Word Translation Entropy

Evidence of Early Target Language Activation During Reading for Translation

Schaeffer, Moritz; Dragsted, Barbara; Hvelplund, Kristian Tangsgaard; Balling, Laura Winther; Carl, Michael

Document Version
Final published version

Publication date:
2015

Citation for published version (APA):
Schaeffer, M., Dragsted, B., Hvelplund, K. T., Balling, L. W., & Carl, M. (2015). *Word Translation Entropy: Evidence of Early Target Language Activation During Reading for Translation*. Paper presented at Technical Committee on Thought and Language (TL). 2015, Tokyo, Japan.

Word translation entropy: evidence of early target language activation during reading for translation

Moritz Schaeffer[1], Barbara Dragsted[1], Kristian Tangsgaard Hvelplund[2], Laura Winther Balling[1], Michael Carl[1]

[1] Copenhagen Business School, Department of International Business Communication

[2] University of Copenhagen, Department of English, Germanic and Romance Studies

Abstract

This study reports on an investigation into the relationship between the number of translation alternatives for a single word and eye movements on the source text. In addition, the effect of word order differences between source and target text on eye movements on the source text is studied. In particular, the current study investigates the effect of these variables on early and late eye movement measures. Early eye movement measures are indicative of processes that are more automatic while late measures are more indicative of conscious processing. Most studies that found evidence of target language activation during source text reading in translation, i.e. co-activation of the two linguistic systems, employed late eye movement measures or reaction times. The current study therefore aims to investigate if and to what extent earlier eye movement measures in reading for translation show evidence of co-activation. Results show that the number of translation alternatives for a single word and differences between source and target text in terms of word order have an effect on very early and late eye movement measures. Results are interpreted in terms of semantic and structural cross-linguistic priming: items which have a similar word order in source and target texts are likely to have similar syntactic structures. These items are therefore more likely to prime structurally. Source items which have few translation alternatives are more likely to share a semantic representation and are hence more likely to prime semantically than items with more translation alternatives. Findings support the literal translation hypothesis.

Keywords

Co-activation, priming, translation, entropy, eye movements

1. Introduction

It has been a subject of debate in translation process research (TPR) whether translation is a sequential process or whether and to what extent comprehension and production activities may occur in parallel (Carl and Dragsted 2012, Balling et al 2014). In the sequential, or vertical perspective, human translation is described (Gile 1995) as a process in which the translator first reads a source-language (SL) segment, then formulates a "meaning hypothesis", i.e., assigns a meaning to the translation segment by drawing on SL and general world knowledge, and possibly external information sources, and then checks the meaning hypothesis for plausibility. Having finished the processes involved in understanding the source text (ST), the translator moves on to reformulating the meaning hypothesis in the target language (TL), drawing again on general world knowledge and on knowledge of the TL, and checks for fidelity and general acceptability, continuously revising the target text (TT) until a satisfactory version has been arrived at. In the same vein, according to the Interpretive Model (Lederer 1994) translation is a process in which the translator understands the text, deverbalizes its language and re-expresses its sense in the TL.

Common to these models is that they view ST reading as a phase distinct from the reformulation phase and characterised largely by the same processes as reading for monolingual comprehension. In contrast to this, the horizontal/parallel view holds that TL reformulation commences during ST comprehension, and that the process involved in reading for translation is different from reading for monolingual comprehension (see e.g. Jakobsen and Jensen 2008, Schaeffer et al *forthcoming*). In line with this view, Carl and Dragsted (2012) propose that the ST is understood and meaning hypotheses are generated only to the extent required to keep on producing target text. Deep ST understanding is prompted by problems occurring in the TT. If TT production is interrupted, for instance because the translator is not able to retrieve an appropriate TL equivalent or is considering which translation to choose out of several alternatives (see below), the missing information needs to be retrieved. This may lead to increased eye movement activity and gaze time on a ST word or passage with a view to verification or reinterpretation (ibid.: 143-144).

Schaeffer and Carl (2013: 185) propose a different kind of model in which "...both horizontal and vertical processes are always active at the same time." Schaeffer and Carl (ibid) argue that "...that the horizontal process is an early process while the vertical processes depend on context which becomes available later, as processing advances in the chunk or text..."

This study assumes that translators read the ST with TT production in mind; hence, different processes are involved in reading for translation than in reading for monolingual comprehension.

Previous studies which found evidence of co-activation of the two linguistic systems during ST reading, i.e., studies which found support for the hypothesis that translation is a parallel/horizontal process, employed late eye movement or other late behavioural measures. This study tests the hypothesis that target-language-specific aspects have an impact during very early stages of ST processing. If target language specific aspects have an impact on early eye movement measures, this would allow for a much stronger claim regarding the horizontal/parallel view, because early eye movement measures are more indicative of automatic processes than late measures, and any effect is more likely to allow for conclusions regarding bilingual lexicon.

2. First fixation durations

First fixation durations (*FFDur*) represent the first contact with a $word_n$, before the eyes either re-fixate $word_n$ or move on to word to the left or right. All the low-level aspects of word recognition such as integration of visual features of letters occur during a first fixation duration in addition to processing of morphological and phonological aspects all of which result in lexical access. In addition to the processing of $word_n$, $word_{n+1}$ is pre-processed in terms of visual features such as word length.

The effect of word frequency on *FFDur* was significant and in the expected direction. The effect of word length (in character) was significant, but in the opposite direction of what would intuitively be expected. However, *FFDur* is the duration of a single fixation, the first on the word, which does not automatically become longer for longer words (e.g. Hyönä and Olson 1995). Longer reading times for longer words are due to re-fixations.

We observe significantly shorter first fixation durations towards the end of the text, suggesting that translators become faster as they progress in the translation – even for such an early measure.

Target-language-specific aspects play a role at the earliest stages of reading the source text, indicating that SL and TL are co-activated from the very first visual encounter with an ST word. In particular two aspects of target language properties were investigated:

1. Word translation entropy (*HTra*) indicates how many translation choices a translator has for a given source text word, i.e. how many equally likely translations may be produced for a source word in a given context.

HTra values of source words correlate with their *FFDur*, i.e. words with fewer alternative translations possibilities have shorter first fixation durations than source text words for which there exists a large number of possible alternative translations.

2. The cross-lingual distortion (*Cross*) of ST-TT word alignments represents the relative distance between the position of a word in a source sentence and the position of its translation in the target sentence.

   Source text words which require more (relative) syntactic reordering in the target language correlate with higher *FFDur* values than words that can be translated in a similar order to that in the source language.

In (Schaeffer et al, 2015) we argue that the *HTra* and the *Cross* features indicate the degree of literality of a translation. These findings indicate, thus, that ST words with lower *HTra* and *Cross* values are more likely to prime and facilitate their TT equivalents than words with a higher word translation entropy and which require long distance reordering in the target language. The *Cross* effect was relatively large, suggesting that re-ordering and structural priming play a large role during the early stages of reading during translation. This further lends support to the literal translation hypothesis, in that the default rendering procedure during ST reading in translation is to generate an interim representation in which ST word order and TT word order are identical, where ST and TT items correspond one-to-one and in which each ST word has only one possible translated form. When this is not possible, because of context, target norms or for any other reason, cognitive effort increases.

Mean first fixation durations, gaze durations and regression path durations are relatively short compared to monolingual reading. Effects of word length and frequency are rather small in comparison to monolingual reading: the effect of e.g. frequency on *FFDur* is typically in the region of 20-30ms and on gaze duration normally around 50-60ms while here, it is around 6ms and 20ms, respectively. However, it is unlikely that this is task related. It is more likely that this is because of the way fixations are calculated in the different studies.

3 First Pass Gaze Duration (FPDurS)

First pass gaze duration (*FPDurS*) is the sum of all fixation durations on a word before the eyes move to a different word and these, hence, represent a later stage in lexical processing. A reader

might re-fixate a word either because it is long or because it is difficult to understand or integrate, or because it is ambiguous in some way.

As expected, word frequency is negatively correlated with first pass gaze duration; as mentioned above, word length also had a significant positive effect, also as expected. Somewhat surprisingly, word translation entropy did not have a significant effect on *FPDurS*. However, *Cross* had a positive slope and was significant. While the effect of *Cross* lingers on into gaze durations, the effect of word translation entropy appears very early on, and it only surfaces again in total reading time. This suggests that initial activation of shared representations is relatively automatic and that these automatically activated shared representation serve as a reference in the production of the target text.

4. Total reading time (TrtS)

TrtS is a very late measure which includes all fixations on a $word_n$ – irrespective of when these have taken place.

For total reading time, all effects were highly significant and mirrored those on first fixation durations (apart from word length). Both the effect of *Cross* and the effect of *HTra* on total reading times were relatively strong, positive and highly significant. Again, these findings suggest that the initially and automatically activated shared structural and semantic representations serve as a basis for later regeneration of the ST in the TL and for later monitoring processes.

5. Conclusion

The picture that emerges from our findings is that reading for translation is fundamentally different from reading for monolingual comprehension. Monolingual reading in L1 is the most well-researched type of reading, but no target-language-specific aspects play a role in this kind of reading. This is the first study, to the authors' knowledge, which employs earlier eye movement measures and such a broad range of target languages and such a large corpus of eye movements. Early eye movement measures are crucial if the time-course of the cognitive model is to be investigated and they are also important if conclusions regarding the organisation of the mental representations are to be drawn from the findings: late eye movement measures are likely to be indicative of willed behaviour while early eye movement measures are likely to be indicative of

more automatic processing. It is not very surprising that target-language-specific aspects play a role during the later processes in reading for translation where TT production is involved, unlike in monolingual reading, which does not involve text production.

References

Balling, L.W., Hvelplund, K.T. and Sjørup, A.C. 2014. Evidence of Parallel Processing During Translation. Meta, Vol. 59, No. 2, 2014, p. 234-259.

Carl, M., & Dragsted, B. (2012). Inside the monitor model: processes of default and challenged translation production. In Translation: Corpora, Computation, Cognition. Special Issue on the Crossroads between Contrastive Linguistics, Translation Studies and Machine Translation 2(1), 127-145.

Carl, M., & Schaeffer, M. (forthcoming). Literal Translation and Processes of Post-editing.In: Edited volume: Translation in transition: Between cognition, computing and technology. Amsterdam: Benjamins

Gile, D. 1995. Basic Concepts and Models for Interpreter and Translator Training. Amster-dam/Philadelphia: John Benjamins

Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: effects of word length and word frequency. Journal of Experimental Psychology. Learning, Memory, and Cognition, 21(6), 1430–1440. doi:10.1037/0278-7393.21.6.1430

Jakobsen, Arnt Lykke, and Kristian T.H. Jensen. 2008. "Eye Movement Behaviour Across Four Different Types of Reading Task." In Looking at Eyes. Eye-Tracking Studies of Reading and Translation Processing, edited by S. Göpferich, A. L. Jakobsen, and I. M. Mees, 36:103–124. Copenhagen: Samfundslitteratur.

Lederer, M. 1994. La traduction aujourd'hui. Le modèle interprétatif. Paris: Hachette.

Schaeffer, Moritz Jonas and Michael Carl. (2013). Shared Representations and the Translation Process: A Recursive Model. Translation and Interpreting Studies, 8(2), 169 – 190.

Schaeffer, Moritz Jonas, Kevin Paterson, Victoria A. McGowan, Sarah J. White, Kirsten Malmkjær (forthcoming) Reading for translation

Schaeffer, Moritz Jonae, Barbara Dragsted, Kristian Tangsgaard Hvelplund, Laura Winther Balling, Michael Carl. "Word translation entropy: evidence of early target language activation during reading for translation". In "New Directions in Empirical Translation Process Research: Exploring the CRITT TPR-DB". Carl, Bangalore, Schaeffer (eds), Springer, 2015