the

> source and target text, while fluency is a monolingual notion referring to proper-
> ties of the target text such as grammar, spelling, and cohesion.

In what follows, translation quality can be assessed in terms of two constructs: adequacy/accuracy, and fluency/acceptability (Daems et al 2013), but the assessment of quality depends on whose expectations a translation should meet—those of the client, the translator or the end-user (Koby et al 2014). In effect, there is no objective measure of translation quality and the consensus is that both accuracy and fluency should be fit-for-purpose and match the expectations of end users (O'Brien 2012).

Measuring translation quality is a methodological challenge and the assessment method is often designed for a specific purpose. For example, the PACTE group (2009) used three evaluation parameters: language, meaning, and function. Most translation quality assessment (TQA) methods (overview in Martínez 2014) assume a hypothetical *golden standard*—an ideal flawless translation and use a point-off system where errors or infelicities are marked by a reviser or proof-reader. In professional translation, quality assessment is required by the ISO 17100:2015 standard for translation quality assurance because it is assumed that all translated texts need to be checked—no reference to directionality is made.

Among the few studies investigating directionality and the quality of the output, Tommola & Helvä (1998) found no significant effect of directionality on information transmission in trainee interpreters working into Finnish (L1) and English (L2). Donovan (2003, 372) reported that conference participants were not much more critical of the L1 → L2 language interpreters' performance as compared to the performance of those who translated into their L1. Chang & Schallert (2007) reported that interpreters were more accurate when working into their L1.

In translation, Pokorn (2005, 112) reported that 46 competent native-English speakers, who participated in her questionnaire study, could not agree whether the seven selected fragments of literary translations from Slovenian into English were produced by L1 → L2 translators or by L2 → L1 translators. In educational settings, Pavlović (2007) used the so called revisabilty score to assess the quality of L2 → L1 and L1 → L2 translations—a system in which three revisers were asked to mark in red, i.e., 'give' a red card to TTs' unpublishable stretches when they distorted the ST meaning or included a blatant language error. When a TT stretch was not exactly unpublishable but could benefit from improvement, the revisers would mark it in yellow (i.e., 'give' it a yellow card). Pavlović (2007, 182) reported that trainee translators scored fewer red cards in L2 → L1 direction than in L1 → L2 direction which, she reckons, was not surprising considering the participants' L2 proficiency level and translation competence. Rodríguez (2017) reported that there was a larger similarity between the translations performed

into L1 by translators and foreign language teachers than between the translations performed into L2, irrespective of the group and language pairs involved—"possibly due to the subjects' different levels of linguistic competence in foreign languages" (266). This article reports on a large scale project investigating the effects of directionality on the translation process and quality of L2 → L1 and L1 → L2 TTs.

2.   The EDiT project[1]

2.1   Participants, materials, procedure

Thirty professional translators, who regularly translate from English (L2) into Polish (L1) and vice versa, participated in the study and produced 26 data sets fit for analysis (4 sets were discarded as incomplete). The inclusion criteria required at least 3 years of professional experience and a minimum of 50 pages of translated text per month. The participants were remunerated for their work.

The materials included a series of tests to measure language dominance—a verbal fluency test and a typing test in both languages (details in Whyatt 2018) and four short STs (ca. 162 words each)—two for each direction. The STs for each translation direction (see Appendix) belonged to different text types according to their predominant function (Reiss 1976): descriptive (a product description) and expressive (a film review). This was crucial to examine whether text type modulated the effects of directionality—an observation made in some studies on directionality (Pavlović 2007, Ferreira et al 2018). In sum, the participants translated 324 words into L1 and 326 words into L2, combined into 15 sentences in each direction, with an average number of 21.6 words per sentence. The readability score measured by the Gunning Fog text scale was 14.1 for L2 → L1 and 14.2 for L1 → L2 direction. Translators came for individual sessions and spent on average two hours to do the experimental tasks. The translation direction was counterbalanced and the texts were randomised to minimise task order effects (Mellinger & Hanson 2018).

2.2   Data collection tools and methods

The performance of participants was keylogged with Translog II (Carl 2012), eye-tracked with EyeLink 1000 Plus and screen-recorded with Morae in the same conditions for the participants, including a translation brief, lighting, access to Internet resources and privacy while they worked. Having completed the translation in one direction, the participants took a short break before translating in the other direction. At the end, they filled in a short questionnaire about their professional experience, directionality preferences, and proportion of their commissions into and out of their L1. They were also asked to rate the perceived difficulty of the STs using a 1 to 5 Likert scale.

All TTs were proofread by experienced proofreaders whose L1 was the target language. To improve reliability, there were two proofreaders for each direction. They were instructed to correct the TTs where they felt it was necessary, so that they could be published. There were no other specific guidelines. The proofreaders did not have access to the STs and they assessed the acceptability/fluency aspect. They were asked to indicate the time when they started and finished correcting each TT. Whenever they came across a stretch of text which they felt they could not correct because the meaning was not clear, they were asked to leave a comment. The adequacy/accuracy of the translations was assessed by an experienced professional translator who highlighted TT stretches which departed from the ST meaning in a way that could not be justified by the purpose of the TT. The corrections and comments made by the proofreaders and the adequacy evaluator were converted into penalty points scores–1 point for a minor error or infelicity was given to a correction which was judged as not impeding meaning construal and 5 penalty points were given to corrections which gravely distorted the ST meaning, or made the meaning construal by a target reader impossible (sense/nonsense problems). All corrections entered by the proofreaders were classified into the following categories: spelling, typography, punctuation, vocabulary, grammar, style, cohesion, and cultural adaptation. Penalty points for adequacy and fluency were added to produce a score of overall quality—the higher the number of penalty points, the lower the translation quality.

2.3   Research questions and data analysis

The aim of the study was to measure the effects of directionality on total task time and the quality of the products, and zoom in on the process using a case-study approach to compare the performance of two translators who produced the best quality L1 → L2 translation. Five research questions will guide the data analysis:

RQ1. Is the L1 → L2 translation more time consuming and thus globally more effortful that L2 → L1 translation, and is task time modulated by text type?

RQ2. Do the proofreaders correcting L1 → L2 translations need more time to make them publishable than proofreaders correcting L2 → L1 translations, and is the allocation of time modulated by text type?

RQ3. Do L1 → L2 translations score more penalty points than L2 → L1 translations and is the number of penalty points modulated by text type?

RQ4. Which aspects of text fluency are significantly influenced by the translation direction and are they modulated by the text type?

RQ5. Is the performance of translators who produce high quality L1 → L2 translation affected by directionality and is the effect modulated by text type?

RQ1 addresses the frequent finding that L1 → L2 translation is more time consuming than L2 → L1 translation (Buchweitz & Alves 2006; Ferreira 2012). The total task time was recorded by the keylogging program in seconds. RQ2 looks at the amount of time needed to correct the translations in both directions as an index of their quality—more time invested in making the translations publishable indicates lower translation quality. RQ3 also evaluates quality but now the number of penalty points scored by each translation is an index of its quality—the more penalty points, the lower translation quality. RQ4 is more qualitative and asks in detail about the nature of problems which were corrected by the proofreaders in order to see which areas of L1 and L2 were significantly affected by directionality. Areas problematic for professional translators could inform translation training programs and draw attention to directionality issues. RQ5 adds to the quantitative and qualitative aspects targeted in RQ 1–4 by zooming in on the process/product relationship in a case study of two participants who produced high quality L1 → L2 translations. Their performance is operationalized by selected measures used in TPR studies (total task time, dwell time when reading the ST for the first time, typing speed, number of long pauses, time for end revision) and compared for both directions.

To analyse the data and assess whether the effect of directionality is significant, linear mixed models (LMM) were used for inferential statistics with translation direction and text type as fixed factors and participants and proofreaders as random factors. As observed by Mellinger & Hanson (2018), LMMs are able to account for the individual differences between participants—in this study between the translators and proofreaders.

3.   Results

RQ1. No statistically significant effects of translation direction and text type were found on total task time. The interaction of translation direction and text type did not reach statistical significance. Table 1 shows the descriptive statistics for total task time in seconds for L2 → L1 and L1 → L2 direction, the two text types: film review (REVIEW) and product description (PRODUCT), and the interaction between translation direction and text type.

Table 1. Total task time—no significant effect of direction and text type

Effect	Factor level	Factor level	N	Mean	SD	SE
	Descriptive statistics for Task duration in seconds					
direction	L2 → L1		52	1210.538	405.0657	56.17250
	L1 → L2		52	1195.808	420.8292	58.35852
text type	review		52	1238.885	395.2679	54.81379
	product		52	1167.462	427.1402	59.23369
dir*type	L2 → L1	review	26	1283.154	381.4176	74.80215
		product	26	1137.923	422.2231	82.80476
dir*type	L1 → L2	review	26	1194.615	411.2993	80.66243
		product	26	1197.000	438.3016	85.95802

If we take total task time—Kring's (2001) temporal effort—as a global measure of cognitive effort, translating into L2 is not more effortful than translating into L1 for the participants in this study.

RQ2. In raw numbers, two L1 proofreaders took 512 minutes to correct the texts and the L2 proofreaders took 455 minutes to correct the L1 → L2 translations. No statistically significant effects of translation direction and text type were found. The interaction of translation direction and text type reached statistical significance ($b = -2.58$, $SE = 0.54$, $t = -4.71$, $p < 0.001$). Post-hoc tests (Bonferroni correction) for the statistically significant interaction effect showed that the L1 proofreaders spent significantly more time (at $p < 0.05$) correcting film reviews ($M = 6.48$) than the L2 proofreaders ($M = 4.50$). Additionally, the L1 proofreaders spent significantly more time correcting film reviews ($M = 6.48$) than product description texts ($M = 3.65$) at $p < 0.001$. The results show that the effect of translation direction is modulated by text type.

RQ3. In raw numbers 1,458 penalty points were scored by L2 → L1 translations and 1, 864 by L1 → L2 translations. As explained earlier on, the total score included penalty points for lack of adequacy, cultural adaptation and for text fluency (language problems). No statistically significant effects of translation direction and

text type were observed. The interaction effect of translation direction and text type reached statistical significance ($b=-4.46$, $SE=1.46$, $t=-3.06$, $p<0.01$). Post-hoc tests for the statistically significant interaction of effects showed that the total penalty score as a measure of overall text translation quality was significantly lower ($p<0.01$) for product description texts translated into L1 ($M=12.02$) than film reviews translated into L1 ($M=15.5$).

RQ4. Four significant effects of directionality were found in all categories into which corrections were classified. Table 2 summarizes the significant effects of directionality, text type and interaction between directionality and text type on all categories which affect text fluency (Koby et al 2014, 414).

Directionality had a statistically significant effect on grammar, punctuation, typographical corrections and a sense/nonsense category—the stretches of text that proofreaders could not correct because the sense was unclear. There were more corrections improving grammar in L1 → L2 translations, but punctuation, typographical and sense/nonsense problems were prevalent in L2 → L1 translations. Text type also had a significant effect on grammar with more problems in film reviews than in product description texts. The effect of interaction between directionality and text type was also significant for grammar corrections and confirmed that directionality effect is modulated by text type—both text types needed more corrections in L1 → L2 translation, but film reviews translated into L1 had more corrections than the much easier and more formulaic product description texts.

Table 2. Statistically significant effects of directionality, text type and directionality/text type interaction in all categories relevant for translation quality (fluency aspect)

Category	Directionality	Text type	Interaction
Spelling	–	more problems in film reviews ($b=-0.62$, $SE=0.17$, $t=-3.53$, $p=0.001$)	–
Typography	more problems in L2 → L1 translation ($b=0.60$, $SE=0.20$, $t=2.97$, $p<0.01$)	–	more problems in L1 product description texts and film reviews than in L2 more problems in L1 product description texts than in L1 film reviews ($b=0.82$, $SE=0.28$, $t=2.89$, $p<0.01$)
Punctuation	more problems in L2 → L1 translation	–	–

Table 2. *(continued)*

Category	Directionality	Text type	Interaction
	($b = 2.13$, $SE = 0.51$, $t = 4.16$, $p < 0.05$)		
Vocabulary	–	more problems in product description texts ($b = 2.35$, $SE = 0.40$, $t = 5.89$, $p < 0.001$)	more problems in L2 product description texts than in L1 more problems in L2 → L1 translation of film reviews more problems in L2 product description texts than in L2 film reviews ($b = -3.73$, $SE = 0.56$, $t = -6.62$, $p < 0.001$)
Sense/ nonsense	more problems in L2 → L1 translation (marginally significant effect) ($b = 1.92$, $SE = 0.67$, $t = 2.86$, $p = 0.058$)	–	more problems in L2 → L1 translation of film reviews ($b = -1.25$, $SE = 0.63$, $t = -1.99$, $p < 0.05$)
Grammar	more problems in L1 → L2 translation ($b = -2.46$, $SE = 0.50$, $t = -4.96$, $p < 0.05$)	more problems in film reviews ($b = 1.02$, $SE = 0.34$, $t = 3.03$, $p < 0.01$)	more problems in L2 product descriptions and film reviews more problems in L2 → L1 translation of film reviews ($b = -2.12$, $SE = 0.48$, $t = -4.45$, $p < 0.001$)
Style	–	–	more problems in L2 → L1 translation of film reviews ($b = -1.85$, $SE = 0.47$, $t = -3.95$, $p < 0.001$)
Cohesion	–	more problems in film reviews ($b = -0.83$, $SE = 0.22$, $t = -3.75$, $p < 0.001$)	–

The more prevalent number of typographical and punctuation corrections when translating into L1 can be explained by the assumed L1 production advantage (faster text production, hence more typos and careless punctuation). However, the significantly higher penalty scores for sense/nonsense errors in L2 → L1 translation are a matter of concern, even though the effect at $p = 0.058$ was only marginally significant.

No statistically significant main effect of translation direction was observed on the number of penalty points classified as corrections to spelling, vocabulary, and cohesion. In these categories, text type had a statistically significant effect on the number of corrections. Translators had more problems with vocabulary

use when they worked on product description texts (especially into L2), possibly because more lexical precision is required in such texts than in film reviews. But when translating into L1, they had more vocabulary problems in film reviews than in the more formulaic product description text. Film reviews also were more problematic in terms of spelling and cohesion, irrespective of the translation direction (no significant interaction effects). Penalty points for stylistic errors showed a combined effect of directionality and text type—the effect of translation direction and text type was significant only for L2 → L1 translations, with L1 film reviews showing more corrections to style than the product description texts.

RQ5. Translator P29 received the lowest penalty score for the L2 film review and P06 for the L2 product description text from at least one L2 proofreader, therefore they were selected to establish whether the performance of translators who produce high quality L1 → L2 translation differs depending on directionality. Both translators declared in their questionnaires that they did not have directionality preferences. P06 estimated that translating into L2 (English) amounts to 30 to 50 per cent of the entire workload and for P29, L1 → L2 translation constituted more than 50 per cent of the workload. P06 had 18 years of professional experience and P29 has worked as a translator for 15 years. Table 3 shows the penalty scores from the L1 and L2 proofreaders. Interestingly, both translators also produced high quality L2 → L1 translations, although the discrepancy between the proofreaders' corrections is visible.

Table 3. Overall quality scores per direction and text type based on Proofreaders 1 and 2 (P1//P2)—the lower the number of penalty points, the higher the overall quality

Quality	L1 → L2		L2 → L1	
	Review	Product	Review	Product
P29	7//11	10//11	5//2	15//15
P06	9//19	8//16	9//26	7//8
mean score	17	18	15.5	12

The mean score for all participants is given as a reference point. Although both translators produced high quality L1 → L2 translation the difference in their total task time per direction and text type is substantial (Table 4). The mean time for all the participants is given as a reference point.

P06 was much closer to the mean average for all texts but P29, although taking considerably more time, did so in L2 → L1 and L1 → L2 directions. However, when the time for both texts in each direction is added, it turns out that both translators took more time translating into their L1. The difference in the task time is reflected in their performance measures throughout the entire translation process. Table 5

Table 4. Total task duration in minutes and seconds

Task duration	L1 → L2		L2 → L1	
	Review	Product	Review	Product
P29	35:53	30:53	40:02	40:10
P06	21:22	10:18	24:53	15:20
Mean	**19:55**	**19:57**	**21:23**	**18:58**

shows dwell time as recorded by an eyetracker—the overall amount of time the translator's eyes focus on the ST when reading it prior to translation.

Table 5. Dwell time on the source text in seconds

ST dwell time	L1 → L2		L2 → L1	
	Review	Product	Review	Product
P29	23	19	33	29
P06	40	50	59	51
mean	**45.5**	**48**	**53**	**44**

The eyetracking measures show that P06 is much closer to the mean dwell time for all the participants and there is no substantial difference dependent on directionality or text type. P29 spent much less time becoming familiar with all the STs. Further comparisons confirm that both translators had different working styles throughout the entire text production process. P29 worked slowly and P06 was a much faster translator. Table 6 shows the typing speed measured by the number of total user events per minute, as recorded in Translog II.

Table 6. Typing speed in total user events per minute

Typing speed	L1 → L2		L2 → L1	
	Review	Product	Review	Product
P29	47.37	46.44	58.94	47.53
P06	80.05	128.26	77.7	104.62
mean	**106**	**95**	**100**	**105**

P06 typed the translation considerably faster when translating product description texts irrespective of the direction, implying that the effect of text type was much stronger than the effect of directionality. P29 worked more or less half the average speed, irrespective of the direction or the text type. A closer look at the number of long pauses measured at two thresholds—longer than 5 seconds and longer than 10 seconds, shows that the typing speed was affected by the need to pause, most likely for problem-solving. Table 7 shows that P29 needed nearly twice as many

pauses longer than 10 seconds as P06, whose measures were much closer to the mean average for all the translators who participated in the experiment.

Table 7. Number of long pauses measured at two thresholds—pauses >5s//pauses >10s

	L1 → L2		L2 → L1	
Pauses >5s//>10s	Review	Product	Review	Product
P29	65//42	46//26	73//45	50//26
P06	41//17	31//17	51//23	20//10
mean	35//18	35//19	42//21	27//14

Interestingly, the most effortful task in terms of long pauses was the L2 → L1 translation of the film review for both translators. The last performance measure in this study is time for end revision (Table 8), as the last effort taken to ensure a good quality translation product.

Table 8. Time for end revision in seconds

	L1 → L2		L2 → L1	
	Review	Product	Review	Product
P29	325	200	226	527
P06	256	99	134	167
mean	240	241	244	235

P06 spent much less time on end revision than P29 and much less time than the mean average for all participants, when revising the product description texts—with the L1 → L2 direction taking less time than the L2 → L1 direction. P29 interestingly spent more time on end revision when working into L1. The differences in the performance of the two translators who produced high quality L1 → L2 translations are corroborated by their self-rating of translation difficulty (range 1 to 5) in Table 9.

Table 9. Self-assessment of translation difficulty

	L1 → L2		L2 → L1	
Translation difficulty	Review	Product	Review	Product
P29	3	4	3	3
P06	3	2	4	1
Mean	2.9	3.1	3.4	2.3

While P29 rated all the texts as similar in terms of translation difficulty, P06 rated the product description texts as less difficult than the film reviews, irrespective of directionality—possibly due to more extensive experience with translating

descriptive texts. Thus, the level of familiarity with the text type might override the assumed directionality effect.

4.   Discussion

The findings presented in Section 3 show complex directionality effects on the translation process and end product, and question the simplified assumption that L1 → L2 translation is more demanding than L2 → L1 translation and that its quality is far inferior. Directionality effects interplay with text-type effects, and possibly with the translators' individual experience and working style. For the 26 professional translators, directionality does not have a statistically significant effect on the overall task time. This finding is not in line with Buchweitz & Alves (2006) and Ferreira (2012), although the latter study investigated translation students. The difference in the findings can be explained by many factors, including the number of participants, language-pair specificity and the length or intensity of professional experience in providing L1 → L2 translation. Similar justification was offered by da Silva et al (2017) who did not find differences either between temporal effort in L2 → L1 and L1 → L2 translation and postediting.

A central contribution of this study should be bringing the contentious quality issue into the discussion of directionality effects. Research questions 2–4 tackled the alleged divide between L2 → L1 and L1 → L2 translation quality by measuring time needed to correct the translated texts by experienced proofreaders (RQ2), by comparing the penalty points scored for all the corrections (RQ3) and by analysing the effect of directionality on the acceptability/fluency of the translated texts (RQ4). Such a comprehensive approach to quality assessment is guided by the concern to see translation quality assurance as a collaborative objective—a matter of cooperation between translators and revisers, proofreaders and editors (ISO 17100:2015). All TTs, irrespective of directionality and text type, needed correcting, and surprisingly the L1 proofreaders spent nearly an hour more (57 minutes) altogether to make the L2 → L1 translations publishable. Still, the directionality effect was not significant and neither was the effect of text type—most likely due to large individual variation between the quality of translations and the discrepancies in the proofreaders' corrections. The interaction between directionality and text type proved significant. Correcting the L2 → L1 film review took significantly more time than correcting the L1 → L2 translation of the same text type, and more time than correcting the more formulaic product description text. This result points to text type as a factor modulating the effect of directionality. Another factor might be the translator's experience and the degree of familiarity with certain text types. Product-description texts are

most likely to be frequently commissioned to translators, whereas film reviews might be less frequent and therefore require more processing effort. Additionally, film reviews in English are often more elaborate in terms of sophisticated vocabulary and syntactic complexity than Polish reviews. It therefore seems essential to prepare translation students to recognise text-typological differences between languages as they might contribute to higher levels of translation difficulty more than directionality.

A very detailed analysis of the impact of directionality on the language categories which affect the quality of translation (the fluency aspect) revealed a complex interplay between translation direction and text type and/or interaction between both—with the pendulum swinging to either L2 → L1 or L1 → L2 translation as more likely to show the need for correction. Directionality had the most significant effect on grammar in L1 → L2 translations, but it also interacted with text type (in L2 → L1 translation of a film review). Interestingly, the significant effect of directionality on punctuation and typographical errors showed that they needed more corrections in L2 → L1 translation—possibly a side effect of the faster typing speed and lower self-monitoring in L1 direction. The translators might have paid less attention to typos also because they are automatically alerted to typos by a spell checker in their everyday practice—when working in Translog II this function is not available. Problems with L1 punctuation cannot be as easily explained and require a more thorough investigation looking into possible cross-linguistic interference.

Interestingly, directionality did not have a significant effect on vocabulary, style and cohesion but it strongly interacted with text type when translators were selecting lexical items and deciding whether the TT was stylistically appropriate for its purpose. The higher penalty scores for vocabulary and cohesion were significantly caused by text type as a factor.

The most surprising finding was revealed by comparing the penalty points scored for the comments left by the proofreaders when they could not correct a stretch of text because they did not understand what the translator had in mind. Such errors were more common in the L2 → L1 translation of a film review (self-assessed as a more difficult task in Table 9), although the significance level was marginal. An explanation can be found in the cognitive constraints of the translation process (limited processing capacity)—Gile's (2005, 11) observation that production may be compromised if comprehension of the ST requires a larger cognitive effort. Another explanation can be due to the procedural and declarative nature of L1 and L2 knowledge. When translating more complex stretches of text from L2 (the more explicitly encoded knowledge), switching into the more implicitly encoded knowledge of L1 might require more effort and more efficient self-monitoring to produce a desirable effect. Still, the translators who participated

in the study worked in lab conditions, where they could not count on 'drawer time'—that is, being able to revise the translated text after some time. A more detailed analysis of how these and other errors surfaced in the process and made their way to the end product is necessary to understand the effect of directionality.

Finally, the case study of two translators who produced high quality translations shows that for them directionality does not have a clearly visible effect on the selected measures of their performance, such as total task time, dwell time in ST, typing speed, number of long pauses and time for end revision. Their translation behaviour shows that familiarity with the text type and personal working style as a product of their professional experience are more decisive factors than directionality. Whether this finding is also true for the translators whose L1 → L2 translation needed more improvement remains to be investigated.

5.   Conclusions and further research

The research presented in this article is not without limitations. Despite the rigorous design and methodology, the texts selected for the experiment presented various degrees of difficulty for the participants and their translations might have been assessed with a different *golden standard* in the proofreaders' mind. Still, the results have statistical power to show that directionality effects question the tacit assumption about L1 → L2 translation—what we think we know. The findings also show that what we do not know about L2 → L1 translation, or have failed to mention, is that it can be equally demanding, equally unsuccessful and in need of improvement. For the translators in the study, L1 → L2 translation was not cognitively more demanding in terms of total task time than L2 → L1 translation. It was not less successful than L2 → L1 translation, as all the texts needed to be improved by the proofreaders. Indeed, L1 → L2 translation abounds in language mistakes but so does L2 → L1 translation. Understandably in line with the unequal language proficiencies in L1 and L2, the nature of problems differs depending on the direction in which translation is done. A more detailed analysis of the directionality effects observed in this study is in order to see if these are language-pair specific or whether some of them occur in other language combinations. Most importantly, the study speaks for the need to reassess and re-embed the thorny issue of directionality in cognitive translation studies. A more profound understanding of how directionality affects both the process and the end product can inform the training of future bidirectional translators by helping them optimise their bilingual resources and recognise the error-prone nature of the translation process.

Acknowledgements

References

Bartłomiejczyk, Magdalena. 2006. "Strategies of Simultaneous Interpreting and Directionality." *Interpreting* 8(2): 149–174. https://doi.org/10.1075/intp.8.2.03bar

Beeby Lonsdale, Allison. 2011. "Directionality." In *Routledge Encyclopedia of Translation Studies.* 2nd edition. Edited by M. Baker, and G. Saldanha, 84–88. London: Routledge.

Buchweitz, Augusto, and Fabio Alves. 2006. "Cognitive Adaptation in Translation: an Interface between Language Direction, Time and Recursiveness in Target Text Production." *Letras de Hoje* 41(2): 241–272.

Carl, Michael. 2012. "Translog-II: A Program for Recording User Activity Data for Empirical Reading and Writing Research." *LREC*: 4108–4112.

Chang, Chia-chien, and Diane L. Schallert. 2007. "The Impact of Directionality on Chinese/English Simultaneous Interpreting. *Interpreting* 9(2): 137–176. https://doi.org/10.1075/intp.9.2.02cha

Chmiel, Agnieszka. 2018. "Meaning and Words in the Conference Interpreter's Mind." *Translation, Cognition & Behavior* 1(1): 21–41. https://doi.org/10.1075/tcb.00002.chm

Daems, Joke, Lieve Macken, and Sonia Vandepitte. 2013. "Quality as the Sum of its Parts: A Two-step Approach for the Identification of Translation Problems and Translation Quality Assessment for HT and MT+PE." In *Proceedings of MT Summit XIV Workshop on Post-editing Technology and Practice*. Edited by S. O'Brien, M. Simard, and L. Specia, 63–71.

da Silva, Igor A., Fabio Alves, Marcia Schmaltz, Adriana Pagano, Derek Wong, Lidia Chao, Ana L. Leal, Paulo Quaresma, Caio Garcia, and Gabriel E. da Silva. 2017. "Translation, Post-Editing and Directionality: A Study of Effort in the Chinese-Portuguese Language Pair." In *Translation in Transition*. Edited by A. L. Jakobsen, and B. Mesa, 108–134. Amsterdam: John Benjamins. https://doi.org/10.1075/btl.133.04lou

Donovan, Clare. 2003. "Teaching Simultaneous Interpretation into B." In *La Direccionalidad en Traducción e Interpretación. Perspectivas Teóricas, Profesionales y Didácticas*. Edited by D. Kelly, A. Martin, M.-L. Nobs, D. Sánchez, and C. Way, 367–380. Granada: Editorial Atrio.

Ferreira, Aline. 2012. "Investigando o processamento cognitivo de tradutores profissionais em tradução direta e inversa no par linguístico inglês português." *Cadernos de Tradução* 29: 73–92.

Ferreira, Aline, and John Schwieter. 2017. "Directionality in Translation." In *The Handbook of Translation and Cognition*. Edited by J. W. Schwieter, and A. Ferreira, 90–105. New York: John Wiley & Sons. https://doi.org/10.1002/9781119241485.ch5

Ferreira, Aline, John W. Schwieter, Alexandra Gottardo, and Jefferey Jones. 2016. "Cognitive Effort in Direct and Inverse Translation Performance: Insight from Eye-tracking Technology." *Cadernos de Tradução* 36(3): 60–80.
https://doi.org/10.5007/2175-7968.2016v36n3p60

Ferreira, Aline, Alexandra Gottardo, and John W. Schwieter. 2018. "Decision-making Processes in Direct and Inverse Translation through Retrospective Protocols." *Translation, Cognition & Behavior* 1(1): 98–118. https://doi.org/10.1075/tcb.00005.fer

Fonseca, Norma B. 2015. "Directionality in Translation: Investigating Prototypical Patterns in Editing Procedures", *Translation & Interpreting* 7(1): 111–125.

Gile, Daniel. 2005. "Directionality in Conference Interpreting: a Cognitive View." In *Directionality in Interpreting. The 'Retour' or the Native?* Edited by R. Godijns, and M. Hinderdael, 9–26. Ghent: Communication and Cognition.

Gumul, Ewa. 2017. "Explicitation and Directionality in Simultaneous Interpreting." *Linguistica Silesiana* 38: 311–329.

Koby, Geoffrey S., Paul Fields, Daryl Hague, Arle Lommel, and Alan Melby. 2014. "Defining Translation Quality." *Tradumàtica* 12: 413–420. https://doi.org/10.5565/rev/tradumatica.76

Krings, Hans P. 2001. *Repairing texts: Empirical Investigations of Machine Translation Post-editing Process*. Kent, OH.: The Kent State University Press.

Ladmiral, Jean-René. 1979. *Traduire: Theorems pour la Traduction*. Paris: Payot.

Martínez Mateo, Roberto. 2014. "A Deeper Look into Metrics for Translation Quality Assessment (TQA): A Case Study. *Miscelánea: A Journal of English and American Studies* 49: 73–94.

Mellinger, Christopher D., and Thomas A. Hanson. 2018. "Order Effects in the Translation Process." *Translation, Cognition & Behavior* 1(1): 1–20. https://doi.org/10.1075/tcb.00001.mel

Newmark, Peter. 1981. *Approaches to Translation*. Oxford: Pergamon Press.

O'Brien, Sharon. 2012. "Towards a Dynamic Quality Evaluation Model for Translation." *The Journal of Specialised Translation* 17: 55–77.

PACTE. 2009. "Results of the Validation of the PACTE Translation Competence Model: Acceptability and Decision-making." *Across Languages and Cultures* 10(2): 207–230. https://doi.org/10.1556/Acr.10.2009.2.3

Paradis, Michel. 2009. *Declarative and Procedural Determinants of Second Languages*. Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.40

Pavlović, Natasa. 2007. *Directionality in Collaborative Translation Processes. A Study of Novice Translators*. PhD Dissertation. Tarragona: Universitat Rovira i Virgili.

Pavlović, Natasa, and Kristian T. H. Jensen. 2009. "Eyetracking Translation Directionality." In *Translation Research Projects* 2. Edited by A. Pym and A. Perekrestenko, 93–109. Tarragona: Intercultural Studies Group.

Pokorn, Nike. 2005. *Challenging the Traditional Axioms: Translation into a Non-mother Tongue*. Amsterdam: John Benjamins. https://doi.org/10.1075/btl.62

Reiss, Katharina. 1976. *Texttyp und Übersetzungsmethode*. Kronberg: Scriptor.

Rodríguez Inés, Patricia. 2017. "Analysis of the Translation Competence Corpus from PACTE's Experiment." In *Researching Translation Competence by PACTE Group*. Edited by A. Hurtado, 243–266. Amsterdam: John Benjamins. https://doi.org/10.1075/btl.127.16rod

Talaván Zanón, Noa, and Pilar Rodríguez Arancón. 2014. "The Use of Reverse Subtitling as an Online Collaborative Language Learning Tool." *The Interpreter and Translator Trainer* 8(1): 84–101. https://doi.org/10.1080/1750399X.2014.908559

Tommola, Jorma, and Marketta Helvä. 1998. "Language Direction and Source Text Complexity: Effects on Trainee Performance in Simultaneous Interpreting. In *Unity in Diversity? Current Trends in Translation Studies*. Edited by L. Bowker, M. Cronin, D. Kenny, and J. Pearson, 177–186. Manchester: St. Jerome.

Toury, Gideon. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam: John Benjamins. https://doi.org/10.1075/btl.4

Traxler, Matthew J. 2012. *Introduction to Psycholinguistics: Understanding Language Science*. New York: John Wiley & Sons.

Whyatt, Bogusława. 2018. "Old Habits Die Hard: Towards Understanding L2 Translation." *Między Oryginałem a Przekładem* 24(41): 89–112. https://doi.org/10.12797/MOaP.24.2018.41.05

Whyatt, Bogusława, and Tomasz Kosciuczuk. 2013. "Translation into a Non-native Language: The Double Life of the Native-speakership Axiom." *mTm Translation Journal* 5: 60–79.

Wu, Yinyin and Posen Liao. 2018. "Re-conceptualising Interpreting Strategies for Teaching Interpretation into a B Language." *The Interpreter and Translator Trainer* 12(2): 188–206. https://doi.org/10.1080/1750399X.2018.1451952

Appendix.   Source texts used in the experiment

L2 → L1 translation

Circulating fans include ceiling fans, table fans, floor fans, and fans mounted to poles or walls. These fans create a wind chill effect that will make you more comfortable in your home, even if it's also cooled by natural ventilation or air conditioning. Ceiling fans are considered the most effective of these types of fans, because they effectively circulate the air in a room to create a draft throughout the room. If you use air conditioning, a ceiling fan will allow you to raise the thermostat setting about 4°F with no reduction in comfort. In temperate climates, or during moderately hot weather, ceiling fans may allow you to avoid using your air conditioner altogether. Install a fan in each room that needs to be cooled during hot weather. Turn off ceiling fans when you leave a room; fans cool people, not rooms, by creating a wind chill effect. Ceiling fans are only appropriate in rooms with ceilings at least eight feet high.

Silence ends Martin Scorsese's decades-long quest with a thoughtful, emotionally resonant look at spirituality and makes no apologies for its solemn tone and the uncompromising way it grapples with the almighty. But those willing to go along on Scorsese's arduous journey will find themselves eventually rewarded. The drama plays out like an anguished summation of the 74-year-old filmmaker's career-long obsession; the inherent paradoxes of living spiritually in a world beset with earthly temptations. Uneven, sometimes repetitive but also powerfully moving and thought-provoking, Silence is an imperfect movie that's very hard to shake off. Opening in late December in the US and early 2017 in the UK, Silence will be an event movie for cinema goers. The curiosity is further stoked by the fact that the director has wanted to make a film of the Shūsaku Endō novel for nearly three decades. Star power from Liam Neeson will help, but a lengthy viewing time and difficult subject matter will naturally limit commercial appeal.

L1 → L2 translation

Smart Mop Plus rewelacyjnie czyści na sucho i mokro, świetnie nadaje się do sprzątania mebli oraz okien. Nasz Smart Mop plus to nie tylko wysokiej jakości technologia mikrowłókien wyko-

rzystanych w główce mopa, ale także zastosowanie 2-fazowego wirowania, które jest wyko- rzystane dzięki zintegrowanemu systemowi w rączce mopa. Smart Mop, wyposażony jest w specjalny system odwirowywujący wbudowany na dnie wiaderka, który ma zadanie oczyszczać główkę mopa z brudu już w wodzie, natomiast rozkładany koszyk z wirownią odwirowuje mop do suchej nitki! Aby wprawić główkę mopa w ruch, wystarczy umieścić mopa w odpowiedniej komorze, odblokować drążek i przycisnąć parę razy. Proces zachodzi praktycznie bez użycia żadnej siły. Dzięki temu sprzątanie staje się lekkie i bardzo przyjemne. Zastosowanie rewelacyj- nego systemu składania wiaderka do kompaktowych rozmiarów jest nieocenionym atutem. Za pomocą jednego przycisku i ruchu ręką nasz mop złoży się w nieduży, estetyczny sześcian, który nie będzie zawadzał w Twoim domu. Jego wytrzymała konstrukcja wytrzymuje obciążenie do 150 kg, więc możesz go wykorzystać nawet jako podstawkę lub taborecik.

Film Powidoki prezentuje osadzoną w latach 50. ubiegłego wieku w Łodzi historię Władysława Strzemińskiego. Poznajemy postać artysty, malarza i wykładowcy na łódzkiej uczelni, który musi stawić czoła nowej rzeczywistości PRL. Wajda opowiada tę historię w sposób prosty, oczywisty i trochę idealizuje postać Strzemińskiego. Z jednej strony wydaje się, że za bardzo akcentuje jego niezłomność i bunt przeciwko komunistom, którzy wchodzą z butami w kulturę, ale z drugiej czuć w tym pewnego rodzaju hołd dla tego człowieka. Reżyser podziwia to, jak malarz za wszelką cenę nie ugiął się pod wszelkimi naciskami ustroju, dla którego nie był jednym z najważniejszych artystów XX wieku, ale kimś, kto zwyczajnie przeszkadza. Reżyserowi udaje się klarownie opo- wiedzieć historię, która stanowi zaledwie wycinek z życia Strzemińskiego, dlatego też nie jest to tradycyjny film biograficzny, w którym możemy dokładnie poznać tę postać i dowiedzieć, się, dlaczego warto było stworzyć o niej tę fabułę. Wajda potępia też ustrój komunistyczny, pokazu- jąc go jako bezmyślny twór, który zniszczy wszystko dobre, co stanie mu drodze.

Address for correspondence

Bogusława Whyatt
Department of Psycholinguistic Studies
Wydział Anglistyki UAM
Adam Mickiewicz University
al. Niepodległości 4
61 874 Poznań
Poland

bwhyatt@wa.amu.edu.pl

Biographical notes

Bogusława Whyatt is an associate professor and head of Department of Psycholinguistic Stud- ies at the Faculty of English, Adam Mickiewicz University in Poznań, Poland. She teaches sem- inars in translation studies and practical translation courses. Her research interests include translation process, directionality, bilingual memory and development of translation expertise. She is also a freelance translator.

https://orcid.org/0000-0003-1645-1458