

Article

Not peer-reviewed version

Behavioral vs. Verbal Methods in Translation Quality Evaluation: A Cognitive Experimental Study

[Xin Huang](#) and [Xiang Zhang](#)*

Posted Date: 16 March 2026

doi: 10.20944/preprints202603.1087.v1

Keywords: translation quality evaluation; behavioral metrics; verbal reports; dual-system theory; cognitive effort



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Behavioral vs. Verbal Methods in Translation Quality Evaluation: A Cognitive Experimental Study

Xin Huang¹ and Xiang Zhang^{2,*}

¹ Jinan University, Guangzhou, China

² Macao Polytechnic University, Macao, China

* Correspondence: zhangxiang@mpu.edu.mo

Abstract

This study explores the sensitivity differences between behavioral experiments and verbal reports in translation quality evaluation. Results indicate that behavioral metrics (e.g., response times) are significantly more sensitive to syntactic-pragmatic manipulations (phrase order) than verbal reports. Translations with congruent phrase order received higher ratings and faster response times compared to those with incongruent order. However, most participants explicitly denied phrase order's influence in verbal reports. Lexical equivalence showed no significant impact on explicit ratings but increased cognitive effort, as indicated by slower response times for approximate lexical matches. These findings reveal a critical dissociation between implicit cognitive processes and explicit awareness in translation evaluation. The study highlights that translation assessment involves both implicit System 1 processes and explicit System 2 reasoning, offering new cognitive insights for translation research and practical implications for translator education and machine translation assessment. By bridging cognitive science and translation studies, this research contributes to a paradigm shift: translation quality is not merely what evaluators say it is, but what their cognitive behavior reveals it to be.

Keywords: translation quality evaluation; behavioral metrics; verbal reports; dual-system theory; cognitive effort

1. Introduction

Translation quality evaluation, a cornerstone of cross-cultural communication, has long been dominated by subjective judgments rooted in philosophical debates and linguistic preferences (Afrouz & Parsa, 2023). As the final stage of the translation process, evaluation plays a crucial role in determining the success and communicative effectiveness of the translated text. Numerous scholars have proposed various frameworks, approaches, or theories from philosophical, linguistic, and cultural standpoints—such as those developed by Williams (2004), House (1997, 2001, 2004, 2015), Reiss and Vermeer (1984) and—in an effort to bring structure and rigor to the evaluation process. Traditional approaches, such as equivalence-based criteria (e.g., “faithfulness” or “fluency”), rely heavily on experts’ introspective reports, which are vulnerable to personal biases and lack empirical validation (Newmark, 1988; Nida, 1964). Despite advancements in cognitive science since the 1990s—particularly in bilingual processing (Green, 1998; Kroll & Stewart, 1994)—translation studies remain largely disconnected from experimental methodologies capable of probing implicit cognitive mechanisms.

The emergence of cognitive approaches, such as Think-Aloud Protocols (TAPs) and eye-tracking, has partially bridged this gap by shifting focus from texts to translators’ mental processes (Alves et al., 2010; Bell, 1991; Eftekhary & Aminizadeh, 2012; Mykhaylenko, 2019). However, a critical

limitation persists: most studies conflate explicit verbalization (e.g., post-hoc rationalizations) with implicit cognitive operations (e.g., automatic phrase alignment). For instance, while TAPs reveal strategic decisions (Jääskeläinen, 1999), they fail to capture rapid, unconscious processes like syntactic priming or lexical competition (Bangalore et al., 2016; Schaeffer & Carl, 2013). This methodological ambiguity raises a pivotal question: Can behavioral experiments outperform verbal reports in detecting the cognitive dynamics of translation quality evaluation?

An important principle in translation quality is that the translation should faithfully reproduce the meaning of the source text. In semantics, the so-called “meaning” is often understood through the compositionality principle, which holds that the meaning of a complex expression is determined by the meanings of its constituents and their syntactic arrangement (Johnson, 2020).

While this principle has long been contested for its oversimplified account of meaning construction, especially in real-world language use, its core insight remains relevant: sentence structure and lexical choice are central to how meaning is construed. This study adopts the compositionality principle not in its strict sense, but as a heuristic to identify two key variables—phrase order and lexical choice—that significantly influence translation quality. Rather than assuming these factors exclusively determine meaning, they are treated as primary yet non-exhaustive contributors to interpretation, especially in cross-linguistic contexts where syntactic and lexical choices are closely linked with cultural and pragmatic considerations.

Phrase order refers to the arrangement of sentence components such as the subject, predicate, and complement. It plays a crucial role in determining which elements are foregrounded or backgrounded, and reflects the syntactic-pragmatic alignment between the source and target texts. A number of theorists have emphasized the importance of phrase order and its rearrangement in translation (Catford, 1965; Vinay & Darbelnet, 1995). Lexical choice refers to the selection of words or expressions in translation, which shapes the tone, nuance, and cultural resonance of the target text. Lexical choice determines semantic precision (Newmark, 1988; Nida, 1964). Prior research has examined these factors individually through text-based analyses (Antonio et al., 2020; Dastjerdi & Shoorche, 2011) or retrospective interviews (Bernardini, 2001), yet no study has systematically compared their effects using behavioral metrics (e.g., response times, RTs) against verbal reports. This gap is significant because cognitive theories posit that translation involves both conscious strategy selection (System 2) and automatic pattern recognition (System 1) (Kahneman, 2013), which may be dissociable through methodological design.

This study bridges translation theory with cognitive psychology by investigating whether implicit behavioral responses and explicit verbal accounts converge or diverge in assessing translation quality. Using poetry translation—a domain demanding nuanced semantic-structural balance—we test the hypothesis that behavioral methods capture cognitive dynamics inaccessible to introspection.

2. Literature Review

The cognitive mechanisms underlying bilingual processing, particularly in translation, have long intrigued researchers across disciplines such as linguistics, psychology, and cognitive science. Central to this inquiry is the interplay between implicit cognitive processes and explicit metacognitive awareness during language tasks. This review synthesizes key theoretical frameworks and empirical findings that contextualize the current study’s focus on translation quality evaluation (TQE), emphasizing the methodological and theoretical gaps addressed by our investigation.

2.1. Dual-System Theories and Language Processing

The foundational dual-system theory posits two distinct cognitive systems: System 1 (fast, intuitive, and automatic) and System 2 (slow, deliberate, and reflective) (Evans & Frankish, 2003; Kahneman, 2013). In bilingual processing, this dichotomy aligns with the differentiation between implicit procedural knowledge (e.g., syntactic parsing, lexical retrieval) and explicit declarative

knowledge (e.g., metalinguistic rules, translation norms). Paradis' (2004) neurolinguistic model further distinguishes these systems, arguing that implicit linguistic competence operates independently of conscious control, while explicit knowledge relies on metacognitive strategies stored in episodic memory. Neuroimaging, psycholinguistic experiments, and clinical observations support the system, revealing distinct neural substrates for implicit and explicit language operations (Ullman, 2001).

Empirical support for dual-system dynamics in translation comes from studies on cognitive effort and automaticity. For instance, translators tend to rely on automatic, fast System 1 thinking for literal translations, while engaging effortful, controlled System 2 thinking for more complex problems (Deckert, 2017; Silva & Pagano, 2017). This dual-processing model is supported by studies on trainee subtitlers, where inducing a switch from System 1 to System 2 thinking improved translation quality (Deckert, 2017). Recent studies on translation embody the dual-system dynamics. Carl and Dragsted (2012) showed that during translation, the monitor triggers sequential reading and writing when problems occur, while text copying and translation share similarities, reflecting the interplay of System 1 and System 2. Schaeffer and Carl (2013) proposed a recursive model, suggesting that translation involves activating shared representations between languages, which relates to System 1's automatic processing. However, when complex problem-solving is needed for adaptation to target norms, System 2's controlled processing comes into play.

2.2. Translation Quality Evaluation: From Rationalist Assertions to Cognitive Paradigms

Traditional approaches to TQE have been dominated by rationalist frameworks rooted in linguistic equivalence (Nida, 1964) and functionalist theories (House, 1997). Scholars such as Newmark (1988) and Vermeer (1989) emphasized fidelity to source text meaning or target text purpose, yet these criteria remain abstract and subjective, relying heavily on introspective judgments (Hansen, 2010). Such approaches mirror System 2 operations, privileging explicit, rule-based evaluations over implicit cognitive realities.

The advent of cognitive science methodologies revolutionized translation studies by shifting focus from product to process. Think-aloud protocols (TAPs), introduced by Ericsson and Simon (1984), allowed researchers to access translators' explicit reasoning (Jääskeläinen, 2002). However, TAPs face criticism for their susceptibility to post-hoc rationalization and limited ability to capture unconscious processes (Bernardini, 2001). As Venuti (2000, p. 339) noted, verbal reports reflect declarative knowledge but fail to account for procedural automatisms, "verbalization won't register unconscious factors and automatic processes"—a critical limitation given that translation inherently involves both systems.

Behavioral experiments, by contrast, offer a window into implicit cognition through metrics like response times (RTs) and error rates. Jiang's (2013) work on RTs in bilingual tasks highlights their sensitivity to cognitive load, revealing how lexical access and syntactic integration unfold in real time. In translation research, eye-tracking studies (Jakobsen & Jensen, 2008) and keystroke logging (Jakobsen, 2011) have illuminated the temporal dynamics of decision-making, demonstrating that System 2 processes dominate early stages or translators when dealing with difficult texts or complex translation problems (e.g., analytic mode), while System 1 governs later revisions or professional translators for easy texts or familiar domains (e.g., integrated mode) (Dragsted, 2005). These findings align with Chesterman's (2011) "literal hypothesis," which posits that translators default to literal strategies before engaging in conscious restructuring—a sequential interplay of dual systems.

2.3. The Behavior-Awareness Dissociation in Translation Evaluation

A growing body of evidence suggests a dissociation between behavioral responses and conscious awareness in cognitive tasks. In bilingualism, this phenomenon manifests in implicit language switching (Green, 1998) and covert lexical competition (Dijkstra & Van Heuven, 1998), where behavioral metrics (e.g., RTs) reveal interference effects absent in self-reports. Such

dissociations challenge the validity of introspective methods and underscore the need for triangulating data sources.

For TQE, this gap is particularly pronounced. Traditional evaluations rely on explicit criteria (e.g., “fluency,” “accuracy”), yet behavioral studies indicate that evaluators unconsciously prioritize structural congruence over lexical precision. For example, Schaeffer et al. (2019) found that translators’ eye movements during revision correlate with syntactic coherence rather than explicit error detection. Similarly, with eye tracker, Sjørup (2011) demonstrated that translators often adapted metaphors to align more naturally with the target language and culture, opting for equivalent imagery or paraphrase even when this meant departing from the source text’s lexical choices. These studies suggest that evaluators may lack introspective access to the cognitive heuristics driving their judgments—a disconnect central to the current study’s hypothesis.

2.4. Methodological Implications: Bridging Cognitive Science and Translation Studies

The integration of psycholinguistic methods into translation research has yielded significant advances. Recent studies have explored various aspects, including competence, expertise, mental load, linguistic complexity, and metacognition (Muñoz Martín, 2014). Researchers have investigated lexical and syntactic processing, temporal aspects, memory, executive functions, and reading patterns in translation (Chmiel, 2020). Lachaud’s (2011) multimodal approach, combining EEG, eye-tracking, and keystroke logging, exemplifies how triangulation enhances ecological validity. Similarly, Christoffels et al. (2013) used ERP to identify neural correlates of language conflict during translation, revealing implicit interference undetectable via verbal reports. These innovations align with broader trends in cognitive psychology, where implicit measures (e.g., priming, RTs) increasingly complement explicit methodologies (Jacob et al., 2021; Schoonbaert et al., 2011). This interdisciplinary approach, incorporating cognitive linguistics, psycholinguistics, and neurology, has enhanced our understanding of the translation process and the translator’s work (Rojo, 2015).

However, few studies have directly compared the sensitivity of behavioral and verbal report methods in TQE. While TAPs remain popular in process-oriented research, their limitations in probing System 1 processes are well-documented (Bernardini, 1999, 2001). Behavioral paradigms, by contrast, provide granular insights into automaticity but risk oversimplifying complex metacognitive phenomena. The current study addresses this gap by systematically contrasting these methodologies, thereby advancing theoretical debates on the nature of translation cognition.

3. Research Questions

This study addresses the following three research questions:

1. Do behavioral metrics (e.g., response times and ratings) and verbal reports differ in their sensitivity to syntactic-pragmatic manipulations (phrase order) during translation quality evaluation?
2. To what extent does lexical equivalence (accurate vs. approximate correspondence) modulate implicit cognitive effort, as indexed by processing speed, compared to explicit quality judgments?
3. How can the dissociation between automatic cognitive processes and conscious awareness be explained in the context of translation quality evaluation?

By situating translation studies within the paradigms of cognitive science, this research not only contributes to refining models of bilingual processing but also offers methodological insights for experimental design in interdisciplinary language research.

4. Methodology

4.1. Participants

Forty-four Chinese-English bilinguals (15 males, 29 females; mean age = 22.6 years) were recruited from the University of Macau, including 26 undergraduates and 18 postgraduates. All

participants were native Mandarin speakers with no prior expertise in translation or haiku poetry. Their English proficiency was rigorously controlled: undergraduates had passed standardized exams (e.g., Chinese National College Entrance Examination), while postgraduates met higher thresholds (e.g., TOEFL iBT ≥ 80 or IELTS ≥ 6.0). Participants were unbalanced bilinguals, having acquired English as a second language after age 8. Written informed consent was obtained, and the study protocol was approved by the university's ethics committee.

4.2. Materials

Twenty Chinese-English haiku pairs were selected from *Lighting the Bridge to the Moon: One Hundred and Eleven Macao Haiku* (Kelen, 2012), a corpus characterized by modern, non-rhyming structures to minimize cultural and stylistic biases. *Haiku* was chosen as the source material due to its brevity and structure: typically composed of three concise lines forming a single proposition, its fragmentary yet self-contained nature allows for controlled bilingual alignment and manipulation. Each haiku pair consisted of a three-line Chinese source text and its English translation. For experimental manipulation, four versions of each source text were created by varying two factors:

1. **Phrase Order:** Congruent (matching the target text's line sequence) vs. incongruent (reordered while preserving meaning).

2. **Lexical Equivalence:** Accurate (word-for-word correspondence) vs. approximate (semantically related but non-identical terms).

Ten additional "filler" pairs with deliberate grammatical errors were included to mask experimental manipulations. All materials were validated in a pilot study ($N = 30$ bilinguals), confirming significant differences in perceived equivalence between conditions ($ps < 0.001$).

4.3. Experimental Design

A 2x2 within-subject factorial design was employed, with phrase order (congruent/incongruent) and lexical equivalence (accurate/approximate) as independent variables. Dependent variables included:

1. **Explicit Ratings:** Translation quality scored on a 5-star scale (1 = poor, 5 = excellent).

2. **Implicit Metrics:** Response times (RTs) recorded from stimulus onset to button press.

4.4. Procedure

The experiment consisted of two phases: a behavioral task and a verbal report session.

During the behavioral task, participants were seated approximately 60 cm away from a 23-inch Tobii TX300 eye tracker (sampling rate: 300 Hz) and a response keyboard. Each trial began with the presentation of a fixation cross for 2,000 milliseconds, followed by a Chinese haiku displayed on the left side of the screen. After a 10,000-millisecond delay, its English translation appeared on the right side. Participants were instructed to evaluate the translation quality by pressing labeled keys corresponding to a 1–5 star rating scale, while their response times were recorded using E-Prime 3.0. The trials were distributed across four counterbalanced lists based on a Latin square design to control for potential order effects. The process is shown in Figure 1.

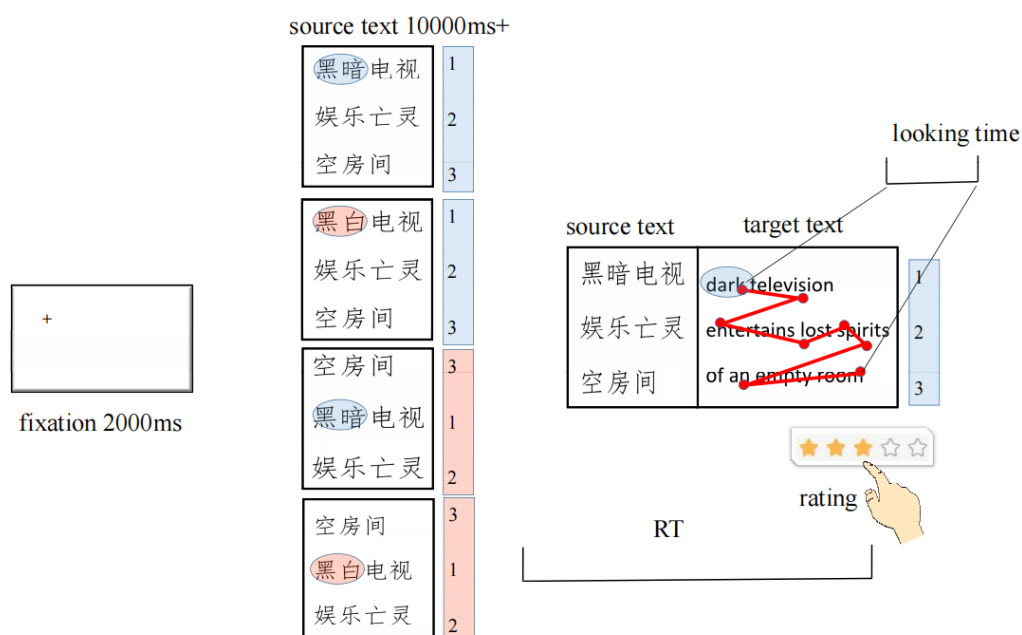


Figure 1. The process of the behavioral task.

In the second phase, participants took part in post-task interviews designed to elicit their explicit reasoning. These interviews involved open-ended questions, such as “Did line order affect your ratings?” Participants’ responses were audio-recorded and later coded in binary form (Yes/No) to facilitate subsequent analysis.

4.5. Data Analysis

All data were analyzed using SPSS 26.0 following a series of predefined steps. First, translation quality ratings were normalized by converting the original 1–5 star scores to a continuous 0–1 scale, with each star corresponding to an increment of 0.2. Descriptive statistics, including mean ratings and response times (RTs), were then calculated for each experimental condition to provide an overview of performance patterns.

To examine the effects of syntactic-pragmatic structure and lexical equivalence, a two-way repeated-measures ANOVA was conducted. This analysis tested both main effects and their interaction. Additionally, paired-sample t-tests were used to compare specific conditions—for example, congruent versus incongruent phrase order—with Bonferroni correction applied to adjust for multiple comparisons. To evaluate the verbal report data, a one-sample t-test was conducted against the chance level of 0.5, assessing whether participants’ explicit awareness significantly diverged from random guessing.

Effect sizes were reported to supplement statistical significance, with Cohen’s *d* used for t-tests and partial eta-squared (η^2) reported for the ANOVA. All inferential analyses adopted a significance threshold of $\alpha = 0.05$.

5. Results

5.1. Behavioral Data (Phase One)

A 2×2 repeated-measures ANOVA revealed a significant main effect of phrase order on translation quality ratings, $F(1, 43) = 9.469, p = .004, \eta^2 = .18$, and a marginally significant effect on response time, $F(1, 43) = 3.723, p = .06, \eta^2 = .08$. Translations with congruent phrase order were not only rated higher but also evaluated faster than those with incongruent order.

A paired-sample *t*-test confirmed a significant difference in ratings between translations with congruent order ($M = 0.75, SD = 0.11$) and those with incongruent order ($M = 0.696, SD = 0.16$), $t(43) = 2.720, p = .009, d = 0.35$ (see Figure 2). Response times were also marginally different between the

congruent order condition ($M = 20,314.05$ ms, $SD = 7,904.13$) and the incongruent order condition ($M = 21,868.57$ ms, $SD = 8,994.09$), $t(43) = 2.012$, $p = .050$, $d = 0.30$ (see Figure 3).

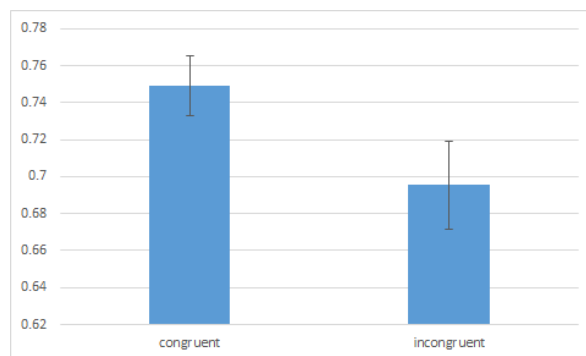


Figure 2. ratings of “goodness” with standard error lines in translations of congruent and incongruent orders.

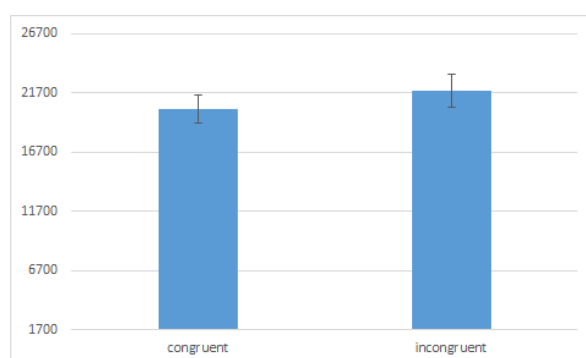


Figure 3. response time (ms) with standard error lines in the judgment of translations of congruent and incongruent orders.

The main effect of lexical choice was not statistically significant for quality ratings; both conditions yielded similar average scores around 0.72 (see Figure 4). In terms of response time, translation pairs with approximate lexical correspondence ($M = 21,371.31$ ms) were judged slightly slower than those with accurate lexical correspondence ($M = 20,579.46$ ms) (see Figure 5), though this difference was not significant.

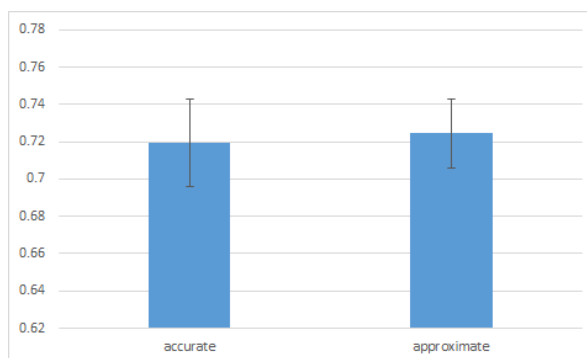


Figure 4. ratings of “goodness” with standard error lines in translations with accurate lexical correspondence and with approximate lexical correspondence.

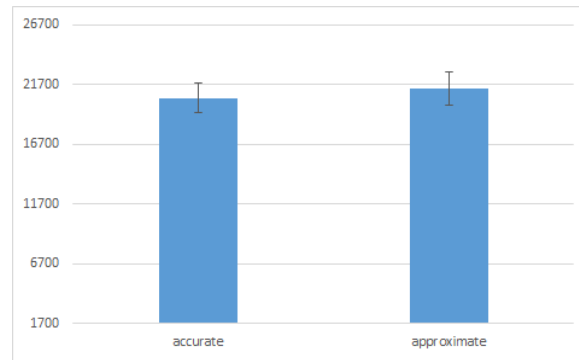


Figure 5. response time (ms) with standard error lines in the judgment of translations with accurate and translations with approximate lexical correspondence.

Importantly, a significant interaction between phrase order and lexical choice was found for quality ratings, $F(1, 43) = 6.166, p = .017, \eta p^2 = .125$ (see Figure 6). The effect of lexical choice on rating was more pronounced in the congruent order condition than in the incongruent order condition. However, the interaction effect between phrase order and lexical choice on response time was not significant, $F(1, 43) = 0.092, p > .05$ (see Figure 7).

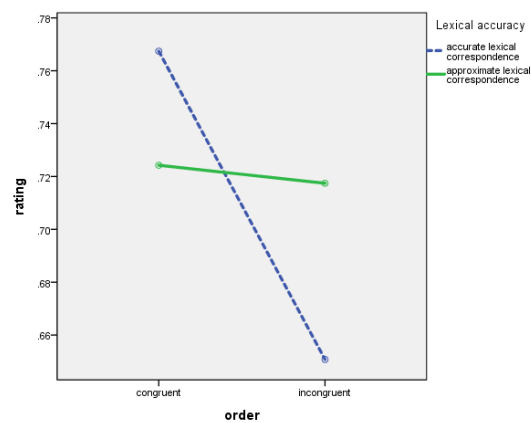


Figure 6. the interaction between order and accuracy in quality rating.

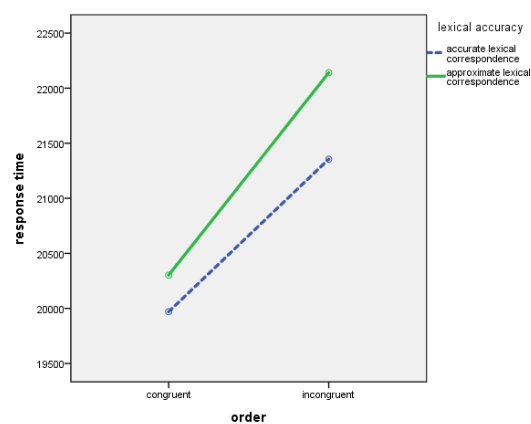


Figure 7. the interaction between order and accuracy in response time (ms).

5.2. Verbal Report (Phase Two)

When asked, “*Did the reversed line order in the target text affect your quality judgment, for example, lower your rating?*”, results showed that 72.7% of participants ($n = 32/44$) explicitly denied any influence of phrase order on their evaluations.

For the question “*Did non-equivalent lexical choices in the target text affect your judgment?*”, 80% of participants initially provided ambiguous answers (e.g., “*It depends*”). When pressed to give a definitive Yes/No response, answers split nearly evenly: 21 participants (47.7%) affirmed an influence, while 23 (52.3%) denied it.

6. Discussion

6.1. Sensitivity of Behavioral Metrics vs. Verbal Reports to Phrase Order Manipulation

Our findings provide robust empirical evidence that behavioral metrics are significantly more sensitive to phrase order manipulations than verbal reports, revealing a critical dissociation between implicit cognitive processes (captured by behavioral data) and explicit rationalizations (captured by verbal reports).

6.1.1. Implicit Processing and Structural Fluency: A Dual-System Account

In translation evaluation, phrase order—a syntactic-pragmatic feature requiring rapid syntactic alignment and pragmatic inference—primarily engages System 1. Participants’ faster RTs ($\Delta = 1,554$ ms) and higher ratings ($\Delta = 0.35$) for congruent-order translations (Figure 2–3) suggest that intuitive fluency drives preference for structurally aligned texts. Conversely, verbal reports, which require System 2 to articulate causal explanations, failed to detect this influence: 72.7% of participants explicitly denied that phrase order affected their judgments, despite behavioral evidence to the contrary.

This mismatch mirrors findings in bilingualism research, where implicit measures (e.g., priming, eye-tracking) often reveal language co-activation unnoticed by explicit tasks (Godfried & Winke, 2015; Wu & Thierry, 2012). Our study extends this paradigm to translation evaluation, demonstrating that structural fluency operates subconsciously, while conscious reflection prioritizes post hoc rationalizations (e.g., adherence to “free translation” ideals).

6.1.2. Conflict Monitoring and Subconscious Sensitivity to Phrase Order

The dissociation between behavioral metrics and verbal reports may also be interpreted through the lens of conflict monitoring theory (Botvinick et al., 2001). Incongruent phrase orders likely triggered a low-level conflict between participants’ syntactic expectations and the actual structure of the translation. Although participants did not report caring these discrepancies, the elevated response times suggest that such conflicts were detected and processed implicitly.

According to conflict monitoring theory, the brain continuously evaluates incoming stimuli for mismatches and activates control mechanisms when inconsistencies arise—even without conscious awareness. In this study, the increased cognitive effort required to resolve phrase order conflicts may have slowed RTs and subtly influenced quality ratings, despite participants’ denial of any syntactic influence. This supports the idea that translation evaluation involves not only higher-level rational processes (System 2) but also automatic conflict detection (System 1), which can shape preferences at a subconscious level.

Incorporating conflict monitoring thus deepens our understanding of the cognitive mechanisms underlying translation evaluation and reinforces the value of behavioral measures in capturing implicit processing effects that escape conscious introspection.

6.1.3. Methodological Implications: Implicit Measures and Cognitive Resolution

These findings carry important methodological implications. In some contexts, behavioral measures—such as reaction times—may reveal cognitive dynamics that participants themselves are

not consciously aware of, thereby outperforming self-reported data. Behavioral metrics, particularly RTs, are sensitive to online cognitive effort. The marginal RT difference ($p = 0.05$) between congruent and incongruent orders (Figure 3) suggests that incongruent structures demand additional parsing effort, even when participants are unaware of this cost. This aligns with psycholinguistic models of sentence processing, where syntactic violations increase cognitive load (Seeber & Kerzel, 2012; Vos et al., 2001).

Verbal reports, however, lack temporal resolution to capture such micro-level processes. When asked retroactively, participants reconstructed judgments based on explicit beliefs (e.g., “good translations prioritize meaning over form”), rather than recalling transient cognitive states. This explains why 72.7% denied phrase order effects, despite their behavioral choices favoring congruent structures.

Verbal reports can be vulnerable to social desirability bias (Grimm, 2010). In post-task interviews, several participants questioned the meaning of “quality” or replied “it depends,” indicating discomfort with simplistic Yes/No prompts. When pressed, many defaulted to socially sanctioned responses (e.g., “literal translation is rigid”) rather than introspecting actual decision-making. This bias is well-documented in translation studies, where professionals often disavow literal strategies despite behavioral evidence of their prevalence (Tirkkonen-Condit, 2005; Tirkkonen-Condit et al., 2008).

Behavioral metrics circumvent such biases by quantifying actions (e.g., button presses) rather than beliefs. The consistent pattern of higher ratings for translations with congruent phrase order reinforces the value of behavioral data in capturing implicit preferences that may not be accessible—or admissible—through verbal reports.

6.2. Lexical Equivalence Modulates Implicit Cognitive Effort Beyond Explicit Judgments

Our findings reveal a striking dissociation between the modulation of implicit cognitive effort by accurate vs. approximate correspondence. While lexical equivalence had no measurable impact on explicit ratings (both conditions: $M = 0.72$; Figure 4), it elicited a substantial increase in cognitive effort during evaluation, as reflected by slower response times (RTs) for approximate lexical pairs ($\Delta = 792$ ms; Figure 5). This divergence underscores the critical role of implicit processes in translation evaluation and challenges traditional frameworks that equate subjective judgments with cognitive reality.

6.2.1. Temporal Sensitivity to Lexical Competition

The observed dissociation is consistent with predictions from the Bilingual Interactive Activation (BIA) model (Dijkstra & Van Heuven, 1998), which suggests that lexical access in bilinguals involves the parallel activation of both languages, even when only one is task-relevant. For instance, while “dark television” represents an accurate lexical match to “黑暗电视” (“dark television”), and “black-and-white television” is only approximately equivalent, participants assigned nearly identical ratings to both translation pairs. This suggests that explicit judgments—such as quality ratings—may be governed by top-down strategies (e.g., prioritizing fluency or idiomaticity) that mask underlying lexical conflict.

In contrast, response times revealed a different picture. Judging approximate lexical matches like “黑白电视–dark television” took longer, indicating increased cognitive effort. This supports the notion of implicit semantic competition: approximate translations activate overlapping but non-identical semantic networks, prompting bilinguals to suppress competing meanings. Although participants may not consciously register this conflict, their behavioral responses—captured through reaction times—provide a more sensitive index of cognitive load. Thus, while ratings reflect participants’ conscious preferences or evaluative strategies, reaction times illuminate the automatic, lower-level processes involved in bilingual lexical access.

6.2.2. Methodological Implications: Task Demands and Cognitive Thresholds

This result has important methodological implications, because RTs usually capture what ratings miss. Explicit quality ratings require participants to integrate multiple evaluative cues—such as overall fluency, fidelity, and aesthetic appeal—which can dilute the salience of any single factor, including lexical equivalence. This task demand imposes a higher cognitive threshold, favoring global impressions over localized sensitivities. In contrast, response times (RTs) are sensitive to micro-level disruptions in processing and can effectively isolate the cognitive cost of specific manipulations, such as mismatched lexical items. The null effect of lexical equivalence on explicit ratings ($F < 1$) should not be interpreted as evidence of its irrelevance; rather, it underscores the limitations of holistic quality scales in capturing subtle, yet cognitively taxing, dissonances. RTs offer a more granular view of these implicit difficulties, revealing cognitive effort where conscious awareness—and thus verbalized evaluation—fails to detect it.

6.3. *Dissociation Between Conscious Awareness and Automatic Processes in Translation Evaluation*

Our findings provide compelling evidence for a dissociation between conscious awareness (verbalized preferences) and automatic cognitive processes (behavioral responses) in evaluating translation quality. What evaluators believe influences their judgments often diverges sharply from what actually drives their decisions.

In translation evaluation, phrase order manipulation—a syntactic-pragmatic feature—was processed rapidly and implicitly, engaging System 1, while verbal reports, which require causal explanations, relied on the more deliberate System 2. Our data highlight this distinction: In the behavioral data, congruent phrase order not only improved ratings ($d = 0.35$) but also marginally sped up judgments ($d = 0.30$), reflecting System 1's preference for structural fluency. However, when asked explicitly about the influence of phrase order, 72.7% of participants denied its impact, suggesting that System 2 either lacked access to or actively rationalized System 1's automatic processing. This pattern mirrors previous findings where syntactic processing is automatic and unconscious, revealing that, like all bilinguals, translators may lack metacognitive insight into the grammar-driven preferences that shape their decisions.

This dissociation between structural fluency and conscious attribution becomes even more evident when we consider the interaction between phrase order and lexical choice. A significant interaction was found: the effect of lexical accuracy on quality ratings was notably stronger when the phrase order was congruent than when it was incongruent. This conditional effect aligns with dual-system predictions. In predictable syntactic contexts (congruent order), automatic processes (System 1) have cognitive resources available to prioritize lexical-semantic coherence, making evaluations more sensitive to lexical precision. However, when phrase order is incongruent, the structural disruption imposes higher cognitive demands and triggers controlled processing (System 2) to resolve the mismatch. As a result, attention is diverted from finer-grained lexical features, weakening their influence on quality ratings. This dynamic reveals an adaptive interplay: System 1 dominates when structural fluency allows seamless integration of meaning, while System 2 steps in when disruption requires resolution. The modulation of lexical effects by phrase order thus provides strong support for the claim that implicit and explicit systems are context-sensitive in translation evaluation.

6.3.1. Attribution Errors and Fluency Heuristics

Participants in this study likely misattributed their processing fluency—such as the ease of parsing congruent phrase orders—to abstract notions of “quality,” rather than recognizing the role of syntactic alignment. According to processing fluency theory (Reber et al., 2004), fluency is hedonic in nature: people perceive fluent texts as better but often cannot pinpoint the underlying reasons. When the phrase order of the target text matched that of the source text, it created a more fluent processing experience. Detecting conflict requires cognitive effort and engagement of monitoring

processes (Botvinick et al., 2001). This congruence in syntactic structure aligns with the readers' linguistic expectations, making the text appear more natural and easier to comprehend, thereby enhancing the perceived translation quality. Conversely, an incongruent phrase order disrupted this fluency, making the text seem less natural and more difficult to process, which led to lower quality ratings. One participant, for instance, noted, "The good translations just flowed," without recognizing that this sense of fluency was facilitated by a priming effect between the source and target texts' phrase order. This misattribution helps explain why verbal reports tended to emphasize lexical choices (47.7% of participants affirmed their influence), even though behavioral data showed no impact on ratings. When forced to rationalize their choices, participants defaulted to more consciously accessible, salient factors like word choice, while overlooking the implicit influence of syntactic fluency that guided their behavior.

6.3.2. Social Desirability and Normative Biases

Verbal reports are also susceptible to social desirability bias, a tendency for participants to align their responses with socially accepted norms rather than reflecting on personal behavior. In post-task interviews, participants frequently invoked translation norms such as, "literal translation is rigid," responding in ways that were more in tune with academic or professional ideologies than their actual cognitive processes. This bias is particularly pronounced among unbalanced bilinguals—like our L1-Chinese participants—who may overcompensate for perceived deficiencies in their second language (L2) by espousing ideals of "free translation." Furthermore, when asked about the influence of lexical choices, 80% of participants initially responded with the equivocal phrase "It depends," indicating an awareness of the context-dependent nature of translation but an inability to articulate how specific lexical mismatches (such as "黑白电视" vs. "dark television") taxed their cognitive processing.

6.3.3. Temporal Decoupling: The "Fading" of Implicit Cues

Another factor contributing to the dissociation between conscious awareness and automatic cognitive processes is the temporal decoupling between implicit cues and verbal reports. Automatic processes, such as phrase order parsing, occur in mere milliseconds, whereas verbal reports depend on delayed, reconstructive memory. By the time participants were asked to reflect on their judgments in the interview phase, they had forgotten the transient cognitive states, such as brief moments of confusion when confronted with incongruent orders. Instead, they reconstructed their judgments based on more stable, long-standing beliefs, such as "I prioritize meaning over form." This "fading" of implicit cues is well-documented in decision neuroscience, where neural signatures of early sensory processing rarely reach conscious awareness, further supporting the idea that automatic processes shape our judgments in ways that are not always accessible to conscious reflection.

7. Conclusion

This study systematically compared the sensitivity of behavioral experiments and verbal reports in uncovering the cognitive mechanisms underlying translation quality evaluation, addressing three core questions: how syntactic-pragmatic manipulations (phrase order) differentially influence implicit and explicit judgments, to what extent lexical equivalence modulates cognitive effort, and how the dissociation between automatic cognitive processes and conscious awareness can be explained in this evaluative context.

Our results demonstrate that behavioral measures revealed systematic cognitive sensitivities to both syntactic and lexical manipulations, which were largely absent from participants' conscious reports. Phrase order significantly affected behavioral metrics, with congruent structures eliciting higher quality ratings and faster response times, yet most of the participants explicitly denied being influenced by order in their verbal accounts. Similarly, lexical equivalence modulated processing speed—participants took longer to respond to approximate translation pairs—despite assigning nearly identical quality ratings across conditions. These findings underscore a robust dissociation

between automatic cognitive processes and verbalized preferences: while behavior aligned with syntactic fluency and lexical precision, participants' explanations often defaulted to post hoc rationalizations, with about half citing lexical choice as influential even when rating data showed no such effect. That says, there is a dissociation between conscious awareness and automatic processes in translation evaluation. While evaluators believe they prioritize semantic fidelity or creativity, their behavior betrays a hidden preference for structural fluency—one they cannot articulate or even recognize.

This dissociation can be further illuminated through conflict monitoring theory (Botvinick et al., 2001), which posits that the brain continuously detects and resolves conflicts between competing representations or response tendencies. In our study, syntactic incongruities likely triggered internal conflict signals that delayed responses or increased cognitive load, even when participants remained unaware of these disruptions. The behavioral data—slower reaction times and increased hesitation—may thus reflect conflict detection processes operating below the threshold of conscious awareness. By contrast, the congruent sentence order is hedonically processed and thus perceived as more preferable. This provides a mechanistic account for why participants' explicit reports failed to align with their behavioral judgments: what they 'felt' to be right was governed by rapid conflict resolution processes that bypassed verbal articulation.

These findings extend dual-process theories to the domain of translation studies, demonstrating that syntactic fluency functions as a System 1 heuristic—rapid, automatic, and often inaccessible to conscious awareness—while verbal reports reflect System 2's slower, reflective reasoning processes that tend to rationalize rather than reveal underlying cognitive mechanisms. By showing that phrase order effects are robustly registered in behavioral responses but remain invisible in introspective reports, this study challenges traditional equivalence-based paradigms, such as Nida's (1964) model, which presuppose that translators and evaluators can reliably access and articulate the factors influencing their judgments. Our findings further contest equivalence-based frameworks (Nida, 1964; House, 1997) that prioritize explicit criteria such as "faithfulness" and "fluency." If evaluators cannot reliably introspect how syntactic features shape their judgments, such criteria risk reflecting normative ideals rather than cognitive reality. For instance, while participants verbally endorsed principles aligned with "dynamic equivalence" (e.g., naturalness and readability), their behavioral data revealed a consistent preference for "formal equivalence" (e.g., structural fidelity), suggesting that automatic cognitive responses may diverge sharply from consciously endorsed evaluation norms.

The study has significant practical implications for translator education and professional training. Our findings highlight a critical gap between what evaluators think influences their translation judgments and the cognitive mechanisms that actually guide them. For translator training, this dissociation underscores the need to cultivate metacognitive awareness of implicit biases—particularly those shaped by syntactic fluency and lexical familiarity. Pedagogical interventions might include encouraging trainees to compare their behavioral responses (e.g., reaction times, eye movements) with their own post hoc justifications, thereby making implicit processing more accessible to reflection. Such training could foster a deeper understanding of how translation quality is not only judged but felt, often before it is consciously articulated.

In addition to educational settings, these insights offer valuable directions for technological applications in machine translation (MT) and quality estimation systems. Current MT evaluation frameworks rely heavily on explicit metrics such as BLEU scores or post-editing distance, which often fail to capture the cognitive load experienced by users. Our findings suggest that integrating behavioral indicators—like processing speed or hesitation patterns—could improve the granularity and ecological validity of MT evaluation. By incorporating such cognitive metrics, automated systems could begin to approximate human processing effort more accurately, paving the way for more human-centered approaches to translation technology development.

Despite its contributions, the study has several limitations that constrain the generalizability of its findings. First, the experimental focus on short Chinese-English haiku-like texts—while ideal for

isolating specific linguistic manipulations—may not reflect the complexity of longer, more contextually embedded translations. Syntactic and lexical effects observed in concise segments might interact differently in extended discourse. In longer or more context-rich texts, syntactic and lexical cues may not act independently but interact dynamically—for instance, structural incongruities might amplify the processing cost of approximate lexical choices, or vice versa—thus altering the overall evaluative outcome. Moreover, the sample consisted primarily of unbalanced bilinguals whose dominant L1 likely heightened sensitivity to source-language structures. As such, the observed effects may not fully capture the behavior of balanced bilinguals or native English speakers.

Second, several important questions remain unresolved. Notably, the extent to which professional translators—who possess greater training, experience, and possibly metacognitive awareness—exhibit similar dissociations between verbalized judgments and implicit processing remains untested. Furthermore, individual cognitive variables such as working memory capacity, attentional control, or tolerance for ambiguity could moderate sensitivity to syntactic and lexical manipulations. Future research should investigate these factors using a broader range of text types, participant profiles, and neurocognitive tools to build a more nuanced understanding of how translation judgments emerge across contexts.

Looking ahead, future research should further explore both theoretical mechanisms and applied contexts of implicit processing in translation. Neuroimaging techniques like EEG could help identify neural correlates of automatic preferences—such as beta oscillations linked to syntactic alignment—thereby refining dual-process models of bilingual cognition. At the applied level, adapting this paradigm to real-world translation tasks (e.g., technical manuals) could test how well behavioral metrics like response times scale to professional settings, enhancing both assessment and training. By bridging cognitive science and translation studies, this work advances a paradigm shift: translation quality is not merely what evaluators say it is, but what their minds and behaviors reveal it to be.

Acknowledgments: This research was supported by Macao Polytechnic University [grant number RP/FLT-08/2022], which encourages faculty research through internal grant funding. The University provided financial support for this study but had no involvement in the study design, data collection, data analysis, or the preparation of the manuscript.

References

1. Afrouz, M., & Parsa, R. N. (2023). Book review: Bittner, H. (2020). Evaluating the evaluator: A novel perspective on translation quality assessment. *The International Journal for Translation & Interpreting Research*, 15(2), 184-188. doi:10.12807/ti.115202.2023.r03
2. Alves, F., Pagano, A. S., & Silva, I. A. L. d. (2010). A new window on translators' cognitive activity: Methodological issues in the combined use of eye tracking, key logging and retrospective protocols. *Copenhagen Studies in Language*, 267-291.
3. Antonio, J. D., Marins, L. C., & Pereira, L. P. (2020). Estratégias de segmentação e de tradução utilizadas por tradutores humanos: Da combinação de orações à estrutura retórica. *Gramática do Uso*, 21(1), 261-277.
4. Bangalore, S., Behrens, B., Carl, M., Ghankot, M., Heilmann, A., Nitzke, J., . . . Sturm, A. (2016). Syntactic variance and priming effects in translation. In M. Carl, S. Bangalore, & M. Schaeffer (Eds.), *New directions in empirical translation process research* (pp. 211-238): Springer.
5. Bell, R. T. (1991). *Translation and translating: Theory and practice*. London/New York: Longman.
6. Bernardini, S. (1999). Using think-aloud protocols to investigate the translation process: Methodological aspects. *RCEAL Working papers in English and applied linguistics*, 6, 179-199.
7. Bernardini, S. (2001). Think-aloud protocols in translation research: Achievements, limits, future prospects. *Target. International Journal of Translation Studies*, 13(2), 241-263.
8. Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol Rev*, 108(3), 624-652. doi:10.1037/0033-295x.108.3.624

9. Carl, M., & Dragsted, B. (2012). Inside the monitor model: Processes of default and challenged translation production. *Translation: Computation, Corpora, Cognition*, 2(1), 127-145.
10. Catford, J. C. (1965). *A linguistic theory of translation: An essay in applied linguistics*. Oxford: Oxford University Press.
11. Chesterman, A. (2011). Reflections on the literal translation hypothesis. In C. Alvstad, A. Hild, & E. Tiselius (Eds.), *Methods and strategies of process research* (pp. 23-36). Amsterdam/Philadelphia: John Benjamins.
12. Chmiel, A. (2020). Translation, psycholinguistics and cognition. In F. Alves & A. Jakobsen (Eds.), *The routledge handbook of translation and cognition* (pp. 219-238): Routledge.
13. Dastjerdi, H. V., & Shoorche, E. M. (2011). Word choice and symbolic language: A case study of persian translations of the scarlet letter. *International Journal of English Linguistics*, 1, 186.
14. Deckert, M. (2017). Asymmetry and automaticity in translation. *Translation and Interpreting Studies*, 12(3), 469-488. doi:10.1075/tis.12.3.06dec
15. Dijkstra, A., & Van Heuven, W. J. B. (1998). The bia model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189-225). Mahwah, NJ: Erlbaum.
16. Dragsted, B. (2005). Segmentation in translation: Differences across levels of expertise and difficulty. *Target*, 17(1), 49-70. doi:10.1075/target.17.1.04dra
17. Eftekhary, A., & Aminizadeh, S. (2012). Investigating the use of thinking aloud protocols in translation of literary texts. *Theory and Practice in Language Studies*, 2. doi:10.4304/tpls.2.5.1039-1047
18. Ericsson, K., & Simon, H. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
19. Evans, J. S. B. T., & Frankish, K. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, 7(10), 454-459. doi:10.1016/j.tics.2003.08.012
20. Godfroid, A., & Winke, P. (2015). Investigating implicit and explicit processing using L2 learners' eye-movement data. In *Implicit and explicit learning of languages* (pp. 325-348): John Benjamins Publishing Company.
21. Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1(2), 67-81.
22. Grimm, P. (2010). Social desirability bias. In *Wiley international encyclopedia of marketing*: Wiley-Blackwell.
23. Hansen, G. (2010). Some thoughts about the evaluation of translation products in empirical translation process research. *Copenhagen Studies in Language*, 389-402.
24. House, J. (1997). *Translation quality assessment: A model revisited*. Tübingen: Gunter Narr Verlag.
25. House, J. (2001). Translation quality assessment: Linguistic description versus social evaluation. *Meta: Journal des Traducteurs/Translators' Journal*, 46(2), 243-257.
26. House, J. (2004). Translation quality assessment: A model and its consequences. In M. P. N. Errasti, R. L. Sanz, & S. M. Omat (Eds.), *Pragmatics at work: The translation of tourist literature* (pp. 81-102). Bern: Peter Lang.
27. House, J. (2015). *Translation quality assessment: Past and present*. London/New York: Routledge.
28. Jääskeläinen, R. (1999). *Tapping the process: An explorative study of the cognitive and affective factors involved in translation*. Joensuu: University of Joensuu.
29. Jääskeläinen, R. (2002). Think-aloud protocol studies into translation: An annotated bibliography. *Meta: Journal des traducteurs*, 14(1), 107-136.
30. Jacob, G., Schaeffer, M., Oster, K., Hansen-Schirra, S., & Allen, S. E. (2021). Towards a methodological toolset for the psycholinguistics of translation: The case of priming paradigms. *Cognitive Linguistic Studies*, 8(2), 440-461.
31. Jakobsen, A. L. (2011). Tracking translators' keystrokes and eye movements with translog. In C. Alvstad, A. Hild, & E. Tiselius (Eds.), *Methods and strategies of process research* (pp. 37-55). Amsterdam/Philadelphia: John Benjamins.
32. Jakobsen, A. L., & Jensen, K. T. H. (2008). Eye movement behaviour across four different types of reading task. In S. Göpferich, A. L. Jakobsen, & I. M. Mees (Eds.), *Looking at eyes: Eye-tracking studies of reading and translation processing* (pp. 103-124). Frederiksberg, Danmark: Samfundslitteratur.
33. Jiang, N. (2013). *Conducting reaction time research in second language studies*. New York/London: Routledge.

34. Johnson, M. T. (2020). Compositionality. In D. Gutzmann, L. Matthewson, C. Meier, H. Rullmann, & T. E. Zimmermann (Eds.), *The Wiley Blackwell companion to semantics*: Wiley-Blackwell.
35. Kahneman, D. (2013). *Thinking, fast and slow* by Daniel Kahneman & Mindset - updated edition: Farrar, Straus and Giroux.
36. Kelen, K. (Ed.) (2012). *Lighting the bridge to the moon: One hundred and eleven Macao haiku*. Macao: Association of Stories in Macao.
37. Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33(2), 149-174.
38. Lachaud, C. M. (2011). Eeg, eye and key: Three simultaneous streams of data for investigating the cognitive mechanisms of translation. In S. O'Brien (Ed.), *Cognitive explorations of translation* (pp. 130-153). London: Continuum.
39. Muñoz Martín, R. (2014). A blurred snapshot of advances in translation process research. *MonTI, Special Issue 1*, 49-84. doi:10.6035/MonTI.2014.ne1.1
40. Mykhaylenko, V. (2019). Think-aloud protocol in translation: A pilot project. *International Humanitarian University Herald. Philology*.
41. Newmark, P. (1988). *A textbook of translation*. New York: Prentice-Hall International.
42. Nida, E. (1964). Principles of correspondence. In *Toward a science of translating* (pp. 156-192). Leiden: E. J. Brill.
43. Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins.
44. Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and social psychology review: an official journal of the Society for Personality and Social Psychology, Inc*, 8, 364-382. doi:10.1207/s15327957pspr0804_3
45. Reiss, K., & Vermeer, H. J. (1984). *Towards a general theory of translational action: Skopos theory explained* (C. Nord, Trans.). London/New York: Routledge.
46. Rojo, A. (2015). Translation meets cognitive science: The imprint of translation on cognitive processing. *Multilingua-Journal of Cross-Cultural and Interlanguage Communication*, 34(6), 721-746. doi:10.1515/multi-2014-0066
47. Schaeffer, M., & Carl, M. (2013). Shared representations and the translation process: A recursive model. *Translation and Interpreting Studies*, 8(2), 169-190.
48. Schaeffer, M., Jean, N., Anke, T., Katharina, O., Silke, G., & and Hansen-Schirra, S. (2019). Eye-tracking revision processes of translation students and professional translators. *Perspectives*, 27(4), 589-603. doi:10.1080/0907676X.2019.1597138
49. Schoonbaert, S., Holcomb, P. J., Grainger, J., & Hartsuiker, R. J. (2011). Testing asymmetries in noncognate translation priming: Evidence from rts and erps. *Psychophysiology*, 48(1), 74-81.
50. Seeber, K. G., & Kerzel, D. (2012). Cognitive load in simultaneous interpreting: Model meets data. *International Journal of Bilingualism*, 16(2), 228-242.
51. Silva, I. A. L. d., & Pagano, A. S. (2017). Chapter 6 cognitive effort and explicitation in translation tasks. In S. Hansen-Schirra, O. Czulo, & S. Hofmann (Eds.), *Empirical modelling of translation and interpreting*: Language Science Press.
52. Sjørup, A. C. (2011). Cognitive effort in metaphor translation: An eye-tracking study. In S. O'Brien (Ed.), *Cognitive explorations of translation* (pp. 197-214). London: Continuum International Publishing Group Ltd.
53. Tirkkonen-Condit, S. (2005). The monitor model revisited: Evidence from process research. *Meta: Journal des traducteurs/Meta: Translators' Journal*, 50(2), 405-414.
54. Tirkkonen-Condit, S., Mäkisalo, J., & Immonen, S. (2008). The translation process-interplay between literal rendering and a search for sense. *Across Languages and Cultures*, 9(1), 1-15.
55. Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, 4(2), 105-122.
56. Venuti, L. (Ed.) (2000). *The translation studies reader*. London/New York: Routledge.
57. Vermeer, H. J. (1989). Skopos and commission in translational action. In L. Venuti (Ed.), *The translation studies reader* (pp. 221-232). London/New York: Routledge.

58. Vinay, J.-P., & Darbelnet, J. (1995). Comparative stylistics of french and English: A methodology for translation. Amsterdam/Philadelphia: Benjamins.
59. Vos, S. H., Gunter, T. C., Kolk, H. H., & Mulder, G. (2001). Working memory constraints on syntactic processing: An electrophysiological investigation. *Psychophysiology*, 38(1), 41-63.
60. Williams, M. (2004). Translation quality assessment: An argumentation-centred approach: University of Ottawa Press.
61. Wu, Y. J., & Thierry, G. (2012). Unconscious translation during incidental foreign language processing. *NeuroImage*, 59(4), 3468-3473. doi:http://dx.doi.org/10.1016/j.neuroimage.2011.11.049

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.