Errors in Specialized Translation Training: A Corpus-Based Study on the Sight Translation of a Popular Science Article

Errores en el marco de la enseñanza de la traducción especializada: un estudio sobre un corpus de traducción a la vista de un artículo de divulgación científica

Kiara Giancola
University of Mons
kiara.giancola@umons.ac.be

Charlène Meyers
University of Mons
charlene.meyers@umons.ac.be

Abstract: In recent years, sight translation (SiT), the oral translation of a written text, has received little attention in the sphere of translator training. Rather, this practice has been almost exclusively studied within the framework of interpreting. Yet, based on the examination of the master's programs in translation (Giancola, 2022) recognized by the European Master's in Translation (EMT) (European Commission, 2022c), SiT is used, although rarely, as a course per se or as occasional exercises in translation training. Additionally, Dragsted and Hansen (2009) have emphasized its relevance and its benefits for translation students, which include improving speed rate, acquiring automatisms, and

Kiara Giancola y Charlène Meyers
Errors in Specialized Translation Training: A Corpus-Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

85

being more fluent. Based on these preliminary observations, the present study focuses on error patterns and speech disfluencies in SiT as part of specialized translation training. The study rests upon a corpus of twenty English-to-French SiTs of a popular science article about astrophysics performed by specialized translation students (Meyers, 2022). Building on existing work on errors in SiT, this study aims to find out the different error patterns observed in the corpus and the impact of speech disfluencies on the overall quality of the output. To analyze the corpus, an error typology in line with SiT has been developed based on established error analysis frameworks, such as MQM (Lommel et al., 2015) and Falbo's typology (1998), combined with new categories. Analysis of the corpus demonstrated the significant prevalence of meaning errors ($p < 0.05$), with a high number of terminological issues. This study also established the statistically significant association between errors and speech disfluencies ($p < 0.05$). The results suggest that SiT error patterns in specialized translation training differ from the observations made in the context of interpreting and that speech disfluencies negatively influence the overall performance. On this basis, it is recommended to explore SiT in specialized translation training independently and target speech disfluencies to enhance performance. Future research is needed to identify the ins and outs of SiT as part of specialized translation training —for instance, in legal translation training— and to determine whether it should become an integral part of the curriculum.

Keywords: Error assessment; Sight translation; Specialized translation training; Speech disfluencies.

Resumen: En los últimos años, la traducción a la vista, es decir, la traducción oral de un texto escrito, ha recibido poca atención en el ámbito de la formación de traductores. Más bien, esta práctica se ha estudiado casi exclusivamente en el marco de la interpretación. Sin embargo, el examen de los programas de máster en traducción (Giancola, 2022) reconocidos por la red *European Master's in Translation* (EMT) (Comisión Europea, 2022c) ha demostrado que la traducción a la vista se utiliza, aunque raramente, como curso propio o como ejercicios ocasionales en la formación en traducción. Además, Dragsted y Hansen (2009) han destacado su relevancia y sus beneficios para los estudiantes de traducción, incluso la mejora de la velocidad, la adquisición de automatismos y una mayor fluidez. Partiendo de estas observaciones preliminares, el presente estudio se centra en las tendencias de error y las disfluencias del habla en la traducción a la vista en el ámbito de la enseñanza de la traducción especializada. El estudio se basa en un corpus de veinte traducciones a la vista del inglés al francés de un artículo de divulgación científica sobre astrofísica realizadas por estudiantes de traducción especializada (Meyers, 2022). Partiendo de los trabajos existentes sobre los errores en la traducción a la vista, este estudio pretende averiguar las distintas tendencias de error observadas en el corpus y el impacto de las disfluencias del habla sobre la calidad general. Para analizar el corpus, se ha desarrollado una tipología de errores acorde con la traducción a la vista basada en tipologías de errores establecidas, como MQM (Lommel et al., 2015) y la tipología de Falbo (1998), y combinada con nuevas categorías. El análisis del corpus demostró la prevalencia significativa de errores de significado ($p < 0,05$), con un elevado número de problemas terminológicos. Este

Kiara Giancola y Charlène Meyers
Errors in Specialized Translation Training: A Corpus-Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - cc by-nc-sa

86

estudio también estableció la correlación estadísticamente significativa entre los errores y las disfluencias ($p < 0,05$). Los resultados sugieren que los errores de traducción a la vista en el ámbito de la formación en traducción especializada difieren de las observaciones realizadas en el contexto de la interpretación y que las disfluencias del habla influyen negativamente en el rendimiento general. Sobre esta base, se recomienda explorar, de forma independiente, la traducción a la vista en la formación en traducción especializada y centrarse en las disfluencias del habla para mejorar el rendimiento. En el futuro, habrá que investigar los pormenores de la traducción a la vista en el contexto de la formación en traducción especializada —por ejemplo, en la formación en traducción jurídica— y determinar si debería formar parte integrante del plan de estudios.

Palabras clave: evaluación de errores; traducción a la vista; enseñanza de la traducción especializada; disfluencias del habla.

1. INTRODUCTION

1.1. Aim of study

Translation pedagogy acknowledges that «there is clearly a place for a plurality of approaches and learning activities within a given curriculum, depending on student profiles, group heterogeneity, and intended outcomes» (Massey, 2017, p. 510). Among these approaches, we find exercises that stand out from written translation, such as sight translation (SiT), which is taught to interpreting and translation students.

SiT is defined as a hybrid task involving the oral translation of a written text (Dragsted & Hansen, 2009; Mellinger, 2017). It is a mix of translation and interpreting, the ST being in a written format and the delivery of the message being oral, like an interpreter who translates aloud a written speech. Generally, the exercise is divided into two main steps: students are given a text and have 5 to 10 minutes to read and prepare it without checking resources such as dictionaries or the Internet; then, they orally translate the written text. Although the practice of SiT by translation students has been shown to be helpful to improve speed rate, acquire automatisms, analyze the ST quicker, translate quicker and become more fluent (Dragster & Hansen, 2009), the study of SiT has most exclusively been analyzed in the realm of interpreting and has received little attention in the sphere of translation training. Yet, based on our own professional experience and that of colleagues, this tool is used in translation by international organizations, such as the UN and the European Commission. Translators are sometimes required to sight translate short communications, especially in times of crisis. It is therefore a real asset to have students, our future translators, capable of mastering this type of translation once on the job market.

Furthermore, the absence of any existing typology to study errors in SiT raises the challenge to find or create the appropriate tool to gauge SiT output. SiT being a hybrid

Kiara Giancola y Charlène Meyers
Errors in Specialized Translation Training: A Corpus-Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

87

between written translation and interpreting, as previously stated, it would be relevant to have a typology that is at the crossroads of these two practices.

In that regard, the aim of the present paper is to build an error typology for SiT and analyze errors and speech disfluencies made by translation students during their specialized translation course in order to portray error trends, suggest avenues of reflection to teachers, and ultimately contribute to the development and improvement of SiT training for translation students.

A corpus of twenty English-to-French SiTs performed by master translation students was selected for the analysis. To analyze this corpus, we developed an error typology by adapting and merging three existing typologies which were aimed at other purposes: the MQM typology (Lommel et al., 2015) developed for machine translation quality assessment, Falbo's typology of interpreting errors (1998), and Shreve et al.'s speech disfluency taxonomy (2011).

Although we specifically focused on a group of master students, the developed typology is not exclusively intended to analyze SiTs made by master students and could also be applied to SiTs by any other translation students.

1.2. *The training of students on specialized translation*

Translation pedagogy has emerged as a whole branch of translation studies in the study of and the search for «the best way of training translators and interpreters, and […] has remained a key concern of applied translation and interpreting studies» (Whyatt, 2022, p. 244). The aim of translation pedagogy is to equip students with sets of competences during their curriculum that will be useful. It is therefore centered around the notion of translation competence.

Nevertheless, there is a noticeable lack of research in the training of students in specialized translation. Although specialized translation is a fundamental backbone of the translation training as stated in the «competence framework» developed by the *European Master's in Translation*, literature on the training of students is scarce. The EMT is a network of universities across Europe, which have been awarded a quality label for their master program in translation (European Commission, 2022b). The selected programs must meet a set of professional standards and market requirements to become part of the network. The skills set by the European Commission reflect the market needs in the EU context, and therefore the current trends. Although it is important to highlight that the quantity of translation produced by other stakeholders, including specialized translation, is vast. These criteria encompass five areas of competence: language and culture, translation, technology, personal and interpersonal, and service provision. According to the translation competence of the EMT 2022 competence framework, translation students should «acquire, develop and use thematic and domain-specific knowledge relevant to translation needs (mastering systems of concepts, methods of reasoning, presentation standards, terminology and phraseology, specialised sources,

Kiara Giancola y Charlène Meyers
Errors in Specialized Translation Training: A Corpus-
Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

88

etc.)» (European Commission, 2022a, p. 8). In other words, students are expected to master specialized translation by learning how to acquire more knowledge depending on the field of study of the ST and how to use specific tools and specialized sources to build a robust terminology database. Learning specialized translation holds therefore a central place in the translator's curriculum and is considered essential to provide students with all the necessary working tools before entering the professional market.

1.3. *Sight translation in the classroom*

The concept of translation is large and encompasses a variety of practices. SiT, which can be defined as the oral translation of a written text (Agrifoglio, 2004; Dragsted & Hansen, 2009; Setton & Motta, 2007), is one of them. SiT has received little attention in the training of translation students and has rather been studied under the scope of interpreting training, probably because it «appears to be closer to interpreting than to translation because the output is oral» (Dragsted & Hansen, 2009, p. 590).

Yet, an examination of 92 Master's programs (Giancola, 2022) recognized by the EMT (European Commission, 2022c) reveals that SiT is, although in a small proportion, present in the training of translation students, either as a course *per se* or as practical exercises within their curriculum. The practice of SiT in the classroom comes with several benefits that are closely associated with two key concepts of translation pedagogy: deliberate practice (e.g*.,* Shreve, 2006) and self-reflection (e.g*.*, Gile, 1994, 2004).

In general terms, deliberate practice is the regular practice of activities aimed at acquiring professional skills (Ericsson et al., 1993). Transposed to translation training, it can be defined as «regular engagement in specific activities directed at performance enhancement in [translation]» (Shreve, 2006, p. 29). According to Ericsson (1996, p. 21), deliberate practice can only occur under the following conditions: (i) the task must be well-defined, (ii) the level of difficulty of the task is appropriate for the learner, (iii) feedback is included in the task, (iv) there must be opportunities for repetitive practice and error correction, and (v) the practice must run on a significant period of time.

Either as a course *per se* or as practical exercises, SiT meets all the criteria of deliberate practice. First, the task of SiT is well-defined before the beginning of the session. It is often broken down into a preparation time and a production time: translation students are allotted a specific amount of time (generally a few minutes, but it can vary depending on the length and difficulty of the text) to read and annotate a short ST in a way that will help them in their SiT of the ST with no access to any type of resources, such as dictionaries or web pages. During the production task, students are asked to sight translate the ST (through the use of a headset if the exercise is given in a language lab) in the most intelligible way, with the fewest pauses and repetitions possible and according to the standards of translation acceptability (acquired throughout their whole curriculum). Second, texts selected for SiT exercises are adapted to the task: they are short and their difficulty appropriate for learners. This allows translation students to

maintain a good level of attention. Third, since SiT is practiced on short texts, it leaves room for immediate informative feedback during the course. Immediate feedback is beneficial because it allows the translation trainer to correct errors «on the spot» and leave a deeper mark in the students' mind. Fourth, students have multiple opportunities to practice SiT either at home or in class. If SiT is practiced at home, teachers can still provide feedback by asking students to upload their performance on a MOOC platform. If it is practiced in class, several STs can be the object of SiT within a single course. Finally, whenever SiT is included in several courses (usually one or two courses per B languages), students have the opportunity to practice SiT multiple times on a weekly basis throughout the school year.

Additionally, courses of SiT allow students to self-reflect on their practice and progress. Indeed, students practicing SiT in language labs are given the possibility to directly listen to their production and retrospectively compare it to the correction provided. Dragsted & Hansen (2009, p. 602) stress three main benefits associated with the deliberate practice of SiT: improved speed rate of delivery, automatism acquisition in the choice of «fast-and-ready solutions», and enhanced fluence in the processing of larger translation units.

Retrospective self-reflection leads to student empowerment, which is «a process in which the learner occupies a central position in the classroom and is actively involved in experiential collaborative learning followed by self-reflection» (Whyatt, 2022, p. 248). Moreover, the retrospective self-reflection that comes with the practice of SiT is a stepping-stone to experiential learning. Following Kolb's (1984) four-stage experiential learning model, Massey (2019) argues that translation students are empowered by (i) engaging in concrete experience, which triggers (ii) self-reflection on that experience. Students then reach (iii) abstract conceptualization learned from the experience and become able (iv) to apply the acquired knowledge to other settings.

Moreover, a growing number of scholars agree that translation and interpreting students should be given the opportunity to train on off-center practices, that are skills-based and transferable to other activities (Calvo, 2011) in order to «provide more than traditional translation services» (Gaspari et al., 2015, p. 336) and hence to better fit the ever-changing job market.

Other scholars, such as Dragsted & Hansen (2009), even encourage translators and translation students to integrate SiT as a deliberate practice in their daily routine because of its high time and effort efficiency when jointly practiced with a speech recognition software: «Our analyses support our overall assumption that working in the oral modality, *speaking your translation*, using speech recognition technologies, which are rapidly gaining ground, seems to have a lot to offer in terms of saving time and effort without compromising the output quality significantly» (Dragsted & Hansen, 2009, p. 602). They also highlight that time saved can also be repurposed for revision. According to the experience of our colleagues, it was common practice, in the past, to record SiTs using a dictaphone and have them transcribed by typists at a later time (although some translators keep on using this tool to translate more quickly). This process was

Kiara GIANCOLA y Charlène MEYERS
Errors in Specialized Translation Training: A Corpus-
Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

90

a time saver and allowed translators and/or revisers to produce a final, revised version by correcting the typed draft. SiT is also often used in on-site interpreting, especially in public service and institutional settings. These examples highlight the practical applications of SiT and underscore the importance of training students in this skill.

Nevertheless, although the practice of SiT comes with several benefits and although there are many opportunities to study this particular form of translation, especially in any translation classroom where SiT is taught, there is a noticeable lack of research in SiT. It is our belief that translation studies underexploit the potential of case studies (Massey, 2017, p. 510) and that more «hands-on» research on SiT is necessary as case-studies increase our «understanding of the processes involved in the development of translator competence» (Kiraly, 2013, p. 222).

1.4. *Error assessment in translation and sight translation*

Translation quality assessment is an integral part of the translation process, but the evaluation is often based upon personal judgment and subjectivity (Lommel et al., 2014). For instance, an incorrectly or imperfectly translated segment can be categorized as a certain type of error by some revisers and teachers, while others might consider it correct, classify it under another type of error, or just consider that it needs to be finessed. As a result, it is no surprise that analyzing translation errors, either for research or pedagogical purposes, can be very demanding, even with a well-developed error typology. As Baxter (2017) and Martín Martín (2010) have pointed out, it is challenging to establish an error typology that is objective, exhaustive, and capable of being applied globally, especially when assessing translation in pedagogical contexts because teachers tend to have their own way of identifying and assessing their students' productions. In addition, Lommel et al. (2014) consider that without clear general guidelines, quality assessment can fall short of objectivity, but it would be hard to come up with a universal error tagging for translation quality, since translation is not an exact science. However, this underexplored scientific area provides the perfect space for investigating and developing the potential of an error typology for each form of translation (i.e., written translation, machine translation, interpreting and SiT) that could be applied on a large scale with exhaustive guidelines and could provide a framework to reduce subjectivity and increase accuracy.

Although there is no single error framework that has been put forward to be used globally in translation, the literature on the matter agrees on the fact that meaning accuracy holds a central place in translation, and even interpreting, quality. Even if the classification and identification of errors can be knotty (Gile, 1997/2002; Havnen, 2019; Stenzl, 1983), there are also clear aspects that stand out, such as expression and grammar (form errors), on the one hand, and meaning issues (content errors), on the other hand (Agrifoglio, 2004; Falbo, 1998; Gile, 1984, 1989, 1995, 1997/2002).

Kiara GIANCOLA y Charlène MEYERS
Errors in Specialized Translation Training: A Corpus-
Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

91

When it comes to SiT, studies have been conducted on errors, mostly in the realm of interpreting (Agrifoglio, 2004; Gile, 1983, 1985a, 1990, 1995), but it remains under-explored (Li, 2014). Some have analyzed errors between written and SiT (Baxter, 2017; Dragsted & Hansen, 2009), while others have compared SiT with simultaneous and consecutive interpreting, notably Agrifoglio (2004). However, it appears that no error typology specifically designed for SiT has been developed so far. Yet, error analysis, and thus a proper error typology, is essential to investigate trends, but also highlight the advantages and shortcomings of SiT and draw recommendations to improve its practice among translators and translation students.

2. OBJECTIVES

Based on the above observations, the objective of the paper is twofold. The first objective is to present an error taxonomy which can be specifically used to assess the quality of SiT. The second objective is to present a thorough analysis of a SiT corpus through the application of the proposed error taxonomy, and therefore to highlight key trends in different error categories. With this approach, we are also measuring two aspects: how useful and appropriate this combined error detection tool is and how better it helps us to categorize students' performance.

3. METHODOLOGY

3.1. *An error taxonomy for sight translation*

3.1.1. Overview of the typology

Before exploring the details behind the development of our error typology in SiT, it is crucial to provide an overview of the typology itself, including its various categories and subcategories. The typology is constructed based on a set of criteria such as the type of error, its source, and the context in which it occurs. The creation of this specialized error typology for SiT was made possible by the combination of two established error frameworks, one used in translation called Multidmensional Quality Metrics (Lommel et al., 2015b) and one in interpreting called *grille d'erreurs en interpretation* (Falbo, 1998); our own implementation of a more comprehensive approach towards mistranslation errors; and one speech disfluency taxonomy (Shreve et al., 2011).

By organizing errors into distinct categories and subcategories, the typology facilitates the identification and analysis of different types of errors, enabling researchers and teachers to gain a better understanding of their causes and effects. Thus, in order

Kiara Giancola y Charlène Meyers
Errors in Specialized Translation Training: A Corpus-Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

92

to fully comprehend the rationale behind the creation of our error typology, it is essential to first have a clear portray of its categories and subcategories.

Before delving into our final typology presented in Figure 1, the nomenclature and definitions employed to denote the groupings and subdivisions of the ultimate typology utilized is hereby explicitly established to ensure accuracy and lucidity:

1. The term «division» refers to the three categories that divide the typology into three main subgroups: the content division, the form division and the speech division.

2. The term «dimension» relates to each high-level category that falls under one of the divisions and that comprehends all the relevant issue types associated with the dimension (Lommel et al., 2015a).

3. The term «issue type» is employed when mentioning a category from a determined dimension that portrays a particular type of error (Lommel et al., 2015a).

4. Finally, the term «error» is exclusively used when referring to a confirmed error found in a text (Lommel et al., 2015a).

The typology is divided into three main divisions: **content, form, and speech**. Each division contains subcategories called «dimensions».

In the **content division**, which focuses on the shifts in meaning, there are two dimensions: accuracy and terminology.

Accuracy holds eight issue types:
– omission: content is missing from the translation that is present in the source (Lommel et al., 2015b);
– addition: the target text includes text not present in the source (Lommel et al., 2015b);
– lexical mistranslation: the translator uses the wrong translation among the translation possibilities of a word that could have otherwise been translated that way in another context (our own addition to the typology);
– mistranslation: the target content does not entirely represent the source content, with a target segment that can be less or more specific and intense than the source segment (our own addition to the typology);
– nonsense: the translation of the source content makes no sense (our own addition to the typology);
– opposite meaning: the target content conveys the opposite meaning of the source content (our own addition to the typology);
– overly literal: translation is overly literal (Lommel et al., 2015b); and
– false friend: the translation has incorrectly used a word that is superficially similar to the source word and does not correspond to a possible and correct translation of the SL word in any other context (Lommel et al., 2015b).

Terminology holds one issue type:

– incorrect terminology: a term (domain-specific word) is translated with a term other than the one expected for the domain or otherwise specified (Lommel et al., 2015b).

In the **form division**, which focuses on form, grammar, syntax and style, there are two dimensions: grammar and style.

Grammar holds six issue types:

– agreement: two or more words do not agree with respect to case, number, person, or other grammatical features (Lommel et al., 2015b);
– part of speech: a word is the wrong part of speech (Lommel et al., 2015b);
– tense/mood/aspect: a verbal form displays the wrong tense, mood, or aspect (Lommel et al., 2015b);
– function words: a preposition, 'helping verb,' article, determiners used incorrectly (Lommel et al., 2015b);
– word order: the word order is incorrect (Lommel et al., 2015b);
– cohesion: portions of the text needed to connect it into an understandable whole (e.g., reference, substitution, ellipsis, conjunction, and lexical cohesion) are missing (Lommel et al., 2015b).

Style holds three issue types:

– register: the text uses a level of formality higher or lower than required by the specifications or general language conventions (Lommel et al., 2015b);
– lexical awkwardness: words that are inappropriate to the context or do not exist, but thanks to the context, it is possible to understand what the speaker meant to say (Falbo, 1998);
– idiomatic awkwardness: expressions, in the broad sense, that are inappropriate to the context or do not exist, but thanks to the context, it is possible to understand what the speaker meant to say (Falbo, 1998).

In the **speech division**, we refer to speech disfluencies that can arise when sight translating. By speech disfluencies, we mean interruptions or disruptions in the flow of speech. There are four issue types:

– unfilled pause: interruptions in speech of at least one second that do not contain any phonetic content (Shreve et al., 2011);
– filled pause: interruptions in speech that are filled with utterances such as [euh] (Shreve et al., 2011);
– repetition: when the speaker repeats part of the utterance (Shreve et al., 2011);

Kiara GIANCOLA y Charlène MEYERS
Errors in Specialized Translation Training: A Corpus-
Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

94

– repair: when an utterance is repeated, but with an alteration to one or more constituents (for instance by replacement, omission, or insertion of words), and the other constituents generally preserved (Shreve et al.; 2011). Within the repair issue type, there is a distinction between announced and unannounced repairs:

- announced repair: repairs for which students explicitly announce that they are correcting the initial segment by saying «excuse me,» «excuse me let me start again,» «sorry,» «no» (our own addition to the typology to determine whether students were used to announcing their repairs before correcting what they had just said);

- unannounced repair: repairs for which students do not announce that they are correcting the initial segment (our own addition to the typology).
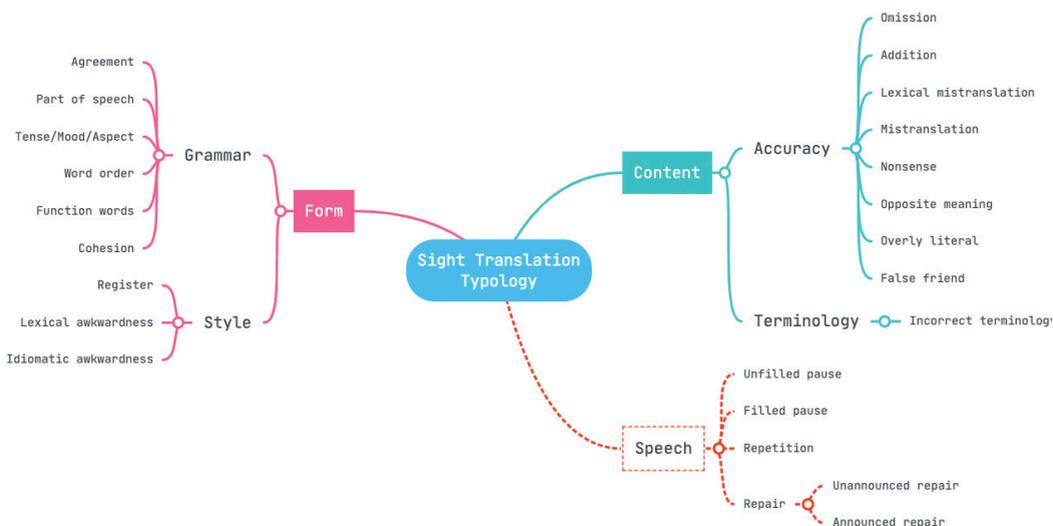

Figure 1: The SiT typology (mind map created with ©Mindmeister [MeisterLabs, 2022]).

3.1.2. The selection of established typologies for the development of the final typology

Considering that SiT is halfway between translation and interpreting (Agrifoglio, 2004), it was important to consider both aspects and thus consider typologies used in the two practices. There might be some error categories that could apply to both interpreting and SiT, since both contain an oral dimension (Gile, 1997/2002), while there might be categories that are only addressed in translation and SiT as they both share a common written denominator that interpreting does not have. Consequently, two established frameworks, one in translation and one in interpreting, were analyzed

Kiara GIANCOLA y Charlène MEYERS
Errors in Specialized Translation Training: A Corpus-
Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

95

and used to develop the final error taxonomy of our study: the Multidimensional Quality Metrics (MQM) by Lommel et al. (2015) and Falbo's interpreting typology (1998). Additionally, we worked on a more thorough approach for dealing with mistranslation errors by creating new categories. Finally, Shreve et al's speech disfluency typology used in SiT was also incorporated into the framework to address the speech disfluencies (i.e., any interruption or disruption in the flow of speech) that arise when sight translating.

First, MQM was promising because it offers structured customization and «provides a framework for describing and defining quality metrics used to assess the quality of translated texts and to identify specific issues in those texts» (Lommel et al., 2015a, «Introduction [non-normative]»). It is designed to assess translation quality and applies to any translated content, be it machine translation or human translation. It is therefore not strictly limited to machine translation. Furthermore, «it provides a functional approach to quality that seeks to see whether a translation meets specifications and to identify aspects that may fall short of expectations.» (Lommel et al., 2015a, «Scope»). The framework is based on a wide range of categories and subcategories of issues, i.e., «a potential error detected in a text» (2015a, 2. Terms and definitions). Each of these issue types is accompanied by a definition and a set of criteria, such as examples in context.

Second, we focused on Falbo's taxonomy (1998), which is used in the interpreting field and focuses on contributing to research in this area and on providing interpreting students with an evaluation grid that highlights the errors they made. As far as her typology is concerned, it was developed as an analytical framework divided into two main categories: content and form. The «content» subcategories revolve around translation loss, that is to say, «the incomplete replication of the source language text in the target language text» (Dizdar, 2014, p. 2). The «form» subcategories revolve around grammatical, lexical, and idiomatic issues.

However, during the analysis of the typologies, we realized that the mistranslation issue type proposed by MQM did not entirely match our needs and was too vague. MQM defines a mistranslation as follows: «the target content does not accurately represent the source content» (Lommel et al., 2015b, 1.2. Detailed listing of MQM issue types). Nevertheless, this definition is overly general, and it does not entirely encompass all the patterns and severity levels of meaning errors that one may encounter in a text. Yet, the core of translation assessment rests upon the fidelity of the ST (e.g., Dragsted & Hansen, 2009; Gile, 1995). In fact, the meaning transfer of a translation perfectly reflecting the ST must be correct, precise, and complete (Dussart, 2005). Consequently, if the ST is mistranslated, then the translation does not faithfully convey the ST message. In the case of students, their main objective is to understand the ST and translate it into French by respecting the meaning conveyed by the ST (Dragsted & Hansen, 2009; Gile, 1995). Therefore, we decided to redefine the mistranslation issue type and add new issues to the accuracy dimension to tackle the different degrees of meaning issues so as to pinpoint flaws, provide an overview of the most common types of meaning errors, and ultimately adapt SiT training, if necessary. In that spirit, instead of only having one mistranslation issue type, we dug deeper and developed a more

Kiara Giancola y Charlène Meyers
Errors in Specialized Translation Training: A Corpus-Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

96

comprehensive approach to mistranslation by making a difference between a mistranslation (slight change in meaning), a lexical mistranslation (use of the wrong translation among the translation possibilities of a word that could have otherwise been translated that way in another context), a nonsense (a translated utterance that makes no sense), and an opposite meaning error (when the translated content is the opposite meaning of the source content).

Finally, a preliminary overview of the corpus revealed at first glance that the SiTs were filled with repetitions, pauses, and similar phenomena. It was therefore impossible to ignore this aspect. As Prak-Derrington (2011) puts it, translation—and by extension, SiT—is «a movement of reformulation from one language to another» (p. 293). SiT is noticeably confronted with this issue considering that sight translators, by translating a written text aloud after having had a few minutes of preparation (without resources), are at risk of introducing repetitions and other speech disfluencies. When translators «speak their translation», speech disfluencies can arise and take many forms, including repetitions of words or sounds, pauses, interjections such as «um» or «ah,» and other types of disruption. In fact, sight translators deliver a translation straightaway and do not have the ability to erase or review segments like in written translation. Moreover, while written translation allows translators to reflect before sending their final output, SiT does not offer such an option and is rather spontaneous.

Shreve et al. (2011) further stress this observation by stating that «because output in SiT must be verbalized, disruptions will manifest not just as errors or deficiencies in rendering, but also as speech disfluencies in the oral performance of the translation.» (p. 94). In fact, Shreve et al. (2011) have dedicated a whole study to SiT and speech disfluency to provide an overview of the cognitive phenomena in action when sight translating (with eye-tracking data, as well). These researchers' work and findings are the most comprehensive in the field of SiT. As a result, we thoroughly examined their work to incorporate and utilize it in our typology. Theirs is believed to be the only study that provides an in-depth focus on speech disfluencies in the realm of SiT. They developed their own speech typology based on a corpus-driven approach and put forward four categories of disfluencies that are thoroughly documented in the literature (Shreve et al., 2011, p. 98), and can be found in our own typology: unfilled pauses, filled pauses, repetitions, and repairs. It should be noted that the corpus for this study already included unfilled and filled pauses. Indeed, Meyers (2022) annotated the transcripts with unfilled and filled pauses of at least one second because pauses of 1-2 seconds have been shown to indicate some translation task-related cognitive processing (Dragsted & Hansen, 2009, p. 593).

The combination of existing issue types from MQM and Falbo, as well as our development of a more comprehensive mistranslation issue approach and the addition of Shreve et al.'s speech disfluency framework, contributed to the completion of an error typology specially designed for SiT, which has been portrayed in Figure 1.

Kiara Giancola y Charlène Meyers
Errors in Specialized Translation Training: A Corpus-
Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

97

3.2. *Compilation of a sight translation corpus*

To build this corpus, twenty Master students were asked, as part of their English-to-French specialized translation course, to sight translate an article from the *New Scientist* called «Looking into the Voids Could Help Explain Dark Energy» (Ananthaswamy, 2015). These students performed the task during their usual scientific and technical class (Meyers, 2021).

The ST is a popular science article written in English with a special focus on astrophysics. The text explains how voids in the universe may explain the existence of dark energy and lead to different scenarios for the faith of the universe. Since the study of voids and dark energy is still in its early stages, much of the information being discussed remains unproven and the observations are largely based on hypotheses. All the participants of the study sight translated the first 375 words of the article. It is important to highlight that we did not reveal the subject, nor did we give information or vocabulary prior to giving the text for the experiment, so we expected the students to struggle with some specialized terms that appear in the article.

To guarantee objectivity and impartiality of the experiment, students were not initially given any information about the conducted experiment. Instead, they were asked to perform a usual type of exercise they were given in class, namely the SiT from English to French of a popular science article from the *New Scientist*. They were given precisely ten minutes to read, prepare, and analyze the text without being allowed to use dictionaries or surf the Internet, relying only on context and their own knowledge. Then, they started recording their output in class through a headset with an integrated microphone. Once they had finished recording their SiT in the language lab, they were asked whether their target text (TT) could be used for research purposes. SiTs of students who had not given their consent for data use were immediately discarded. It is worth mentioning we did not measure time employed by each student when performing the exercise. After the experiment, the recordings were compiled and transcribed to build a corpus, which is available in .docx and .txt formats (Meyers, 2022). This corpus was chosen for the present study in order to analyze the errors made by the students and classify these in accordance with the proposed error typology tailor-made for SiT.

4. HYPOTHESES BASED ON THE FINAL TYPOLOGY

The study is expected to yield several outcomes:

– The participants in the experiment are students who may lack experience and are therefore more likely to make meaning errors. However, as they are master's students with translation skills, we do not anticipate many serious meaning errors that would significantly compromise the source text's message.

- Students may struggle with specialized terms as they are not familiar with the subject of the article and were not given any explanation before analyzing the text.
- Stress and cognitive overload may influence the output and result in different error trends. Baghi & Khoshsaligheh (2019) found that translation students tend to feel anxious and make errors when performing SiT, particularly when their coordination effort is saturated.
- There is a risk that the translator may inadvertently use a more informal register when delivering the oral translation because oral language is generally less formal than written language (Alic, 2020).
- Certain issue types, such as lexical mistranslation or overly literal translations, could be explained by the visual interference phenomenon associated with SiT.
- The study is expected to reveal many agreement errors as previous research has already identified agreement errors as the most common grammar errors in SiT (Agrifoglio, 2004).

5. ANALYSIS

Each TT was analyzed individually based on the SiT typology that we have already described. Each error found in the corpus was associated with the code of the corresponding error type. We decided to count every occurrence of error, even when the student made the same error more than once. The TTs were coded and analyzed, then converted into .txt files. The concordancer *AntConc* was subsequently used to identify, classify, and count the errors for each TT and for the overall output. AntConc (Anthony, 2020) is a «freeware corpus analysis toolkit for concordancing and text analysis» (Anthony, 2022, main page).

6. RESULTS

6.1. *Form or Content: which type prevails?*

Out of the three main categories already described (form, content and speech), content was the most prominently identified in the corpus. To determine the extent of this trend, a chi-squared test on the frequency of the content and form errors in all TTs was performed and gave rise to the following results: $\chi^2(19) = 30.957$, $p = 0.04082$. The results of the test show statistical significance ($p < 0.05$), but the effect size is weak (Cramer's $V = 0.187$). In other words, although the strength of difference is weak, the results are not due to mere chance, and there is a statistically significant difference between content and form errors, with more content errors in all TTs than form errors.

Kiara GIANCOLA y Charlène MEYERS
Errors in Specialized Translation Training: A Corpus-
Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

99

Additionally, the correlation between the two variables was analyzed to identify the general trend in data. In this case, the results show that the correlation is weak ($r = 0.1024633$). That is, if the number of content errors increases, then the number of form errors will slightly increase as well, and vice versa. In Figure 2, the linear regression is portrayed in red, while the blue dots represent the TTs. This linear regression shows the weak correlation between the two variables.
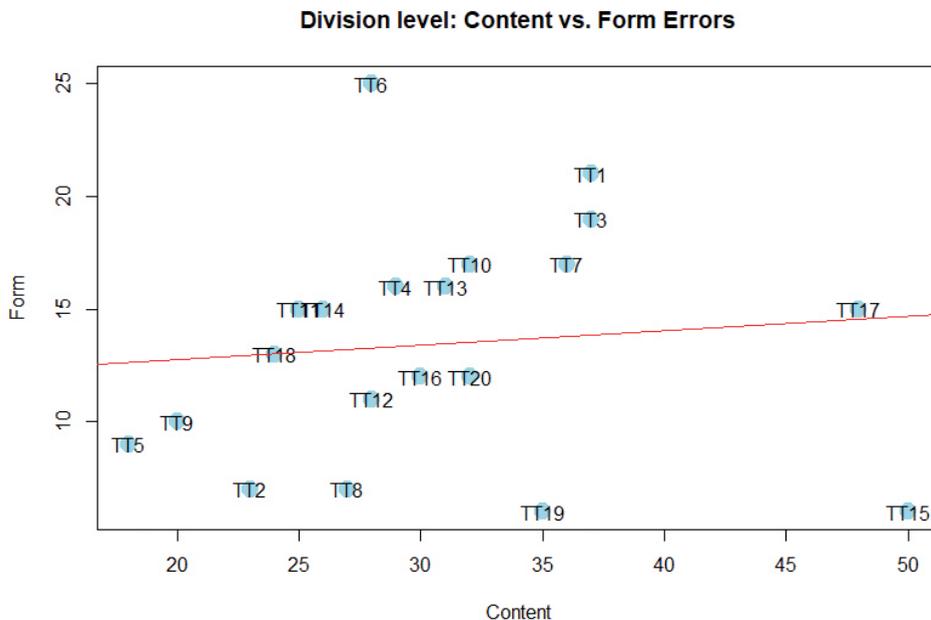

Figure 2: Scatter plot of content vs. form errors.

Given the above, the results contradict that of Agrifoglio (2004) and Gile (1985a), according to which there are more form and expression errors than meaning errors in SiT.

The difference could be explained by the fact that the participants in Agrifoglio's and Gile's experiments were professional interpreters, whereas this research project analyzed the output of translation students who are still in the learning process of becoming professional translators. Although it is true that, in this case, the students were used to practicing SiT (Meyers, 2021), they were still students and were more likely to make content errors as they lacked real-life practice and understanding of the SL. It should also be noted that the popular science article included specialized terms, which made it more difficult for students to understand the meaning of the ST. The text refers to concepts associated to voids and dark energy, which can become challenging when

Kiara GIANCOLA y Charlène MEYERS
Errors in Specialized Translation Training: A Corpus-
Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

100

one cannot consult any databases or dictionaries. The interpretation of the meaning of the ST is therefore based on the general knowledge of the students, who may lack experience.

6.2. *Error frequency in the content division*

Figure 3 showcases the frequency of the total number of the errors that fall under the issue types from the accuracy dimension, which are omission, addition, lexical mistranslation, mistranslation, nonsense, opposite meaning, overly literal, and false friend. Amid the errors that were classified in the accuracy dimension, mistranslation is the most identified issue type (n = 174), followed by omission (n = 82), lexical mistranslation (n = 77), overly literal (n = 45), and addition (n = 42). The frequency of opposite meaning errors is the lowest (n = 2), followed by nonsense (n = 8), and false friend (n = 9).
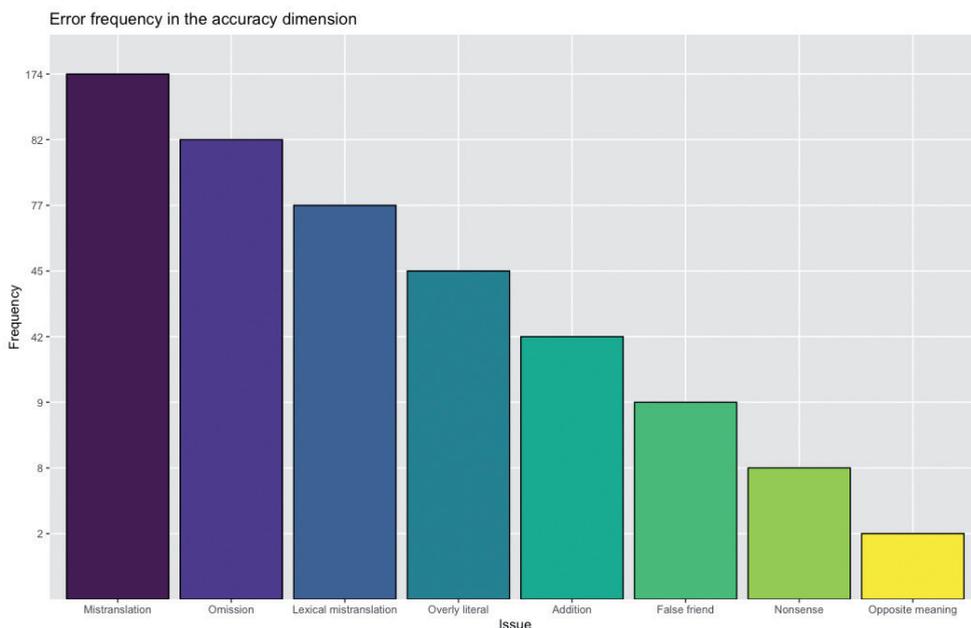

Figure 3: Bar plot of the error frequency in the accuracy dimension.

The fact that mistranslation errors are widespread is not surprising. The shift in meaning is among the most common errors discussed in the literature about translation and interpreting (Havnen, 2019). It is worth restating that the participants of the experiment were students. As they are not on the professional market yet, they lack experience and are more likely to make more errors of this nature.

*Kiara G*ɪ*ᴀɴᴄᴏʟᴀ y Charlène M*ᴇʏᴇʀs
Errors in Specialized Translation Training: A Corpus-
Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - ᴄᴄ ʙʏ-ɴᴄ-sᴀ

101

The omnipresence of omissions in TTs does not fit with the theory that omissions are a limited phenomenon in SiT. Indeed, according to the findings of Agrifoglio (2004) and Gile (1995), omissions are rather rare in SiT, since the ST is always available throughout the exercise. This advantage should therefore prevent any sight translator from omitting segments.

Figure 4 portrays the frequency of terminology and omission errors in each TT. For instance, TT2 and TT9 contain a great percentage of omissions, but one of the lowest numbers of terminology errors. On the other hand, TT15 and TT17 hold the record of incorrect terminology errors but one of the lowest rates of omissions.
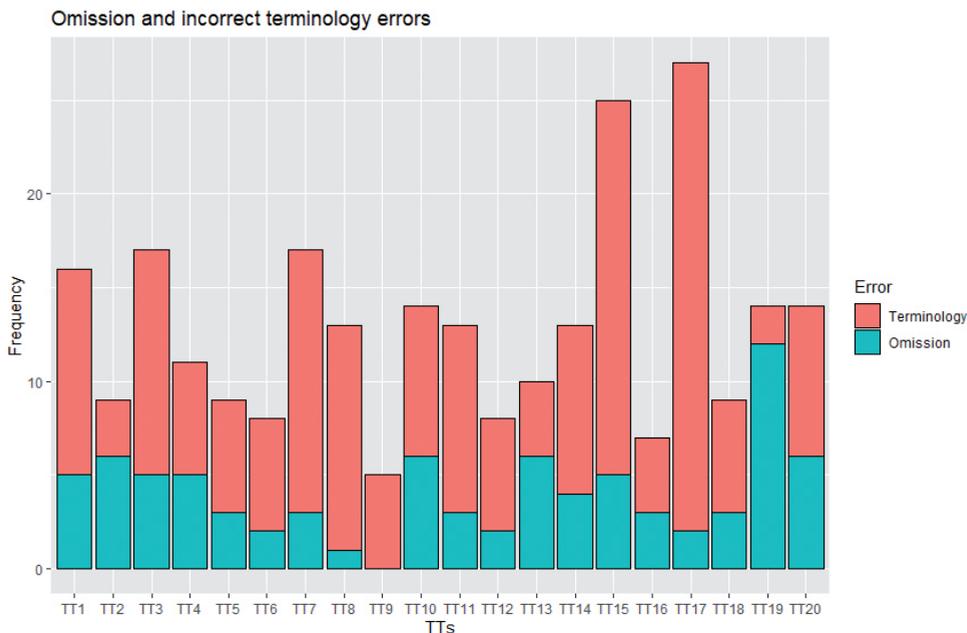

Figure 4: Bar plot of omissions and incorrect terminology errors.

Since there was sometimes an imbalance between the proportion of omission and terminology errors in some TTs, we wondered whether there was a link between omission and terminology errors. In fact, we wondered whether omissions were occasionally employed as a tactic to bypass words or segments considered difficult (since students did not have access to dictionaries and were not familiar with the subject prior to the experiment), especially when confronted with the presence of specialized terms. Gile (2009) argues that omissions are sometimes a deliberate choice in interpreting to avoid words or segments that are considered difficult. In order to determine whether the high number of omissions was partly due to the presence of specialized terms in the text, we classified the omissions into two subcategories: «omission of specialized terms»

Kiara GIANCOLA y Charlène MEYERS
Errors in Specialized Translation Training: A Corpus-
Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

102

and «other» (all the other omissions). We found that 63% of the omissions concerned specialized terms. As previously stated, we expected students to struggle with some specialized terms, since we did not provide any information on the article. As a result, students probably omitted many specialized terms on purpose.

As a way of illustration, Table 1 showcases an example of an omission of a specialized term. The student did not translate «quantum» and only kept the word «fluctuations» by translating it literally, with a generalization in the translation of the word «fabric», which literally means «cloth». The change from a metaphorical but physical aspect of the ST is translated via a more general and open term «creation». In fact, the French translation of this specialized concept is *fluctuations quantiques* (Fliess, 2007).

EN	FR
Just instants after the big bang, **quantum** fluctuations in the fabric of space-time led to regions [. . .]	Quelques instants seulement après le Big Bang les fluctuations **<o>** dans la création de l'espace-temps ont mené à la création de régions [...]

Table 1: Example of the omission of a specialized term in TT1.

However, the results on the omission of specialized terms cannot justify alone why students make so many omissions despite having the text available throughout the performance. Stress is a possible explanation for the omission of segments that were not considered specialized terms. Baghi and Khoshsaligheh's results about stress in SiT (2019) emphasize the fact that students, who have to sight translate on the spot without resorting to resources, are particularly anxious, and «time pressure allows little time for processing all the available translation option» (p. 251). As a result, students, because of stress and time pressure, might be more likely to overlook parts of segments.

Figure 4 also portrays the number of incorrect terminology errors found in each TT. The plot shows that there is strong evidence that errors falling under this issue type are ubiquitous in general. Again, the nature of the ST that encompasses concepts about astrophysics and the fact that students did not receive information before the exercise justify the high number of this type of error. Even if we did not check whether some had background knowledge on the subject, we can assume that most of them in this study tended to struggle when faced with the translation of these specialized terms.

Moreover, it is not surprising to find that lexical mistranslations (and to a lesser extent, false friend errors) are present in many TTs. Agrifoglio (2004) has noted this problem in SiT, arguing that visual interference plays a role in it. This finding points towards the argument that visual interference may indeed play an important role in SiT. Indeed, as the written text is available throughout the process, words in the SL that are similar, if not the same, as potential words in the TL might be translated literally due to their proximity to the TL or their frequent use. When sight translators spot words that could look like a word in the TL or are often translated the same way, they are tempted

to choose this option, rather than looking for a better-suited translation. For instance, in Table 2, the student opted for a word that is one of the correct translations of «rate,» but it is not correct. Although the term «rate» can be translated into French by «taux,» it does not correspond to the correct translation in this context. Indeed, «rate» can both mean «frequency or number of instances of a phenomenon in an area in a given time period» and «the speed at which something occurs or moves» (Druide informatique, 2021, «rate»). In this context, it is the second definition that has to be applied, and thus the translation possibilities are limited to *vitesse* or *rythme*. This issue type differs from mistranslation because students do use one of the potential correct translations of the word from the SL. However, they pick the wrong one for the context.

EN	FR
Pisani and colleagues considered three scenarios, all of which can explain the observed rate of expansion today.	Pisani et ses collègues ont envisagé trois scénarios, chacun expliquant le **taux d'expansion <L>** de l'univers que l'on observe actuellement.

Table 2: Example of a lexical mistranslation error in TT16.

Moreover, the presence of overly literal sentences could be another sign of the visual interference that impedes sight translators from performing a natural and smooth translation, since they tend to be influenced by the SL structures, in this case English structures. For example, in Table 3, the student translated «unavoidable part of» literally. However, «partie inévitable» in this case seems too literal and rather unnatural. The student should have taken a step back by reformulating «unavoidable part». There is a more idiomatic translation in French that could resolve the issue: *«les vides font inévitablement partie intégrante de la répartition de la matière dans l'univers»*.

EN	FR
«Voids are just an unavoidable part of the distribution of matter in the universe,» […]	«Il s'agit juste de **parties inévitables ** de la répartition de matière dans l'univers» […]

Table 3: Example of an overly literal error in TT14.

As for the other issue types, the low prevalence of addition errors does not lead to a particular interpretation, except for a compensation strategy. However, the low percentage of opposite meaning and nonsense errors shows that students still managed to sight translate without introducing major errors that would highly compromise the output.

6.3. *Error frequency in the form division*

Several findings were highlighted based on the observations of the proportions of the errors falling under the issue types belonging to the grammar and style dimensions. Interestingly, as shown in Figure 5, the most identified issue in the grammar dimension is tense/mood/aspect (errors of conjugation), followed by function words (incorrectly used prepositions, helping verbs, etc.), word order, and agreement (subject/verb, etc.). Cohesion errors (missing parts of a sentence or text to form an understandable whole) and part of speech errors (the incorrect syntactic function of a word, e.g., use of an adjective instead of an adverb) were a minority.

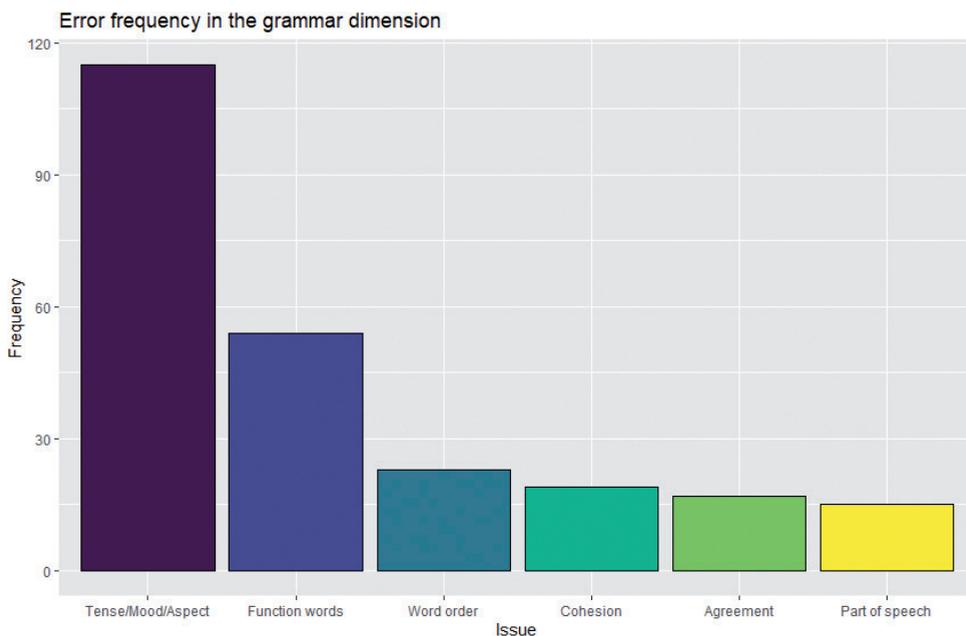**Error frequency in the grammar dimension**

Figure 5: Bar plot of the error frequency in the grammar dimension.

The omnipresence of tense/mood/aspect errors is worth being analyzed. Students made many errors of conjugation. The errors were likely made due to a lack of under-standing of the ST and a lack of distinction between descriptions and actions. As the text is about a study of voids in its early stages, many of the aspects that are discussed have not been proven yet, and the observations are based on hypotheses, as previous-ly stated. However, most of the students struggled with the notion of hypotheses and translated the text into the past tense as if it had happened or as if the hypotheses were facts. In Table 2, there is a clear example of a student translating the hypothesis into a fact because the verb in the ST is in the simple past. However, although the hypotheses

*Kiara G*ɪᴀɴᴄᴏʟᴀ *y Charlène M*ᴇʏᴇʀꜱ
Errors in Specialized Translation Training: A Corpus-
Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - ᴄᴄ ʙʏ-ɴᴄ-ꜱᴀ

105

(scenarios) are put forward in the simple past in English, it is not possible to translate it into a simple past verb in French (with a verb conjugated with the *passé composé* [the French perfect tense]) because it conveys the idea that the scenario really happened. However, the author in the ST states that it has not been proven which scenario really occurred yet.

EN	FR
Pisani and colleagues considered three scenarios, all of which can explain the observed rate of expansion today [...] **The second caused** the expansion to accelerate later but faster than the cosmological constant would have [...]	[...] le deuxième scénario **a provoqué <tma>** un rythme d'expansion plus tardif mais plus rapide que dans le cas de la constante cosmologique [...]

Table 4: Example of a tense/mood/aspect error in TT11.

Additionally, some students mistook the use of the *passé composé* and the *imparfait* [the French imperfect tense], by employing the *imparfait* for concrete actions instead of the *passé composé*, and by using the *passé composé* for descriptions instead of the *imparfait*. Overall, within the category of tense-mood-aspect, tense was particularly a problem to sight-translate properly.

Word order errors are not surprising in EN-FR translation, as the order of words in a sentence in English can differ from that in French. For instance, adjectives often precede the noun in English, whereas adjectives are often positioned after the noun in French. Even when already aware and trained, sight translators can still be tempted to follow the same word order of a given sentence in English. Sometimes, this can lead to an error because the word order in French for that given sentence, or a part of it, should be adapted. As Shreve et al. (2011) argue, the sentence structures proper to a given language make it more difficult for sight translators to rearrange the ideas correctly in another language, especially between English and languages such as Spanish or French. In the corpus, some students fell into that trap of word order errors, as Figure 5 showcases it.

Previous research carried out on SiT errors (Agrifoglio, 2004) found that agreement errors were the main issue of the grammar category. Her findings are supported by the fact that participants seemed «to have suffered from coordination problems and short-term memory failures,» which led them to «[lose] the referent and [forget] the gender, number, and person» (Agrifoglio, 2004, pp. 52-53). Interestingly, our analysis revealed that agreement errors occurred less frequently than errors related to tense/mood/aspect, function words, and word order in the corpus. Despite their lower frequency, however, the occurrence of agreement errors among some students could still be attributed to the factors identified by Agrifoglio.

Although it is not clear why such errors arose in large numbers in the corpus, function word errors were numerous. For instance, there were many preposition errors,

such as shown in Table 5. We can see that the preposition [de] (of) is used incorrectly. The preposition that goes with the adjective «inhérent» is [à] (to), as the word «inhérent» means something that is closely linked to something else (Druide informatique, 2021). This issue type differs from cohesion because it focuses on the wrong use of function words, while the cohesion category focuses on the absence of connecting words.

EN	FR
[. . .] Astronomers realised that the expansion of the universe was accelerating and attributed this to the inherent «dark energy» of space-time.	[…] les astronomes […] ont attribué ce phénomène à l'énergie noire inhérente **de <fw>** l'espace-temps.

Table 5: Example of a function word error in TT13.

As for the part of speech errors, such as using an adjective instead of an adverb, and cohesion errors, these issues were rare. The only cohesion error we found is the one portrayed in Table 6. The dash in the ST is important. It links the first part of the sentence with the second one. It is used to give further information on voids. The text is translated aloud, so it is not possible to put the dash in French given that it is a written sign. However, there is a missing element to link both parts in the TT. One way to explicitly express the dash in French is to say «c'est-à-dire» [that is to say] when the dash is used to provide details about the information given before it. This is therefore a cohesion error, since there is a missing element to connect the sentence as an understandable whole.

EN	FR
The less dense regions became **voids—regions** of space-time almost empty of matter, which can stretch from 30 million to 150 million light years across.	Les régions moins denses sont devenues **des vides des régions <ch>** de l'espace-temps quasiment vides de matière qui peuvent s'étirer sur 30 à 150 millions d'années-lumière.

Table 6: Example of a cohesion error in TT9.

No specific interpretation for the low prevalence of part of speech and cohesion errors can be given apart from the fact that, generally, students did manage to link the sentences together in a proper way, without using an adjective instead of an adverb, for instance, and remained overall faithful to the cohesion of the ST.

As a whole, the most identified issues from the style dimension (see Figure 1) are both lexical clumsiness and idiomatic clumsiness. Register is the least identified issue from that dimension, although the difference with the other two is insignificant. These results are in line with the findings of Agrifoglio (2004), Gile (1985a), and Falbo (1998) in SiT or in interpreting. According to Gile (1985a), students are prone to making these errors due to the particular constraints associated with interpreting, and more generally

Kiara GIANCOLA y Charlène MEYERS
Errors in Specialized Translation Training: A Corpus-Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

107

with oral translation. Indeed, students do not have time to pause and reflect on what they say, and they are thus more likely to mix up two expressions or invent a word without realizing it. However, it should be noted that the frequency of both lexical and idiomatic awkwardness errors is low among the TTs of the corpus.

In the context of our SiT study, there was a possibility that the students' spoken translations might contain errors related to register. There is a risk that the translator may unintentionally use a more informal register when delivering the oral translation because oral language tends to be less formal than written language in general (Alic, 2020). This is why it is important for the translator to be aware of the appropriate level of formality and register for the intended audience and context of the oral translation to ensure accuracy and consistency in the translation. Regardless of the genre of the original text, if the register in the written ST is formal, it should be kept formal in the oral translation as well. This means avoiding the introduction of any informal elements of oral language that could compromise the register of the ST.

Nevertheless, the tendency to resort to a less formal register was very limited in the corpus, with only one highlighted error. In Table 5, we can see that the student introduced an error of register. «Ça» is the less formal version of the demonstrative pronoun «cela» [*this/that*] (Druide informatique, 2021) and is mainly used in spoken contexts or in informal written language, such as text messaging. Since the text is published in the journal called *New Scientist* (journalistic genre), the register both in English and French should therefore be kept rather formal.

EN	FR
[. . .] Astronomers realised that the expansion of the universe was accelerating and attributed **this** to the inherent «dark energy» of space-time.	[...] les astronomes ont réalisé que l'ex l'expansion de l'univers s'accélérait et et ils ont attr et ils ont attribué **ça <R>** à la matière noire [...]

Table 7: Example of a register error in TT17.

6.4. *Cognitive load*

As stated in the methodology, previous studies carried out by Dragsted and Hansen (2009) and Shreve et al. (2011) have postulated that speech disfluencies could influence the quality of SiT, without statistically proving this relationship. However, the present paper provides strong evidence for such a link, by proving the statistically significant association between the number of errors and the number of speech disfluencies in each of the four categories ($p < 0.05$). These findings imply that unfilled pauses, filled pauses, repetitions, and repairs all tend to affect, to different degrees, the quality of SiT.

For the speech disfluency analysis, we wanted to determine whether there was an impact on the number of errors when students corrected themselves aloud and

Kiara Giancola y Charlène Meyers
Errors in Specialized Translation Training: A Corpus-Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

108

thus introduced repairs. A chi-squared test was performed on errors and repairs and resulted in $\chi^2(19) = 40.456$, $p = 0.00285$. The p-value shows that there is a significant association between errors and repairs. Additionally, a correlation coefficient was computed to determine the strength of correlation between the two values and resulted in $r = 0.5359512$, which means that there is indeed a moderate correlation. In other words, there is a moderate tendency for the total number of repairs to rise when the number of errors increases, and vice versa. Overall, it is fair to argue that the correlation is far from being a coincidence.

Figure 6 showcases the scatter plot of errors and repairs in each TT. The regression line moves upward, which implies that there is a positive correlation between the values. TT17, which appears to be an outlier in the scatter plot, has the highest number of both errors and repairs, while TT2 and TT8 are the TTs with the lowest number of errors and repairs.
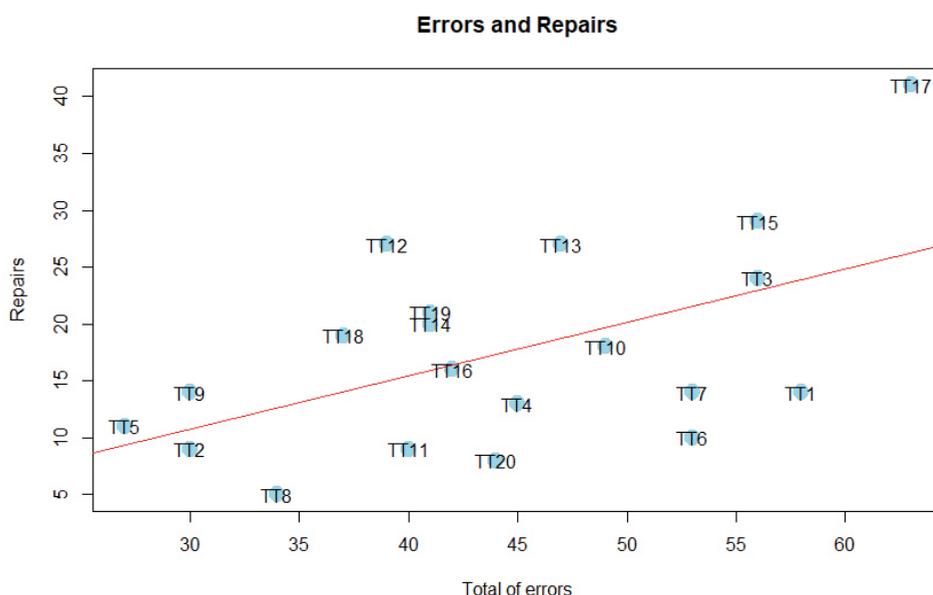

Figure 6: Scatter plot of repairs vs. total of errors in each TT.

The results on the repairs show that students, upon realizing they made errors or mistakes while sight translating, often introduce oral corrections to their oral translation. However, because SiT is processed on the spot, sight translators may avoid explicitly stating that they are correcting themselves and forget to say, for instance, «excuse me, let me correct this sentence». We took this aspect into account for our typology by splitting the repair issue type into two sub-types, announced and unannounced repairs.

109

In the corpus, the observed proportion between unannounced and announced repairs led to the conclusion that, in 95% of the cases, students did not explicitly announce they were about to correct themselves. The speedy delivery that is expected from sight translators when performing the exercise is one of the potential factors of such a phenomenon, alongside stress and cognitive overload.

As SiT is processed, sight translators are expected to translate the text on the spot without stopping, which might prompt them to avoid stating explicitly that they correct themselves. This is probably because, when a sight translator announces a self-correction during translation, it can interfere with their thought process and disrupt the flow of the translation. The focus may shift to the announcement of the correction itself, rather than the correction of said translation, which can disturb the train of thought and mislead the sight translator. Indeed, this type of announced correction is an interruption to the translation process, since it introduces new information that is not included in the ST, such as phrases like «sorry, let me rephrase that».

Additionally, as said earlier, stress and cognitive overload have been proved to play a negative role in the quality of SiT (Baghi & Khoshsaligheh, 2019). Indeed, Baghi & Khoshsaligheh (2019) found that translation students feel anxious when performing SiT and they thus tend to make errors and introduce speech disfluencies when their coordination effort is saturated. Another potential explanation for the predominance of unannounced repairs, which could give further credit to Gile's (1985b, 1997/2002) and Baghi and Khoshsaligheh's (2019) observations, is that when students realize that they might have made an error and decide to introduce a repair, their cognitive load and coordination effort are already saturated with information. They are thus left with no time or space to think about explicitly announcing their revision. This interpretation could partly justify the high number of unannounced repairs in the corpus and their direct impact on the overall quality of the output.

In the example in Table 3, there is a repair between «*être*» and «*expliquer*». The student probably realized that the verb «*être*» [to be] was not the correct choice to translate the ST segment. He therefore said «*tous pouvant expliquer*,» which, in this case, meets the ST meaning. However, upon realizing he had made a mistake, the student did not explicitly say that he was correcting himself. This unannounced repair can create confusion and hinder the understanding of the TT. However, sometimes, some students did announce their correction by stating «excusez-moi, je reprends» [excuse me, let me start again], for instance.

EN	FR
Pisani and colleagues considered three scenarios, all of which can explain the observed rate of expansion [. . .]	Pisani et ses collègues ont considéré trois scénarios, **tous pouvant être pouvant expliquer [h]** le taux d'expansion observé

Table 8: Example of an unannounced repair in TT6.

Kiara Giancola y Charlène Meyers
Errors in Specialized Translation Training: A Corpus-
Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

110

Overall, these results show that the presence of speech disfluencies is not random and does play an active part in the delivery of SiT. By proving their direct negative implication on the overall quality of SiT, this study shows that such phenomena bear witness to the important cognitive load tied to SiT, as pointed out by Shreve et al. (2011). In fact, the speech disfluencies the researchers identified in their corpus «allowed [them] to associate the disfluencies with probable sources of processing difficulty» (Shreve et al., 2011, p. 108). Additionally, they argue that speech disfluencies are a direct reflection of the cognitive processes and efforts at play in SiT and are «associated with specific production problems resulting from the translation task» (Shreve et al., 2011, p. 117).

In light of these observations and the results of this study, it is fair to assume that speech disfluencies are an embodiment of the cognitive efforts associated with SiT. More importantly, the high rate of speech disfluencies in this corpus proves that students struggled to process the specialized text when sight translating it, certainly because they could only rely on their knowledge and ability to understand a specialized text written in English without the use of dictionaries. SiT already implies a cognitive effort, but the added level of difficulty due to the text has had a direct impact on the number of speech disfluencies as well. With these findings, it is fair to argue that a proper preparation is more than needed to help students cope with any type of obstacles when sight translating specialized texts and therefore improve both their training and performance.

7. CONCLUSION

This study has contributed to broadening the scope of SiT in the sphere of translator training. Indeed, the analysis of SiT based on a corpus of translation students made it possible to create a typology specifically designed for SiT and yield results about the error patterns and trends in the context of specialized translation training, which has scarcely been done, if done at all.

While the genre of the text and the size of the corpus limit the generalizability of the results, these still provide new insight into errors in SiT in the context of specialized translation training. Such findings are particularly relevant because they show that the results in the corpus were not always in line with the findings identified by researchers such as Agrifoglio (2004) and Gile (e.g., 1985a; 1995; 2009), who have investigated SiT errors in the realm of interpreting. It is therefore fair to argue that SiT within translator training should be explored separately, because it would provide an independent framework that would address the needs and issues specifically arising among translation students and translators when performing this exercise, and consequently adapt training accordingly, or so we suggest.

Finally, the establishment of the significant impact of speech disfluencies on the overall quality of SiT is by far the most important contribution to this study. Although

Kiara GIANCOLA y Charlène MEYERS
Errors in Specialized Translation Training: A Corpus-
Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

111

the impact of speech disfluencies on SiT quality had already been postulated by Dragsted and Hansen (2009) and Shreve et al. (2011), the new findings are believed to be the first to confirm the statistically significant correlation between errors and speech disfluencies in SiT. In light of the profound implication of speech disfluencies in the overall performance, these should be addressed in training, with the establishment of exercises and workshops to help curb such phenomena and improve the overall quality of the output. Moreover, the high number of speech disfluencies bears witness to the difficulty of the text and confirm the importance of teaching specialized translation and providing students with the necessary tools to progress and eventually master SiT with other types of specialized texts, such as legal texts and research articles.

Based on these findings and previous research, SiT teachers should consider emphasizing on the understanding of the ST to help students avoid a high number of content errors, especially mistranslations, omissions, and incorrect terminology errors. It can be hypothesized that the misunderstanding of the subject has had an impact on the students' performance. This proves that specialized translation training remains central to provide students with a more comprehensive approach to translation, and thus help them improve their skills with regular training.

To develop a full picture of SiT in translator training, additional studies that explore the ins and outs of this tool and its purpose in the translator's toolbox will be needed. Further studies on the error analysis in SiT are also required to explore patterns depending on the genre of the ST, especially in the different fields of specialized translation. It would be relevant to determine which type of text is considered most or least efficient, but also to compare the error trends and patterns between all the genres.

The corpus-driven approach employed for the creation of a SiT error typology awaits more rigorous study. The present work represents a first attempt to address errors in SiT with the help of a typology specifically put forward for the practice of SiT. The investigation of errors in other SiT corpora would therefore be an opportunity to further develop the tool.

Although the generality of the current results must be established by future research, the present study has provided a first in-depth overview of error patterns in SiT in the context of translator training following previous literature. Despite the limitations, not only have the results yielded valuable conclusions on the error typology and error trends and patterns in SiT, but the study has also enhanced our understanding of the relationship between errors and speech disfluencies in SiT settings. Such analyses contribute to a growing body of evidence suggesting that speech disfluencies impact the quality of the output. All aspects considered, this work paves the way for more promising studies on the role of SiT in the field of translator training and specialized translation.

Kiara Giancola y Charlène Meyers
Errors in Specialized Translation Training: A Corpus-
Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

112

REFERENCES

Agrifoglio, M. (2004). Sight translation and interpreting: A comparative analysis of constraints and failures. *Interpreting. International Journal of Research and Practice in Interpreting*, *6*(1), 43-67. https://doi.org/10.1075/intp.6.1.05agr

Alic, L. (2020). La traduction à vue: quelles compétences? In S. Vaupot, A. Mezeg, G. Perko, M. Schlamberger Brezar, M. Zupančič (Eds.), *Contacts linguistiques, littéraires, culturels: cent ans d'études du français à l'université de Ljubljana*, Les presses universitaires de l'Université de Ljubljana, pp. 218-229.

Ananthaswamy, A. (2015). Looking into the voids could help explain dark energy. *New Scientist*. https://www.newscientist.com/article/mg22630164-600-looking-into-the-voids-could-help-explain-dark-energy/

Anthony, L. (2020). *AntConc* (3.5.9) [Computer software]. http://www.laurenceanthony.net/software/antconc/

Anthony, L. (2022). *AntConc* (4.0.5) [Computer software]. http://www.laurenceanthony.net/software/antconc/

Baghi, H. and Khoshsaligheh, M. (2019). Stress in written and sight translation in training setting. *Hikma*, 18(2), 237-255.

Baxter, R. N. (2017). Exploring the effects of computerised sight translation on written translation speed and quality. *Perspectives*, 25(4), 622-639. https://doi.org/10.1080/0907676X.2016.1241287

Calvo, E. (2011). Translation and/or translator skills as organising principles for curriculum development practice. *The Journal of Specialised Translation*, *16*, 5-25.

Dizdar, D. (2014). Instrumental Thinking in Translation Studies. *International Journal on Translation Studies*, 26(2), 206-223.

Dragsted, B., & Hansen, I. G. (2009). Exploring Translation and Interpreting Hybrids. The Case of Sight Translation. *Meta*, *54*(3), 588-604. https://doi.org/10.7202/038317ar

Druide informatique. (2021). Antidote 10 (6.3) [Computer software]. https://www.antidote.info

Ericsson, K. A. (1996). The acquisition of expert performance: An introduction to some of the issues. In *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games.* (pp. 1-50). Lawrence Erlbaum Associates, Inc.

Ericsson, K. A., Krampe, R. Th., & Tesch-Romer, C. (1993). The Role of Deliberate Practice in the Acquisition of Expert Performance. *Psychological Review*, *100*(3), 363-406.

European Commission. (2022a). *European Master's in Translation Competence Framework 2022*. Retrieved November 21, 2022, from https://ec.europa.eu/info/sites/default/files/about_the_european_commission/service_standards_and_principles/documents/emt_competence_fwk_2022_en.pdf

European Commission. (2022b). *European Master's in Translation (EMT) explained*. Retrieved November 21, 2022, from https://ec.europa.eu/info/resources-partners/european-masters-translation-emt/european-masters-translation-emt-explained_en

European Commission. (2022c). List of EMT members 2019-2024. Directorate-General for Translation. https://ec.europa.eu/info/resources-partners/european-masters-translation-emt/list-emt-members-2019-2024_en

Falbo, C. (1998). Analyse des erreurs en interprétation simultanée. *The Interpreters' Newsletter*, *8*, 107-120.

Kiara Giancola y Charlène Meyers
Errors in Specialized Translation Training: A Corpus-Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

113

FLIESS, M. (2007). Probabilités et fluctuations quantiques. Comptes rendus de l'Académie des sciences. Série I, Mathématique, *Elsevier*, 344 (10), pp.663-668. https://hal.inria.fr/inria-00141719/document

GASPARI, F., ALMAGHOUT, H., & DOHERTY, S. (2015). A survey of machine translation competences: Insights for translation technology educators and practitioners. *Perspectives*, *23*(3), 333-358. https://doi.org/10.1080/0907676X.2014.979842

GIANCOLA, K. (2022). *Exploring sight translation as a pedagogical tool in translator training programs across Europe* [Conference]. European Society for Translation Studies Congress 2022 – Advancing Translation Studies.

GILE, D. (1983). Des difficultés de langue en interprétation simultanée. *Traduire*, 117, 2-8.

GILE, D. (1984). Les noms propres en interpretation simultanée. *Multilingua*, 3, 79-85.

GILE, D. (1985a). La sensibilité aux écarts de langue et la sélection d'informateurs dans l'analyse d'erreurs: une expérience. *The Incorporated Linguist*, 24/1, 29-32.

GILE, D. (1985b). Le modèle d'efforts et l'équilibre d'interprétation en interprétation simultanée. *Meta*, 30(1), 44-48. https://doi.org/10.7202/002893ar

GILE, D. (1989). *La communication interlinguistique en réunion multilingue: les différences de la transmission informationelle en interprétation simultanée.* [Unpublished doctoral dissertation]. Université Paris III.

GILE, D. (1990). L'évaluation de la qualité de l'interprétation par les délégués: une étude de cas. *The Interpreters' Newsletter*, 3, 66-71.

GILE, D. (1994). The process-oriented approach in translation training. In C. Dollerup & A. Lindegaard (Eds.), *Teaching Translation and Interpreting 2: Insights, aims, visions* (pp. 107-112). John Benjamins.

GILE, D. (1995). *Regards sur la recherche en interprétation de conférence*. Presses universitaires de Lille.

GILE, D. (1997/2002). Conference interpreting as a cognitive management problem. In F. Pöchhacker and M. Shlesinger (Eds.), *The Interpreting Studies Reader*, Routledge, 163- 176. [Original work published in 1997].

GILE, D. (2004). Integrated Problem and Decision Reporting as a Translator Training Tool. *The Journal of Specialised Translation*, *2*, 2-20.

GILE, D. (2009). *Basic Concepts and Models for Interpreter and Translator Training: Revised edition.* John Benjamins. https://doi.org/10.1075/btl.8

HAVNEN, R. (2019). Multimodal and interactional aspects of sight translation: a critical review. *FITISPos International Journal*, 6, 91-106.

KIRALY, D. (2013). Towards a View of Translator Competence as an Emergent Phenomenon: Thinking Outside the Box(es) in Translator Education. In D. Kiraly, S. Hansen-Schirra, & K. Maksymski (Eds.), *New Prospects and Perspectives for Educating Language Mediators* (pp. 197-224). Narr Francke Attempto.

KOLB, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice-Hall.

LI, X. (2014). Sight translation as a topic in interpreting research: Progress, problems, and prospects. *Across Languages and Cultures*, 15(1), 67-89. https://akjournals.com/view/journals/084/15/1/article-p67.xml

LOMMEL, A., BURCHARDT, A., GÖRÖG, A., USZKOREIT, H., and MELBY, A. K. (Eds.). (2015, December 30). *Multidimensional Quality Metrics (MQM) Issue types* [version 1.0]. German Research

Kiara GIANCOLA y Charlène MEYERS
Errors in Specialized Translation Training: A Corpus-Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - CC BY-NC-SA

114

Center for Artificial Intelligence and QTLaunchPad. Retrieved March 27, 2022, from https://www.qt21.eu/mqm-definition/issues-list-2015-12-30.html

Lommel, A., Uszkoreit, H. and Burchardt, A. (2014). Multidimensional Quality Metrics (MQM): A Framework for Declaring and Describing Translation Quality Metrics. *Tradumàtica: Tecnologies de La Traducció*, 12, 455-463. https://doi.org/10.5565/rev/tradumatica.77

Martín Martín, J.M. (2010). Sobre la evaluación de traducciones en el ámbito académico? *Revista española de lingüística aplicada*, 23, 229-245.

Massey, G. (2017). Translation Competence Development and Process-Oriented Pedagogy. In J. W. Schwieter & A. Ferreira (Eds.), *The Handbook of Translation and Cognition* (1st ed., pp. 496-518). Wiley. https://doi.org/10.1002/9781119241485.ch27

Massey, G. (2019). Learning to Learn, Teach and Develop. Co-emergent Perspectives on Translator and Language-mediator Education. *TRAlinea*, *Special Issue: New Insights into Translator Training*, 1-10.

Meyers, C. (2021). *Qualitative and quantitative analyses of metaphors in the LSP of astrophysics: A comparative linguistic study in three registers of English and a translation study in English and French followed by pedagogical considerations* [Thèse de doctorat en langues, lettres et traductologie]. Université de Mons.

Meyers, C. (2022). *Transcription of 20 sight translations (EN >FR) by translation students coded with voiced pauses and silent pauses* [Data set]. Social Sciences and Digital Humanities Archive – SODHA. https://doi.org/10.34934/DVN/KHQE0P

Mellinger, C. D. (2017). Equal access to the courts in translation: A corpus-driven study on translation shifts in waivers of counsel. *Perspectives*, 25(2), 308-322. https://doi.org/10.1080/0907676X.2016.1248985

Prak-Derrington, E. (2011). *Traduire ou ne pas traduire les répétitions.* Nouveaux Cahiers d'Allemand: Revue de linguistique et de didactique, Association des Nouveaux Cahiers d'Allemand, 29 (3), pp. 293-305.

Riccardi, A. (1999). Interpretazione simultanea: strategie generali e specifiche. In Ed. par C. Falbo, M. Russo and F. Straniero Sergio, *Interpretazione simultanea e consecutiva: problemi teorici e metodologie didattiche.* Hoepli, pp. 161-174.

Setton, R., & Motta, M. (2007). Syntacrobatics: Quality and reformulation in simultaneous-with-text. *Interpreting. International Journal of Research and Practice in Interpreting*, 9(2), 199-230. https://doi.org/10.1075/intp.9.2.04set

Shreve, G. (2006). The Deliberate Practice: Translation and Expertise. *Journal of Translation Studies*, 9(1), 27-42.

Shreve, G. M., Lacruz, I. and Angelone, E. (2010). Cognitive Effort, Syntactic disruption, and Visual Interference in a Sight Translation Task. In G. M. Shreve and E. Angelone (Eds.), *Translation and Cognition*, John Benjamins Publishing Company, 63-84.

Shreve, G. M., Lacruz, I. and Angelone, E. (2011). Sight Translation and Speech disfluency: Performance Analysis as a Window to Cognitive Translation Processes. In B. Englund Dimitrova, C. Alvstad, A. Hild and E. Tiselius (Eds.), *Methods and Strategies of Process Research: Integrative Approaches in Translation Studies*, John Benjamins Publishing Company, 93-121.

Stenzl, C. (1983). *Simultaneous interpretation: ground work towards a comprehensive model*. [Unpublished Master's thesis]. University of London.

Whyatt, B. (2022). Training and Pedagogy. In A. Ferreira & J. Schwieter (Eds.), *Introduction to Translation and Interpreting Studies* (pp. 244-269). Wiley Blackwell. https://www.wiley.com/en-us/Introduction+to+Translation+and+Interpreting+Studies-p-9781119685326

Kiara Giancola y Charlène Meyers
Errors in Specialized Translation Training: A Corpus-Based Study on the Sight Translation of a Popular…

CLINA
vol. 9-2, 2023, 85-115
eISSN: 2444-1961
Ediciones Universidad de Salamanca - cc by-nc-sa

115