

## Evidence of Parallel Processing During Translation

Laura Winther Balling, Kristian Tangsgaard Hvelplund et Annette C. Sjørup

Volume 59, numéro 2, août 2014

URI : <https://id.erudit.org/iderudit/1027474ar>

DOI : <https://doi.org/10.7202/1027474ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (imprimé)

1492-1421 (numérique)

[Découvrir la revue](#)

Citer cet article

Balling, L. W., Hvelplund, K. T. & Sjørup, A. C. (2014). Evidence of Parallel Processing During Translation. *Meta*, 59(2), 234–259.  
<https://doi.org/10.7202/1027474ar>

Résumé de l'article

Dans trois études oculométriques, nous testons l'hypothèse selon laquelle la traduction suppose un traitement parallèle plutôt que séquentiel des textes source et cible. Dans la première étude, un groupe de traducteurs professionnels ont traduit vers l'anglais des textes rédigés dans leur langue maternelle, c'est-à-dire le danois. Les textes comprenaient aussi bien des segments où l'ordre verbe/sujet était le même entre texte source et texte cible que des segments où l'ordre était différent. Nous avons relevé que les traducteurs regardaient plus longtemps les segments où l'ordre verbe/sujet différait entre les deux textes, ce qui indique une anticipation de la structure du texte cible dès la lecture du texte source. Nos deux études ultérieures sur la lecture en langue maternelle et en langue seconde démontrent deux choses : d'une part, l'ordre des mots en traduction ne découle pas du fait que les segments danois dont l'ordre diffère sont plus difficiles que les segments dont l'ordre est identique. D'autre part, cet effet n'est pas généralisé dans la lecture bilingue. Cela nous permet de conclure que la traduction est un processus parallèle et que la traduction littérale est probablement la stratégie de traduction par défaut. Cette conclusion est soutenue par le fait que les trois études étaient menées dans un cadre relativement naturel grâce à l'utilisation d'un système d'oculométrie et à une modélisation de régression à effets mixtes.

# Evidence of Parallel Processing During Translation

**LAURA WINTHER BALLING**

*Copenhagen Business School, Copenhagen, Denmark*  
lwb.isv@cbs.dk

**KRISTIAN TANGSGAARD HVELPLUND**

*University of Copenhagen, Copenhagen, Denmark*  
bnm486@hum.ku.dk

**ANNETTE C. SJØRUP**

*Copenhagen Business School, Copenhagen, Denmark*  
acs@wordplay.dk

## RÉSUMÉ

Dans trois études oculométriques, nous testons l'hypothèse selon laquelle la traduction suppose un traitement parallèle plutôt que séquentiel des textes source et cible. Dans la première étude, un groupe de traducteurs professionnels ont traduit vers l'anglais des textes rédigés dans leur langue maternelle, c'est-à-dire le danois. Les textes comprenaient aussi bien des segments où l'ordre verbe/sujet était le même entre texte source et texte cible que des segments où l'ordre était différent. Nous avons relevé que les traducteurs regardaient plus longtemps les segments où l'ordre verbe/sujet différait entre les deux textes, ce qui indique une anticipation de la structure du texte cible dès la lecture du texte source. Nos deux études ultérieures sur la lecture en langue maternelle et en langue seconde démontrent deux choses : d'une part, l'ordre des mots en traduction ne découle pas du fait que les segments danois dont l'ordre diffère sont plus difficiles que les segments dont l'ordre est identique. D'autre part, cet effet n'est pas généralisé dans la lecture bilingue. Cela nous permet de conclure que la traduction est un processus parallèle et que la traduction littérale est probablement la stratégie de traduction par défaut. Cette conclusion est soutenue par le fait que les trois études étaient menées dans un cadre relativement naturel grâce à l'utilisation d'un système d'oculométrie et à une modélisation de régression à effets mixtes.

## ABSTRACT

Three eye tracking experiments test the hypothesis that translation involves parallel rather than sequential processing of the source and target texts. In Experiment 1, a group of professional translators translated texts from their native language Danish into English. The texts included both segments where the order of verb and subject was congruent between source and target text and segments that were non-congruent. Translators gazed significantly longer at the non-congruent segments of the source text, indicating that the structure of the target text is anticipated during source text reading. Two follow-up experiments on first and second language reading demonstrate that this congruence effect in translation is not the result of the non-congruent Danish segments being inherently more difficult than the congruent ones and that the effect is not a general effect in bilingual reading. We conclude that translation is a parallel process and that literal translation is likely to be a universal initial default strategy in translation. This conclusion is strengthened by the fact that all three experiments were relatively naturalistic, due to the combination of remote eye tracking and mixed-effects regression modeling.

**MOTS-CLÉS/KEYWORDS**

processus de traduction, oculométrie, traitement séquentiel/parallèle, transfert syntaxique, lecture en L1 et en L2

translation processes, eye tracking, parallel/sequential processing, syntactic transfer, L1 and L2 reading

**1. Introduction*****1.1. Sequential and parallel processing***

A central question in translation studies (TS) has been whether translation is a sequential or a parallel process. The sequential view (Seleskovitch 1976; see also Gile 1995 and Angelone 2010) holds that source text (ST) content is decoded and deverbilised before being reformulated as target text (TT); according to this so-called deverbilisation hypothesis, the form of the ST is lost before target language processing occurs. In this view, processing would shift back and forth between ST and TT at certain intervals, each interval including processing in one of the two languages only. By contrast, the parallel view (Gerver 1976) states that ST and TT processing occur in parallel, with the consequence that purely formal aspects of the ST may be transferred to the TT.

The traditional distinction between sequential and parallel processing provides two opposing but simplified views of possible information processing routes during translation. However, the distinction falls short of capturing the complex nature of translation processing. In this paper, we will discuss and refine the traditional TS definitions of parallel and sequential processing and consider the distinction between parallel and sequential processing from a cognitive perspective. A cognitive approach to translation processing is likely to offer a richer and more nuanced basis for considering the matter of parallel and serial processing in translation and for understanding translation as a function of cognitive activity.

From a cognitive viewpoint, parallel processing in translation is when translation processes overlap, and sequential processing is when translation processes do not overlap. While it is not possible for a person to devote *attention* to two tasks at the same time, it is possible for a person to devote *cognitive resources* to two tasks at the same time where one of those tasks is at the centre of attention and the other task is performed without attentional resources being devoted to its execution. In translation, the ability to use such subconscious processing, or automaticity, has been suggested as an explanation of why skilled translators can produce translation faster and better than novice ones (Jääskeläinen and Tirkkonen-Condit 1991; Hvelplund 2011).

Based on work by Kintsch (1998) and Kellogg (1996), Hvelplund (2011: 48-58) provides an itemisation of (some of) the cognitive processes involved in translation. The stages involved in ST processing include initial orthographic analysis of the characters manifested in the ST and lexical and propositional analyses of the ST meaning. The stages involved in TT processing include identification of target language equivalents to represent the meaning of the ST words and verification of meaning congruence between the TT representation and the ST representation. Thus, in parallel processing, several processes which are relevant to the processing of the translation run simultaneously. For instance, when reading a ST word, a TT processing stream is triggered, during which possible target language realisations of

the ST word or segment are considered more or less at the same time. Conversely, in sequential processing, any consideration of possible target language realisations of the ST word or segment is postponed until comprehension of the ST segment is finalised.

Hvelplund (2011) found that translators regularly engage in simultaneous source text reading and target text typing during the translation process; though parallel processing is not necessarily restricted to situations where it is so explicitly manifested, this is clear evidence that parallel processing does actually take place, for both the student and the experienced translators involved in Hvelplund's study. Dragsted (2010) reports similar evidence of simultaneous ST reading and TT typing (which she calls integrated processing), especially for professionals. Further, the differences between reading for comprehension and reading for translation observed by Jakobsen and Jensen (2008) also point in the direction of parallel processing.

We hypothesise that parallel processing is not restricted to such explicit cases as those shown by Hvelplund (2011) and Dragsted (2010), but that anticipation of the TT is ubiquitous during source language processing. There is some support for this hypothesis in the work of Ruiz, Paredes *et al.* (2008) and Macizo and Bajo (2006) from the University of Granada. These authors present evidence that properties of words (cognateness and ST frequency) and syntactic structures (congruence of adjective-noun order) in the source language affect target language processing. The evidence comes from experiments where ST sentences were presented to translators one word at a time, using self-paced presentation, and the translators were asked to either repeat the sentence or translate it orally into English. When participants were reading source language words with the object of translating them later, the frequency of target language translations, which was manipulated in Experiment 1 of Ruiz and colleagues (2008), affected the reading of the source language words, whose frequency was kept constant. Similarly, Macizo and Bajo (2006) found effects of cognateness between source and target language words on source language reading when participants were reading for later translation, but not for normal monolingual reading. Thus, the results of the Granada group are in line with our hypothesis, but the evidence comes from a rather artificial task.

In contrast to the Granada group, we tested the hypothesis that processing is parallel in a naturalistic experimental setup, asking professional translators to translate while their eye movements were monitored. There are several advantages to this: In addition to the fact that a naturalistic task is desirable, this setup allows us to measure processing online. Further, the task relates specifically to translation, whereas the task used by the Granada group is in a way a hybrid between translation and interpreting, because the source text is presented one word at a time and because the target text is produced orally.

The hypothesis that translation is parallel is tested using differences in subject-verb (SV) order between English and Danish declarative clauses. English declarative main clauses have what we call SV order, the subject always precedes the verb. In Danish declarative main clauses, by contrast, the order of subject and verb is determined by the principle of verb second or V2, which refers to the fact that the finite verb is always the second constituent of the sentence. V2 has the consequence that if the subject is the first constituent of the main clause, Danish has SV order, which is categorised as the canonical word order. An example could be the sentence Det sker

*i dag* (literally: *It happens today*). Very often, however, an adverbial is the first constituent in the main clause, the verb is still in second position and the subject third, that is, verb-subject (VS) order as in the clause *I dag sker det* (literally: *Today happens it*). In other words, the order of subject and verb in a Danish declarative main clause may be congruent with its English translation if it has SV order, or incongruent if it has VS order. If translation is sequential, this congruence should not affect processing time in translation from Danish to English, while we would expect differences between congruent and incongruent segments if translation is a parallel process.

The subject-verb order variation between Danish and English is well-suited for this investigation because it is purely formal; although VS order contributes to marking non-subject focus in Danish sentences, a similar focus is achieved in English translation equivalents that have SV order purely by fronting of the adverbial. Interestingly, although VS order is normally categorised as non-canonical in Danish, it is quite frequent – large-scale statistics of this for Danish do not exist, but in the other V2 languages, including Swedish, which is closely related to Danish, VS order is found in approximately 40% of declarative main clauses (Håkansson 1997).

### 1.2. Indicators and hypotheses

The main hypothesis of the experiments reported below is that translation is a parallel process. The parallel processing would apply to both congruent and non-congruent segments, but would have the consequence for non-congruent segments that they should be more difficult to process, because they require a reordering of constituents. We test this using professional translators as participants in Experiment 1 (see section 2). Because our participants were professional translators and the difference between English and Danish that we are investigating is a very basic one, we do not expect a large congruence effect on production time – and much less an effect on the actual product, resulting in grammatical errors – but we do expect an effect on ST reading time. Such an effect would indicate that the translators anticipate the TT structure while reading the ST and as such would provide evidence of parallel processing.

We index processing difficulty in three different ways: Firstly, we analyse for how long the translators gaze at the congruent (SV) as opposed to the non-congruent (VS) segments in the ST. Based on the eye-mind assumption (Just and Carpenter 1980), longer gaze times are taken as evidence of more difficult processing. Therefore, we hypothesise that VS segments will elicit longer gaze times than SV segments, when a range of other variables are controlled (see section 2.2 below). The total gaze time measure includes all fixations on the given segment, but excludes saccadic activity because the eye is generally believed to be blind during saccades (Rayner 1998: 378). There are alternative indicators of fixation activity, for example length of single fixations and first pass reading time which may be obtained with the equipment used here (see Balling 2013), but for the reading of clause-level segments, the more aggregated total gaze time is the more meaningful measure. The number of fixations on the segments may also be analysed, but this does not provide any new information compared to the total gaze time.

Secondly, variation in pupil dilation during attention to the relevant ST segments is analysed. Generally, larger pupils are associated with heavier cognitive load (Iqbal, Adamczyk *et al.* 2005) and we hypothesise that participants' pupils will be more

dilated during reading of VS segments in the ST than during reading of SV segments. Pupil dilation is an interesting supplementary measure because, in contrast to gaze time, it is not under conscious control. Pupil dilation was registered for each 20 ms gaze sample on the critical segments, so, in contrast to total gaze time, pupil dilation is not an aggregate measure. This means that the pupil dilation analyses are based on a larger number of datapoints which is an advantage for the statistical techniques. Previous research has shown that pupils constrict and dilate with some delay relative to the presentation of a stimulus (see Hvelplund 2011: 70 for an overview). It was necessary to take this delay into account; therefore the pupil size values that were used to calculate pupil dilation variation were taken from gaze samples which occurred 100 ms after a fixation inside a critical segment, based on the approach by Hvelplund (2011).

Thirdly, the translation product is considered in an analysis of what the translators typed during the task, using Translog (Jakobsen and Schou 1999). In a quantitative analysis, which is our focus here, we may observe longer production time for the non-congruent VS segments than for the congruent SV segments as a consequence of longer pauses between key presses and more mistakes made in the non-congruent segments, but given the experience of the participants, we do not expect a large effect.

In order to test whether the effects observed in Experiment 1 are specific to translation, Experiments 2 and 3 investigate reading of the source texts of Experiment 1 by first (L1) and second (L2) language readers (see sections 3 and 4).

## **2. Experiment 1: Translation from L1 Danish to L2 English**

### **2.1. Method**

#### *2.1.1. Participants*

The participants were eight professional translators whose native language was Danish and whose language of translation was English.<sup>1</sup> They were all certified translators and had at least two years' professional experience, and although they cannot be said to be expert translators according to Shreve's (2002: 161) 10-year professional experience threshold, they can be assumed to share competences which make them well-equipped to produce quality translations. In addition, data from 11 translators (who had parallel profiles in terms of languages and professional experience to those translators whose data were analysed) had to be discarded due to poor eye tracking quality, discussed in section 2.1.4, but the data from the remaining translators were sufficient to obtain reliable results. All eight translators whose eye-tracking record was analysed had English as their L2; further detail is provided in Appendix B. The keylog data from all 19 translators – the eight whose eye-tracking data were analysed and the 11 whose eye-tracking data had to be discarded – were analysed.

#### *2.1.2. Task*

The participants were asked to translate first a warm-up text and then two experimental texts (A and B) from their native language Danish into their second language English (excepting one translator whose second language was German and third language English, but his eye-tracking data were not analysed). The order of presentation of the experimental texts was random. Participants were asked to translate the

texts to their own satisfaction, using no translation aids. Participants could choose whether they wanted to edit their translations. There was no time constraint.

### 2.1.3. *Texts*

The two source texts (see Appendix A) were constructed in order to obtain a high and equal number of appropriate SV and VS segments, namely a total of 12 of each in both texts. The SV segments are canonical syntactic structures such as, from text A, *disse misforståelser giver* (*these misunderstandings give*), where the subject precedes the verb, while VS segments are non-canonical but still very frequent cases where the verb precedes the subject, as exemplified by the segment *tror de fleste mænd* (literally: *think the most men*), also from Text A (all segments and their contexts are shown in Appendix A along with sample translations). During the analysis, it transpired that one of the VS segments had a markedly lower mean content word frequency than the remaining segments; this segment was excluded from the analyses in order not to distort the results, so the analyses reported below are based on 12 SV and 11 VS segments. In our setup, with statistical control of background variables (see further below, section 2.2), it was unnecessary to lose data by also discarding a matching SV segment. This contrasts with a traditional factorial design based on experimental control, where we would have had to also discard an SV segment and thus lose more data.

All segments contained a simple finite verb, because complex verb phrases are discontinuous in non-canonical sentences. All subjects were complex in order to maximise the chance of observing a “transposition” effect when translating a Danish VS segment into an English SV segment. None of the relevant segments occurred at the beginning and end of sentences since reading times may be longer at these positions (Rayner and Sereno 1994). All VS-segments were in main clauses, while a few of the SV-segments were in subordinate clauses, which as a rule have SV order.

Both texts were relatively simple texts on general topics, containing no specialised or very low-frequent vocabulary items. Their lengths were around 150 words. One text had a higher LIX score (a readability index calculated using the LIX calculator of [elkan.dk](http://elkan.dk))<sup>2</sup> than the other (46 for text B, 37 for text A), but the analyses showed no differences between the two texts and they could therefore be analysed together.

### 2.1.4. *Equipment and data quality*

The experiments were run using a Tobii 1750 eye tracker, which is a remote tracker that looks like a standard computer screen. Participants were seated at a distance between 55 and 80 cm from the screen. The position of their eyes was calibrated on a five-point grid before the translation task started. If the experiment was paused for some reason, the participant's eyes were recalibrated.

The STs were presented using Translog 2006 (version 3.2.5.0) and the translated texts were typed in Translog's target text output window. The font was 18 point Times New Roman and double line spacing was used in the STs in order to maximise the chance of accurately linking fixations to appropriate areas of interest (AOIs), which were spatially defined regions surrounding the SV and VS segments. Due to limitations in the Translog version used for the experiment reported here, single line spacing was used in the output window.



Tobii's analysis software ClearView was used to define fixations to include gaze samples that were spatially within an area of 40 pixels of each other on the screen and temporally within a window of 100 ms of each other.

Since average fixations during reading have been found to range between 200 and 250 ms (Rayner and Sereno 1994), data from participants with average fixations below 175 ms were excluded from the analysis, because such short fixations are likely to indicate measurement error. On examination, it turned out that most of the participants for whom we had to discard data had been seated at a distance of more than approximately 60 cm, which may have contributed to the problem. If this is indeed the case, it can be remedied in future studies to reduce the attrition rate. Another potential problem that cannot so easily be remedied is the fact that participants may move around when working at a remote eye tracker, causing the calibration of the eyes to become imprecise; however, in naturalistic translation tasks alleviating this problem by using a chin rest is highly undesirable.

The loss of so much data is of course regrettable, but the gaze data from the remaining eight translators turned out to be sufficient to obtain significant results (see the analyses below). While we would have preferred to have a larger sample of good-quality data, the problem of small datasets is most acute when expected effects turn out to be non-significant, since in that case the non-significance may be due to reduced statistical power. This is not what happens in the analyses reported below. That being said, replications with different and more participants are of course always desirable.

For the analysis of the translation product, data from all 19 translators can be used, since the keylog record does not depend on the quality of the eye-tracking data.

## 2.2. Statistical analysis

The results were analysed using linear mixed-effects regression models in the statistical environment R (version 2.13.1, R development core team 2011),<sup>3</sup> using the packages lme4<sup>4</sup> and languageR.<sup>5</sup> There are several advantages to this analysis approach. On the most general level, the use of inferential rather than purely descriptive statistics (such as percentages and means) allows us to say much more confidently whether any effects observed in our data are likely to apply more generally to the translation process; if effects are significant at the .05-level, we assume that this is the case.

More specifically, a regression approach is useful because it allows statistical control of a range of variables which would otherwise have to be controlled experimentally or constitute noise in the analysis. This means that we can investigate the effect of congruence once we have taken other relevant factors, such as word frequency and segment length, into account. If such variables had to be controlled experimentally, the result would be fewer and more unrepresentative items. For a formal discussion of these issues, see Baayen, Davidson *et al.* (2008), for a more informal one, see Balling (2008). The fact that these variables can be included in the analysis also serves to make the experiments more informative, an advantage in any kind of research but perhaps especially in a relatively young field such as translation process research.

For the present experiments, we analyse three dependent variables – the indicators of translation processing load outlined in section 1.2 – and different categories



of independent variables, or predictors. Both the dependent variable total gaze time and the predictor word frequency had skewed distributions that are typical of reaction times and lexical frequencies; in order to minimise the harmful impact of outliers, gaze time was logarithmically transformed (but shown untransformed in the Figures for ease of interpretation) and word frequency was square root transformed. Both transformations are standard practice in psycholinguistics, experimental psychology and a range of other disciplines where real data are involved (Bartlett 1947; Box and Cox 1964; Howell 2007; Johnson 2008; Myers and Well 2003; see also the guidelines for dealing with skewed data from the Royal Society of Medicine Press).<sup>6</sup> What the transformations do is equate larger distances at the higher end of the scale with smaller distances at the lower end of the scale of a given variable. This is particularly intuitive for a measure like frequency where the difference in speakers' and readers' familiarity is likely to be bigger when frequency jumps from 1 to 10 occurrences per million words than when it jumps from 101 to 110 occurrences. For the reading times, the log transformation means that differences in reading time at the lower end of the scale have more weight than reading times at the higher end of the scale; though primarily motivated by the fact that it is necessary for the statistical procedure, this also makes some sense since very long reading times may reflect processes other than the ones we are most interested in.

Turning to the predictors, our most important predictor is of course congruence, which is a factor with two levels: congruent (SV) vs. incongruent (VS). Additionally, we included two categories of control predictors.

First, we have predictors that are inherent to the words and AOIs, namely the mean frequencies of words and letter bigrams in the AOIs, the length of the AOIs in letters and whether the AOIs contained words that were cognate between ST and TT. The frequencies were extracted from a combination of two corpora of contemporary Danish, Korpus90<sup>7</sup> and Korpus2000;<sup>8</sup> the word frequencies used were the frequencies of the whole word forms, rather than the lemmas. For the mean word frequency, we included only the content words in the AOI, since function words are likely to be skipped by the eye (Rayner 1998) and moreover are generally of such high frequency as to distort the mean frequency considerably. These ST related predictors are likely to be more important in the analyses of gaze time and pupil dilation than in the analysis of typing time, but we test their significance in all analyses.

Secondly, we include two variables which arise during the experiment or as a result of the presentation of the AOIs in context. The first context variable is the position of the segment in the text, where we may observe speeding up (priming) or slowing down (fatigue or relaxation) as the experiment progresses. The second context variable is the effect of repetition of words, which arises because we present sentences in a textual context rather than individually. Repetition of words was measured as the mean number of times the content words in each AOI had occurred in the preceding parts of the text.

The choice of control variables is based on an analysis of the characteristics of the naturalistic experimental situation that we used and on the literature on experimental effects in translation and reading. Thus, we know from a large number of studies that word frequency affects reading time across different languages (for example Taft 1979 for English; Baayen, Dijkstra *et al.* 1997 for Dutch; Moscoso del Prado Martín, Bertram *et al.* (2004) for Finnish; Balling and Baayen 2012 for Danish)

as well as a range of other cognitive processes (see Hasher and Zacks 1984), making it a reasonable variable to control. Similarly with cognateness (see for instance Dijkstra, Miwa *et al.* 2010; Kroll and Tokowicz 2005), where special circumstances may hold for translation as opposed to reading since translators may have a tendency to avoid using cognates where possible (Jakobsen, Jensen *et al.* 2007). The control predictors in the present study do not constitute all possible influences on the translation process, but together with the fact that we generalise across at least eight participants in all analyses, they provide a reasonable level of control of factors that influence the naturalistic task the participants undertook.

A further advantage of using linear mixed-effects models is the fact that these allow us to model the considerable variation between participants and items which is not captured by the predictors included in the analysis. This is done by the inclusion of so-called random effects of participant and item, which are adjustments to the intercept of the regression model for each participant and item. The intercept of the model may be informally described as the point where the regression line – which may describe the relation between a dependent variable such as total gaze time and a predictor such as word frequency – crosses the vertical axis (for illustration, see for instance Figure 2). The model adjusts this intercept up for relatively slow participants (those with longer total gaze times) and down for relatively fast participants and in this way takes the variation between participants into account. Similarly, random intercepts for AOIs account for the random variation between the different AOIs.

Different models were run for the different indicators or dependent variables, that is, total gaze time, pupil dilation and production time. In each case, the dependent variable was analysed as a function of a range of different predictor variables. We started each analysis by including the most control-oriented variables and then gradually adding the more interesting ones, ending with the congruence variable. Variables that turned out to be non-significant were excluded from the final analyses and are therefore not shown in Tables 1 to 7, but they are mentioned where relevant. All analyses and the variables tested in each are summarised in Table 8 and considered in the general discussion. Using this procedure, we were able to investigate the effect of congruence once other relevant variables were taken into account.

In some of the final models, the residuals (the differences between each actual gaze time, pupil dilation or production time and the prediction of the model) showed a marked deviation from normality which could be remedied by removing data points whose residuals were more than 2 standard deviations from the mean. This was done for the final models for Experiment 1 (translation gaze time, pupil dilation and production time), removing between 2.4% and 4.9% of the data points from the analysis. It results in more robust models and effects that are less likely to be driven by outliers. For the remaining analyses, such trimming of the data was unnecessary.

In regression analyses, especially in models with many predictors, it is necessary to consider collinearity between the predictors, which springs from multiple correlations between the predictors in the models. If collinearity between predictors is high, it becomes impossible to disentangle the contributions of each individual predictor. In all our models, all pairwise correlations between predictors were below 0.2 (where the traditional point for starting to do something about collinearity is 0.5) and the condition number of each model was below 18, which is well below the condition number 30 which Baayen (2008) defines as indicating harmful collinearity.

2.3. Results and discussion

The results of the three analyses, for total gaze time, pupil dilation and production time, are shown in Tables 1 to 3. These Tables, and those for the next two experiments, show the name of each predictor in the first column and its effect size (estimate) in the second column. The remaining columns are based on a run of 10,000 Markov chain Monte Carlo (MCMC) simulations for each analysis; these are simulations based on the data and the analysis of the data and provide more accurate p-values and confidence intervals than those based on the t-distribution (see Baayen, Davidson *et al.* 2008), especially for smaller data sets. The third and fourth columns show the credible intervals within which 95% of the observations are expected to lie; they are similar to standard confidence intervals but provide superior accuracy. Finally, the p-value in the fifth column is also based on the MCMC-sampling.

TABLE 1  
**Linear-mixed effects regression model fitted to total gaze time, using treatment coding for the factor Congruence with SV as the reference level.**

The model includes random intercepts for participant (standard deviation, SD, estimated at 0.3981) and item (SD estimated at 0.1147); the residual standard error was estimated at 0.3857.

	Estimate	HPD95 lower	HPD95 upper	p
Intercept	6.5875	5.9192	7.2477	0.0001
AOI length in characters	0.0428	0.0246	0.0600	0.0001
AOI position (linear)	0.0213	0.0131	0.0292	0.0001
AOI position (quadratic)	-0.0001	-0.0002	-0.0001	0.0002
Square root mean frequency (linear)	-0.0165	-0.0235	-0.0100	0.0001
Square root mean frequency (quadratic)	0.0001	0.0000	0.0001	0.0002
Congruence: VS	0.1618	0.0051	0.3091	0.0362

TABLE 2  
**Linear-mixed effects regression model fitted to pupil dilation, using treatment coding for the factor Congruence with SV as the reference level.**

The model includes random intercepts for participant (SD estimated at 0.2319) and item (SD estimated at 0.0707); the residual standard error was estimated at 0.1651.

	Estimate	HPD95 lower	HPD95 upper	p
Intercept	3.7341	3.6153	3.8487	0.0001
Congruence: VS	0.0167	-0.0397	0.0715	0.5682

TABLE 3  
**Linear-mixed effects regression model fitted to typing time, using treatment coding for the factor Congruence with SV as the reference level.**

The model includes random intercepts for participant (SD estimated at 0.1948) and item (SD estimated at 0.1255); the residual standard error was estimated at 0.2464.

	Estimate	HPD95 lower	HPD95 upper	p
Intercept	7.7022	7.3198	8.0669	0.0001
Characters produced (linear)	0.0723	0.0496	0.0972	0.0001
Characters produced (quadratic)	-0.0007	-0.0012	-0.0003	0.0016
Deletions: YES	0.1900	0.1351	0.2516	0.0001
Square root mean frequency	-0.0018	-0.0031	-0.0006	0.0072
Congruence: VS	0.0771	-0.0412	0.1868	0.1752

Most of the analyses include both categorical predictors such as congruence and numerical predictors such as frequency. The regression analysis includes categorical predictors by letting the intercept of the model represent one level of the factor, the so-called reference level, while the other level of the factor is represented in the model as the difference of this other level relative to the reference level. To give an example, in the models summarized in Tables 1 to 3, SV is represented by the intercept while the line ‘Congruence: VS’ shows the estimated difference of VS relative to SV and the significance of that difference. For numerical predictors, the estimate column represents the slope of the regression line (which is also illustrated, for instance, in the top left Figure 2); when the effect of the categorical predictor is non-linear, it is represented by two lines in the Tables, one for the linear and one for the quadratic parameter.

We first discuss the results with respect to congruence and then the different control predictors, which also provide information about the translation process.

2.3.1. *Congruence*

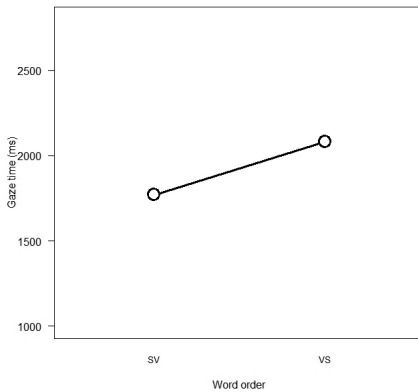
As shown in Figure 1 (see also Table 1), the gaze time analysis showed an effect of congruence: The incongruent VS-segments were read significantly slower than the congruent SV-segments. This is clear evidence of the parallel processing hypothesis: we see the effect of congruence on gaze time, which measures attention to the ST where congruence with the TT is irrelevant under the sequential view of the translation process. Importantly, this holds for experienced professional translators when other relevant factors are controlled. The control of other relevant factors also holds for the illustration of the effect in Figure 1: This is a partial effects plot that shows the effect of congruence once the other predictors in the analysis are taken into account, that is, the effect of congruence all other things being equal.

In contrast, the pupil dilation analysis summarized in Table 2 showed no effect of congruence, and thus no evidence of increased cognitive load for the incongruent segments. However, the pupil dilation analysis is perhaps somewhat questionable, since it shows no significant effects at all, also not for otherwise ubiquitous reading variables such as word frequency.

We hypothesised that congruence would have only a very small effect on production time, or none at all, because the congruence difference we investigated is so very basic and because the participants were skilled L2 users and professional translators.

FIGURE 1

**Partial effects of congruence on total gaze time.**



This prediction was borne out in the analysis of production time summarised in Table 3: There was no significant effect of congruence.

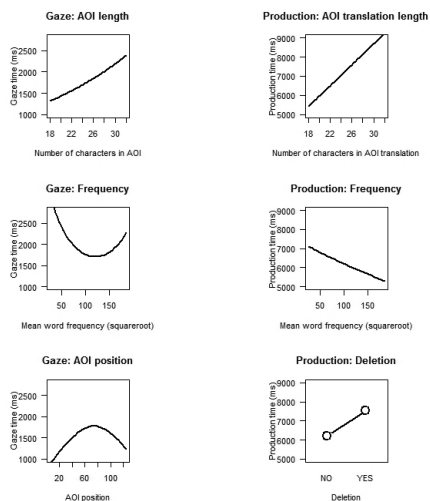
While the difference between gaze time and production time was at least to some extent expected, the difference between gaze time and pupil dilation is more surprising, since both measures are supposed to index processing difficulty. Of the two, gaze duration measures are by far the most well-established (see for example Rayner 1998), while pupil dilation has been used much less as an indicator of cognitive load. There are various possible explanations for the difference: Firstly, it may be that the pupil dilation measure is not precise enough; this is supported by the fact that none of the other predictors had a significant effect on this indicator in the present experiment. Secondly, it may be that the two indicators measure different processes: Perhaps anticipating the reversal of subject and verb necessary for target text production, which we take to be the process underlying the increased gaze time for VS-segments, takes time without being so inherently difficult as to trigger larger pupil dilation. Thirdly, it may be that the 100 ms pupillary delay that we introduced for all participants at all times was either too long or too short to capture the pupil size during the reading of a critical segment. It is also not unlikely that pupillary delay varies among participants. In future studies, one way to overcome this potential problem could be to identify baseline measurements of pupillary delay which could then be used to calculate more confidently the pupil dilation for each participant. The hypothesis that translation is a parallel process is confirmed irrespective of which of these explanations is correct.

### *2.3.2. Effects of control predictors*

As mentioned in section 2.2, statistical control of both item- and experiment-related variables is crucial in a naturalistic experiment like the present one. Additionally, the inclusion of these variables in the analysis makes the experiment yield more information, and the analyses of gaze time and production time showed effects of both item and experiment variables, which are illustrated in Figure 2. The analysis of pupil dilation showed no significant effects at all.

FIGURE 2

**Partial effects of control predictors on gaze time and production time in Experiment 1.**



The top left panel of Figure 2 shows the effect of ST AOI length on gaze time: As expected, translators looked longer at longer segments. The ST AOI length did not have a significant effect on production time, but a related variable did, namely the length of the TT AOI, which is illustrated in the top right panel.

The next item-related variable is the mean frequency of the content words in the segment (square root transformed to reduce skewness). This had significant effects on both gaze time (middle row left panel) and production time (middle row right panel). For gaze time, the effect was non-linear with initial facilitation flattening out for higher values. The apparent inhibition for very high values holds only for relatively few items and may not be reliable.

Finally, the bottom row of Figure 2 shows the effects of two experimental variables: For gaze time, there was a significant effect of the position of the AOI, again a non-linear effect, but predominantly inhibitory, corresponding to a fatigue or relaxation effect as the task progressed. In the production analysis, whether or not participants made deletions in a given AOI had an effect, in the expected direction with shorter production times for AOIs in which no deletions were made. The number of deletions was split into the binary factor – deletion or not – due to the binary distribution of the original variable, with a large number of AOIs with no deletions and the remaining ones distributed across different numbers of deletions between one and 18.

## 2.4. Conclusions and new questions

In sum, our translation experiment shows that segments that are incongruent between the source and target texts are looked at longer than congruent segments, indicating that processing is parallel, since the necessity for transposition of word order in the TT seems to be anticipated during reading of the ST. However, the pupil dilation analysis showed no difference between congruent and incongruent segments, indicating that – although they take longer to process – the incongruent segments may not be inherently more difficult to process, at least not for the experienced

translators taking part in this experiment. Alternatively, the absence of any effects in the pupil dilation analysis may be a sign that this indicator is not sensitive enough for the present purposes, or that the delay between the occurrence of a stimulus and the reaction of the pupils is more variable between participants than we can model. Finally, the analysis of production time showed effects of various control predictors, but no effects of congruence, which is perhaps not surprising given the experience of the translators and the fundamentality and prominence of the difference between English and Danish that is expressed in the congruence variable.

Although the results point in the direction of parallel processing, two questions remain before we can firmly conclude that processing during translation is parallel. The first question is based on the syntactic difference investigated: Since the incongruent segments in Danish are categorized as non-canonical, it may be that the effect observed in the translation experiment is not a translation effect at all, but simply the result of the non-canonical Danish segments, which represent the incongruent category here, being inherently more difficult to read. Given the high frequency of VS-constructions in Danish, we would expect that this is not the case, but that is an empirical question which should be investigated. In Experiment 2 (see section 3), we therefore asked L1 Danish participants to read the Danish texts that were translated in Experiment 1, in order to establish whether the VS segments are more difficult to process than the SV segments in a reading task.

The second question comes from the perspective of research on general L2 processing. There is evidence that both lexical and syntactic structures of a language user's first language affect processing in their second language (see for example De Groot and Nas 1991; Hartsuiker, Pickering *et al.* 2004; and Balling 2013 for an experiment with a similar population to that of Experiment 2) and, at least to some extent, vice versa (Van Hell and Dijkstra 2002; Costa, Caramazza *et al.* 2000). Therefore, the question arises whether the transfer of L1 syntactic structure that we observed in Experiment 1 is a translation phenomenon or the result of more general characteristics of bilingual processing. We address this question in Experiment 3 (see section 4), where the same texts as in the previous experiments were read by a group of participants whose L1 was English and whose highly proficient L2 was Danish.

### 3. Experiment 2: L1 reading

In Experiment 2, the Danish source texts from Experiment 1 were presented to a group of L1 Danish participants, who were asked to read the texts for comprehension. The same eye-tracking indicators were analyzed as for Experiment 1, namely total gaze time and pupil dilation, while there was of course no keylog record to analyze from this reading task.

#### 3.1. Method

Fourteen native speakers of Danish participated in the experiment. All were MA students of English at the Copenhagen Business School with English as a highly proficient second language, but none of them had grown up as bilinguals. Further information on these participants is included in Appendix B.

The participants were asked to read the source texts from Experiment 1 with the object of comprehending them and being able to answer a few simple comprehension



questions after reading (none were actually asked, the warning about questions was merely given in order to ensure that the participants did not read too superficially). No warm-up task was deemed necessary.

The experiment was run using the same equipment as in Experiment 1. Since participants were only reading the two texts, the experiment was quite short, five to ten minutes including calibration, and we observed no problems with data quality according to the criteria outlined in section 2.1.4.

3.2. Results and discussion

The results of the analyses are summarized in Tables 4 and 5. The final models summarized in the Tables were reached in the same way as in Experiment 1: The predictors were added to the model in order of their importance to the hypothesis, starting with the most peripheral control predictors and ending with the central congruence predictor. Non-significant predictors were excluded from the analysis, with the exception of congruence, which is included to allow the reader to verify the size and (non-) significance of the effect.

TABLE 4  
Linear-mixed effects regression model fitted to total gaze time in L1 reading.

The model includes random intercepts for participant (SD estimated at 0.2671) and item (SD estimated at 0.2279); the residual standard error was estimated at 0.5085.

	Estimate	HPD95 lower	HPD95 upper	p
Intercept	6.2965	6.0023	6.5907	0.0001
AOI position	0.0035	0.0000	0.0067	0.0448
Cognateness: NO	0.3011	0.0506	0.5415	0.0182
Congruence: VS	-0.0637	-0.2661	0.1541	0.5348

TABLE 5  
Linear-mixed effects regression model fitted to pupil dilation in L1 reading.

The model includes random intercepts for participant (SD estimated at 0.0452) and item (SD estimated at 0.3083), the residual standard error was estimated at 0.1680)

	Estimate	HPD95 lower	HPD95 upper	p
Intercept	3.7108	3.6203	3.8048	0.0001
AOI position	-0.0026	-0.0032	-0.0020	0.0001
Congruence: VS	0.0373	-0.0022	0.0763	0.0638

Neither analysis shows any significant effect of congruence, though the effect on pupil dilation approaches significance ( $p = 0.0638$ ), with a tendency for readers' pupils to be larger when reading the incongruent VS-segments. Generally, one should be careful about interpreting non-significant results, since the non-significance could be an issue of measurement error or insufficient statistical power. Such care should also be taken here, but given the relatively large effect in Experiment 1 and the very small one here, it appears that the translation effect observed in Experiment 1 is not an artefact of the non-canonical VS-segments being inherently more difficult to process for L1 language users than the canonical SV segments. The question remains

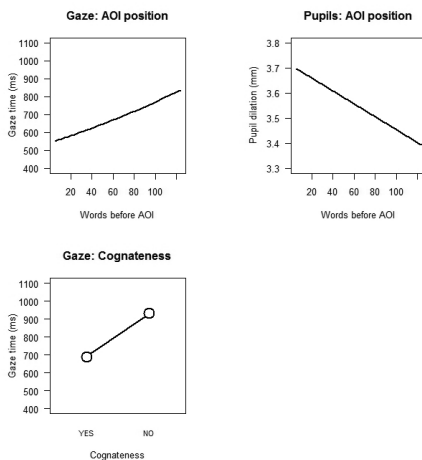
whether the delay for the VS segments also holds more generally in bilingual processing; this question is addressed in Experiment 3, which is reported below.

In the present experiment, only two of the control variables had significant effects on gaze time, namely the position of the AOI, with slightly longer gaze times but slightly reduced pupil dilation early in the experiment, as illustrated in the top row of Figure 3. This may be understood as a relaxation effect. It suggests that reading becomes slower as the experiment progresses, while the cognitive effort as indicated by pupil dilation becomes smaller (this and other differences between effects in the gaze time and pupil dilation analyses are examined in the general discussion, section 5).

More interestingly, the gaze time analysis showed an effect of cognateness with English, the L2 of the participants, as illustrated in the bottom panel of Figure 3. A simple count of the number of cognates in each AOI had a highly skewed distribution; therefore the cognateness variable was reconstructed as a binary indicator of whether the AOI contained cognate words or not. The effect in the gaze time shows that AOIs that did not contain cognates (Cognateness: NO in Table 5) were looked at longer than AOIs that did contain cognates. This is remarkable because the task is reading in the L1 where awareness of L1-L2 correspondences is not in any way relevant or encouraged. It provides further evidence of the effect of cross-linguistic influences on the processing even in the L1 of late bilinguals like the participants in the present experiment (see for example Van Hell and Dijkstra 2002; Costa, Caramazza *et al.* 2000). Although interesting, we interpret this with some caution, since this is the only analysis that shows this effect.

FIGURE 3

Partial effects of AOI position and cognateness on total gaze time and pupil dilation in Experiment 2.



## 4. Experiment 3: L2 reading

### 4.1. Method

The method of this experiment was the same as for Experiment 2, except that the participants were nine native speakers of English (both British and American) who had lived for a long time in Denmark (13–41 years, mean 26) and were highly proficient

L2 speakers of Danish that function in Danish society using the Danish language on a daily basis. Information about these participants and their own assessment of their Danish skills is included in Appendix B.

4.2. Results and discussion

The results of Experiment 3 are shown in Tables 6 and 7. Crucially, the congruence variable had no effect at any stage of the analyses. We did, however, observe effects of several of the control predictors: Firstly, we found a non-linear effect of the mean frequency of the content words, which is illustrated for gaze time in the top left panel of Figure 4 and for pupil dilation in the top right panel. The difference between these effects is discussed in section 5.

Secondly, word repetition, the mean number of times the content words in an AOI had occurred previously in the text, had an effect on total gaze time, which is illustrated in the bottom left panel of Figure 4. The non-linearity of this effect is somewhat puzzling: It seems that optimal processing occurs for AOIs with few or many repeated words, while AOIs with a medium number of repetitions are the most difficult.

Finally, the pupil dilation analysis showed a small effect of AOI position, as in Experiment 2, with smaller pupil dilation later in the experiment, this being the result of a habituation effect.

TABLE 6  
Linear-mixed effects regression model fitted to total gaze time in L2 reading.

The model includes random intercepts for participant (SD estimated at 0.3646); the residual standard error was estimated at 0.6785. In this analysis alone, the effect of the word repetition was significant, but only if an outlier with high word repetition value was excluded. In order to control this source of noise, this outlier was excluded in this analysis. The central congruence predictor remains non-significant in all analyses of this experiment.

	Estimate	HPD95 lower	HPD95 upper	p
Intercept	3.9007	1.7110	6.0797	0.0008
Number of repeated words (linear)	4.5724	1.4246	7.7702	0.0050
Number of repeated words (quadratic)	-1.4352	-2.4778	-0.4206	0.0070
Square root mean frequency (linear)	-0.0141	-0.0247	-0.0030	0.0134
Square root mean frequency (quadratic)	0.0001	0.0000	0.0001	0.0276
Congruence: VS	-0.0389	-0.3282	0.2660	0.7800

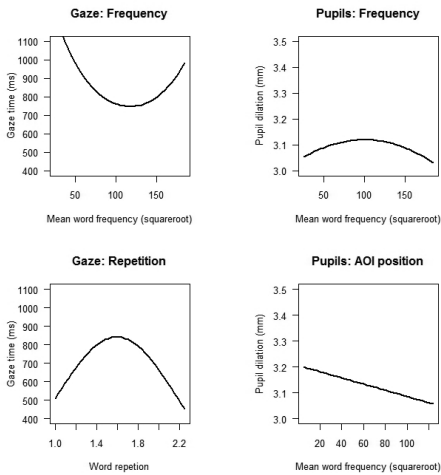
TABLE 7  
Linear-mixed effects regression model fitted to pupil dilation in L2 reading.

The model includes random intercepts for participant (SD estimated at 0.2400) and item (SD estimated at 0.0468); the residual standard error was estimated at 0.1140.

	Estimate	HPD95 lower	HPD95 upper	p
Intercept	3.0851	2.9567	3.2062	0.0001
AOI position	-0.0012	-0.0018	-0.0006	0.0004
Square root mean frequency (linear)	0.0024	0.0007	0.0044	0.0104
Square root mean frequency (quadratic)	0.0000	0.0000	0.0000	0.0104
Congruence: VS	0.0271	-0.0116	0.0697	0.1722

FIGURE 4

**Partial effects of word frequency, word repetition and AOI position on total gaze time and pupil dilation in Experiment 2.**



This experiment with L2 readers of Danish tested whether the congruence effect observed in the translation experiment generalized to L2 processing in general. The translation effect is not directly an L2 effect, since it is an effect on the reading of the L1 before translating into the L2, but the question does arise how far the effect generalizes and whether it is possible to “escape” the syntax of one’s L1 at all – a topic which is also hotly debated in the bilingualism literature (see for example Clahsen and Felser, 2006; Frenck-Mestre, 2005). If the syntax of one’s L1 is inescapable, we would expect an effect of congruence for our group of very experienced L2 users of Danish. This was clearly not the case in this experiment, where the effect of congruence was far from significant, indicating that the syntax of the L1 is not activated to any measurable extent.

## 5. General discussion

In this study, we set out to test the hypothesis that translation is a parallel process in the sense that target text processing is anticipated during source text reading. This hypothesis was confirmed: We observed an effect of congruence between the source and target languages in the translation task in Experiment 1. The fact that the non-congruent VS-segments were looked at longer than the congruent SV-segments during reading of the source text indicates that the syntactic structure of the target language is anticipated during source language processing, showing that the two languages are processed in parallel.

We consider this process parallel in the sense that both source and target language structures are processed while reading the source text. This holds on the relatively general level of reading of source text segments; it cannot be ruled out that very low-level processes are rapidly alternating rather than parallel, but such very low-level processes cannot be accessed using current experimental methods, and indeed at a sufficiently low level, such processes may never be accessible. What may be shown with current methods, and what our experiment clearly shows, is that the translation

process is parallel in the sense that both ST and TT are processed during ST reading, even though attention is only allocated to the ST.

While the translation experiment showed cross-language transfer on the syntactic level, in the form of the word order congruence effect, we saw no solid evidence of such syntactic transfer during reading, in either the L1 or the L2. One possible explanation for this is that it does not matter to reading to what degree a given structure is congruent with its translation equivalent in another language or not; syntactic structures at this level of clausal constituents may simply be too general and abstract to play a role across languages. The more specific lexical level, by contrast, does show some evidence of transfer, in the shape of a cognate effect on L1 reading (Experiment 2, see Table 4 and fFigure 3). Although we interpret this single cognate effect cautiously, it does suggest that the languages of bilinguals are more closely connected at the lexical than at the syntactic level.

It is unclear to us why the cognate effect is significant in the L1 reading experiment but not in L2 and translation. At first sight, cognateness seems likely to matter more in translation than L1 reading, but any positive effect in translation may be cancelled out by a tendency for professional translators to avoid cognates (Jakobsen, Jensen *et al.* 2007). The difference to L2 reading in experiment 3 is more puzzling. It may have to do with the fact that there are more participants in Experiment 2; on the other hand, cognateness is less relevant in L1 reading than it is in L2 reading. A final issue in this relation is that the binary cognateness variable – did the AOI contain cognateness or not – that we had to use because the original cognateness variable was so highly skewed is a relatively crude variable which may not be the most informative. We leave further investigation of this to future research.

Another interpretation of our congruence effect is that it is evidence of literal translation as a default strategy for translation, which may even have the status of translation universal (Tirkkonen-Condit 2005). It seems likely that the time cost for VS-segments is incurred by a rearrangement of constituents that must happen in order to move from a literal and incorrect translation of the VS-segments to a correct and transposed version in the target language, that is, a move from an initial literal to a later non-literal translation. Interestingly, this happens only at the level of ST reading, whereas we see no effects on production – probably because the participating translators were too experienced to actually produce, or even begin to produce, the incorrect literal translations. However, the fact that we observe an effect on ST reading for such a basic phenomenon as the Danish-English word order difference for these experienced translators indicates that this is probably a good candidate for a translation universal, as also argued by Tirkkonen-Condit (2005). This interpretation is supplementary rather than contradictory to our main conclusion that translation is a parallel process.

Alongside the main investigation of congruence, we also investigated a range of other possible predictors of ST reading time and TT typing time (Experiment 1) as well as L1 and L2 reading (Experiment 2 and Experiment 3). Table 8 provides an overview of these different predictors and their effects in the different analyses. This Table shows two main things: that the congruence effect seems to be restricted to the translation task, and that the most consistently significant and therefore most interpretable control predictors are the effects of word frequency and AOI position. The varying shapes of the AOI position effect probably reflects conflicting trends of relaxation, fatigue and habituation, and the differing lengths of the experiment.

TABLE 8

**Overview of the effects of the different explanatory variables on the different dependent variables in the three experiments.**

Light grey cells indicate non-significant (n.s.) effects, darker grey indicates that a given variable is not relevant for the given analysis. Inhibitory means that dependent variable increases with increase in explanatory variable, facilitatory that it decreases. The precise size and shape of each effect is best studied in Figures 1-4.

	Expl: Translation			L1 reading		L2 reading	
VARIABLE	Gaze	Pupils	Keys	Gaze	Pupils	Gaze	Pupils
<b>Congruent (yes/no)</b>	<b>Congruent faster</b>	<b>n.s.</b>	<b>n.s.</b>	<b>n.s.</b>	<b>n.s.</b>	<b>n.s.</b>	<b>n.s.</b>
AOI length in characters	Inhibitory	n.s.	<i>Chars produced</i>	n.s.	n.s.	n.s.	n.s.
Mean frequency of content words	U-shaped	n.s.	Facilitatory	n.s.	n.s.	U-shaped	Highest in middle
Mean letter bigram frequency	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Contains cognate words (yes/no)	n.s.	n.s.	n.s.	Cognates = faster	n.s.	n.s.	n.s.
AOI position in text	Highest in middle	n.s.	n.s.	Inhibitory	Facilitatory	n.s.	Facilitatory
Mean number of word repetitions	n.s.	n.s.	n.s.	n.s.	n.s.	Highest in middle	n.s.
Typing only: characters produced			Inhibitory				
Typing only: deletions (yes/no)			Deletions= slower				

Our study also raises two methodological points: one that opens questions about the validity and interpretation of different eye-tracking indicators and one that suggests new paths for the study of translation processes.

The question about eye-tracking indicators arises because we observe several differences between our gaze time and pupil dilation measures: Firstly, in the translation experiment, there was an effect of congruence on gaze time but not on pupil dilation, while in the L1 reading experiment (Experiment 2), we observed a pattern that approached the reverse: No effect on gaze time, while the effect on the pupil dilation approached significance. Secondly, in Experiment 2, the effect of AOI position had opposite signs for the two indicators, inhibitory for gaze time and facilitatory for pupil dilation. Thirdly, the two indicators show opposite profiles for the frequency effect in Experiment 3.

There are various possible interpretations of this: Firstly and most trivially, it may be that one of the measures is less sensitive or less accurate than the other. Since gaze time is the most well-established of these (see Rayner 1998), we lean towards trusting this indicator the most, but pupil dilation does have the advantage of not being under the conscious control of the participant. A second, more interesting

possibility is that the two measures index different aspects of cognitive processing. In the case of translation, it may be that one indicator reflects changes in workload that are related to comprehending the source text whereas the other indicator reflects changes which are related to reformulation in the target language. When considering the subprocesses involved in comprehension and reformulation, it may further be that the two indicators reflect different cognitive processing aspects such as orthographic analysis, lexical retrieval and propositional analysis. This is a highly speculative explanation which will have to be examined in more detail in future studies.

The second methodological point is that the present experiments demonstrate the advantages of using eye-tracking to execute and mixed-effects models to analyse naturalistic experiments. This setup allows us to investigate translation processes in a laboratory setting but relatively close to how they actually happen in real life. Building on this, our inferential statistical approach indicates to what extent generalization to the wider population of translators is warranted, while the specific mixed-effects regression approach allows the investigation of a single central variable while statistically controlling and also investigating other relevant variables, in our case, exploring the effect of congruence, other things being equal. With this battery of advanced methods, we can firmly conclude that translation processes are parallel in nature, while reading in one of a bilingual's languages seems to draw more on the lexicon than on the syntactic parser of the other language.

#### ACKNOWLEDGMENTS

We are grateful to Inger M. Mees, Fabio Alves and Susanne Göpferich for comments on an earlier article reporting a subset of the data in Experiment 1 (Jensen, Sjørup *et al.* 2009), and to Henrik Selsøe Sørensen for translating the abstract and keywords into French.

#### NOTES

- \* Previously affiliated with the Copenhagen Business School.
- 1. The results from six of these translators were also reported in Jensen, Sjørup *et al.* (2009); the analyses and results are similar to those reported previously, but have been substantially extended and reinterpreted in terms of the distinction between sequential and parallel processing.
- 2. elkan.dk (Last update: 28 November 2011) Visited on 8 February 2012. <[http://elkan.dk/lixtal\\_lixberegner.asp](http://elkan.dk/lixtal_lixberegner.asp)>.
- 3. R DEVELOPMENT CORE TEAM (2011). R: A Language and Environment for Statistical Computing. Version 2.13.1. Downloaded 6 September 2011, <[www.r-project.org](http://www.r-project.org)>.
- 4. BATES, Douglas, MAECHLER, Martin and BOLKER, Ben (2011): lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-42. Downloaded 6 September 2011, <<http://CRAN.R-project.org/package=lme4>>.
- 5. BAAYEN, R. Harald (2011): languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics." R package version 1.2. Downloaded 6 September 2011, <<http://CRAN.R-project.org/package=languageR>>.
- 6. ROYAL SOCIETY OF MEDICINE PRESS: <<http://www.rsmjournals.com/site/misc/stats.xhtml>>.
- 7. Korpus90. Downloaded October 2006, <<http://korpus.dsl.dk/e-resurser/k90info.php?lang=dk>>.
- 8. Korpus2000. Downloaded October 2006, <<http://korpus.dsl.dk/e-resurser/k2000info.php?lang=dk>>.

#### REFERENCES

- ANGELONE, Erik (2010): Uncertainty, uncertainty management, and metacognitive problem solving in the translation task. In: Gregory M. SHREVE and Erik ANGELONE, eds. *Translation and Cognition*. Amsterdam/Philadelphia: John Benjamins, 17-40.



- BAAYEN, R. Harald (2008): *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- BAAYEN, R. Harald, DAVIDSON, Douglas and BATES, Douglas (2008): Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language*. 59:390-412.
- BAAYEN, R. Harald, DIJKSTRA, Ton and SCHREUDER, Rob (1997): Singulars and Plurals in Dutch: Evidence for a Parallel Dual-Route Model. *Journal of Memory and Language*. 37:94-117.
- BALLING, Laura Winther (2008): A brief introduction to regression designs and mixed-effects modelling by a recent convert. *Copenhagen Studies in Language*. 36:175-192.
- BALLING, Laura Winther (2013): Reading authentic texts: What counts as cognate? *Bilingualism: Language and Cognition*. 16(3):637-653.
- BALLING, Laura Winther and BAAYEN, R. Harald (2012): Probability and surprisal in auditory recognition of morphologically complex words. *Cognition*. 125:80-106.
- BARTLETT, Maurice Stevenson. (1947): The Use of Transformations. *Biometrics*. 3:39-52.
- BOX, George Edward Pelham and COX, David Roxbee (1964): An analysis of transformations. *Journal of the Royal Statistical Society, Series B*. 26:211-252.
- CLAHSEN, Harald and Felser, Claudia (2006): How native-like is non-native language processing? *Trends in Cognitive Sciences*. 10:564-570.
- COSTA, Albert, CARAMAZZA, Alfonso and SEBASTIÁN-GALLÉS, Nuria (2000): The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 26:1283-1296.
- DE GROOT, Annette M. B. and NAS, Gerard L. J. (1991): Lexical Representation of Cognates and Noncognates in Compound Bilinguals. *Journal of Memory and Language*. 30:90-123.
- DIJKSTRA, Ton, MIWA, Koji, BRUMMELHUIS, Bianca *et al.* (2010): How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*. 62:284-301.
- DRAGSTED, Barbara (2010): Coordination of reading and writing processes in translation. An eye on uncharted territory. In: Gregory M. SHREVE and Erik ANGELONE, eds. *Translation and Cognition*. Amsterdam/Philadelphia: John Benjamins, 41-62.
- FRENCK-MESTRE, Cheryl (2005): Eye-movement recording as a tool for studying syntactic processing in a second language: a review of methodologies and experimental findings. *Second Language Research*. 21:175-198.
- GERVER, David (1976): Empirical Studies of Simultaneous Interpreting: A Review and a Model. In: Richard W. BRISLIN, ed. *Translation: Applications and Research*. New York: Gardner Press, 165-207.
- GILE, Daniel (1995): *Basic Concepts and Models for Interpreter and Translator Training*. Amsterdam/Philadelphia: John Benjamins.
- HARTSUIKER, Robert J., PICKERING, Martin J. and VELTKAMP, Eline (2004): Is Syntax Separate or Shared Between Languages. *Psychological Science*. 15:409-414.
- HASHER, Lynn and ZACKS, Rose T. (1984): Automatic Processing of Fundamental Information. The Case of Frequency of Occurrence. *American Psychologist*. 39:1372-1388.
- HOWELL, David C. (2007): *Statistical methods for psychology* (6th edition). Belmont, CA: Thomson Wadsworth.
- HVELPLUND, Kristian Tangsgaard (2011): Allocation of cognitive resources in translation. Doctoral thesis. Copenhagen: Copenhagen Business School.
- HÅKANSSON, Gisela (1997): Barnets väg till svensk syntax. In: Ragnhild SÖDERBERGH, ed. *Från joller till läsning och skrivning*. Malmö: Gleerups, 47-60.
- IQBAL, Shamsi T., ADAMCZYK, Piotr D., XIANJUN, Sam Zheng *et al.* (2005): Towards an index of opportunity: understanding changes in mental workload during task execution. *CHI '05: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Portland, Oregon, 2-7 April 2005) New York: ACM, 311-320.
- JAKOBSEN, Arnt Lykke and SCHOU, Lasse (1999): Translog documentation. *Copenhagen Studies in Language*. 24:151-186.

- JAKOBSEN, Arnt Lykke and JENSEN, Kristian Tangsgaard Hvelplund (2008): Eye movement behaviour across four different types of reading tasks. *Copenhagen Studies in Language*. 36:103-124.
- JAKOBSEN, Arnt Lykke, JENSEN, Kristian Tangsgaard Hvelplund and MEES, Inger M. (2007): Comparing modalities: Idioms as a case in point. *Copenhagen Studies in Language*. 35:217-249.
- JENSEN, Kristian Tangsgaard Hvelplund, SJØRUP, Annette C. and BALLING, Laura Winther (2009): Effects of L1 syntax on L2 translation. *Copenhagen Studies in Language*. 38:319-336.
- JOHNSON, Keith (2008): *Quantitative Methods in Linguistics*. Oxford: Blackwell.
- JUST, Marcel Adam and CARPENTER, Patricia A. (1980): A theory of reading: from eye fixations to comprehension. *Psychological Review*. 87:329-354.
- JÄÄSKELÄINEN, Riitta and TIRKKONEN-CONDIT, Sonja (1991): Automated processes in professional vs. non-professional translation: a think-aloud protocol study. In: Sonja TIRKKONEN-CONDIT ed. *Empirical Research in Translation and Intercultural Studies*. Tübingen: Gunter Narr, 89-109.
- KELLOGG, Ronald T. (1996): A model of working memory in writing. In: C. Michael LEVY and Sarah Ellen RANDELL, eds. *The Science of Writing*. Mahwah, NJ: Lawrence Erlbaum.
- KINTSCH, Walter (1998): *Comprehension: A Paradigm for Cognition*. Cambridge: Cambridge University Press.
- KROLL, Judith F. and TOKOWICZ, Natasha (2005): Models of bilingual representation and processing. Looking back and to the future. In: Judith F. KROLL and Annette M. B. DE GROOT, eds. *Handbook of bilingualism: Psycholinguistic approaches*. Oxford: Oxford University Press, 531-554.
- MACIZO, Pedro and BAJO, Maria Teresa (2006): Reading for understanding and reading for translation: Do they involve the same processes? *Cognition*. 99:1-34.
- MOSCOSO DEL PRADO MARTIN, Fermín, BERTRAM, Raymond, HÄIKIÖ, Tuomo *et al.* (2004): Morphological Family Size in a Morphologically Rich Language: The Case of Finnish Compared with Dutch and Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 30:1271-1278.
- MYERS, Jerome L. and WELL, Arnold D. (2003): *Research Design and Statistical Analysis*. 2<sup>nd</sup> ed. Mahwah, NJ: Lawrence Erlbaum.
- RAYNER, Keith (1998): Eye movements in reading and information processing. *Psychological Bulletin*. 124:372-422.
- RAYNER Keith and SERENO, Sara C. (1994): Eye movements in reading: Psycholinguistic studies. In: Morton ANN GERBACHER, ed. *Handbook of Psycholinguistics*. New York: Academic Press, 57-82.
- RUIZ, Carmen, PAREDES, Natalia, MACIZO, Pedro. *et al.* (2008): Activation of lexical and syntactic target language properties in translation. *Acta Psychologica*. 128:490-500.
- SELESKOVITCH, Danica (1976): Interpretation, A Psychological Approach to Translation. In: Richard W. BRISLIN, ed. *Translation: Applications and Research*. New York: Gardner Press, 92-116.
- SHREVE, Gregory M. (2002): Knowing translation: cognitive and experiential aspects of translation expertise from the perspective of expertise studies. In: Alessandra RICCARDI, ed. *Translation Studies: Perspectives on an Emerging Discipline*. Cambridge: Cambridge University Press, 150-171.
- TAFT, Marcus (1979): Recognition of affixed words and the word frequency effect. *Memory & Cognition*. 7:263-272.
- TIRKKONEN-CONDIT, Sonja (2005): The Monitor Model Revisited: Evidence from Process Research. *Meta*. 50:405-414.
- VAN HELL, Janet G. and DIJKSTRA, Ton (2002): Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin and Review*. 9:780-789.

## APPENDIX A

In the Danish texts, SV segments are in bold, VS segments are underlined.

## Text A

Nogle vil påstå, at **mænd og kvinder taler** forskellige sprog. Hvis **en mand i et parforhold siger**, at han ikke ønsker at tale om sine problemer, føler den typiske kvinde, at hun bliver udelukket fra hans verden. Hvis kvinden siger til ham, at hun ikke ønsker at snakke om problemerne, tror de fleste mænd, at hun ikke har behov for at tale om dem. **Disse misforståelser giver** anledning til skænderier. Når **store skænderier opstår**, reagerer kvinder og mænd også forskelligt. **Ny forskning viser**, at **mange kvinder græder**, når de er vrede. I disse situationer vælger skræmte mænd at flygte, og de kan føle sig hjælpeløse. I modsætning til den grædende kvinde bruger den vrede mand hårde ord. På trods af disse forskellige reaktioner formår de fleste mænd og kvinder at leve fredeligt sammen, og vi kan jo nok ikke leve uden hinanden.

## Sample translation of Text A

Some people claim that men and women speak different languages. If the male party in a relationship says that he does not want to discuss his problems, the typical woman will feel excluded from his world. If the woman tells him that she does not want to discuss the problems, most men will think that she has no need to talk about them. These misunderstandings cause arguments. When serious arguments arise, men and women also react differently. Recent research shows that many women cry when they are angry. In such situations, terrified men choose to escape, and they may feel helpless. Contrary to the crying woman, the angry man uses harsh words. In spite of these different reactions, most men and women manage to live peacefully together, and we can hardly live without each other.

## Text B

I mange hjem finder man et kæledyr, og **de fleste mennesker foretrækker** enten hund eller kat. Indtil videre slår den bløde hundehvalp den kære killing i kapløbet om at være Danmarks mest populære kæledyr. Mens **hundens loyalitet forklarer** dens store popularitet, elsker de talrige katteejere kattens selvstændighed. Mens **den almindelige katteejer synes**, at en hund er for uselvstændig og beskidt, opfatter mange hundeejere katte som dovne og upersonlige. **Mange kæledyrsejere vælger** et kæledyr, der afspejler deres egen personlighed. Når **den store stærke mand køber** en stor stærk hund, forstærker hundens træk hundeejerens ego. Omvendt foretrækker den feminine kvinde tit den smukke og elegante kat. **Hunden og katten enes** normalt uden problemer, hvis de er vant til hinanden. Ofte toppes hundeejeren og katteejeren mere end hunden og katten.

## Sample translation of Text B

Many homes have a pet, and most people prefer either dogs or cats. So far, the soft puppy beats the cute kitten in the race to being Denmark's most popular pet. While the dog's loyalty explains its huge popularity, the many cat owners love the cat's independent nature. Whereas cat owners in general find dogs dependent and dirty, many dog owners regard cats as lazy and impersonal. Many pet owners choose a pet that mirrors their own personality. When the big, strong man buys a big, strong dog, the dog's traits emphasise the dog owner's ego. On the other hand, the feminine woman often prefers the beautiful and elegant cat. The dog and the cat normally get along without problems if they are used to each other's company. Often the dog owner and the cat owner are more at odds than the dog and the cat.

APPENDIX B

TABLE 1  
Metadata about the participants in Experiment 1 (translation) - The data from participants 01 to-06 are also reported in Jensen, Sjørup *et al.* (2009).

	Gender	L1	L2	Translator training	Degree finished	Professional experience (Years)
P01	M	Danish	English	Yes (no certification)	2007	3
P02	F	Danish	English	Yes	2005	10
P03	F	Danish	English	Yes	2008	2
P04	F	Danish	English	Yes	1978	15
P05	F	Danish	English	Yes	1979	30
P06	F	Danish	English	Yes	1994	15
P07	F	Danish	English	Yes	2000	6
P08	F	Danish	English	Yes	2005	4

Experiments conducted in 2009-2010

TABLE 2  
Metadata about the participants in Experiment 2 (L1 reading)

Participant number	Gender	L1	Own rating of English proficiency*
Dan01	F	Danish	5
Dan02	F	Danish	5
Dan03	F	Danish	5
Dan04	F	Danish	5
Dan05	F	Danish	5
Dan06	F	Danish	5
Dan07	F	Danish	5
Dan08	F	Danish	5
Dan09	F	Danish	5
Dan10	F	Danish	4
Dan11	F	Danish	4
Dan12	F	Danish	5
Dan13	F	Danish	4
Dan14	M	Danish	5

\*Scale from 1 to 5 with 5 the highest rating

TABLE 3  
Metadata about the participants in Experiment 3 (L2 reading)

Participant	Gender	L1	Years in Denmark	Danish competence*			
				Listening	Reading	Writing	Speaking
Eng01	F	English	39	5	5	5	5
Eng02	M	English	14	5	4	3	4
Eng03	M	English	33	5	5	4	5
Eng04	F	English	31	5	4	1	3
Eng05	F	English	41	4	3	1	4
Eng06	M	English	20	4	5	4	4
Eng07	F	English	13	4	4	2	4
Eng08	M	English	23	5	5	5	5
Eng09	M	English	36	5	5	4	5

\*Scales from 1 to 5 with 5 the highest rating