The Translation of Adolescence Language by means of *Apertium, Systran* and *Google Translate*

La traducción del lenguaje juvenil a través de *Apertium, Systran* y *Google Translate*

MARÍA CIRA NAPOLETANO
ANDRÉS CANGA ALONSO
UNIVERSIDAD DE LA RIOJA

The present paper explores the translation of adolescents' speech from English into Spanish using three free online translation tools (*Apertium, Systran* and *Google Translate*), and two corpora based on oral communication (the *Bergen Corpus of London Teenage Language* and the *Linguistic Innovators Corpus*). Additionally, errors were classified in terms of accuracy and fluency, and a revised version after post-editing is provided in order to overcome these mistranslations. Our findings show that the errors these Machine Translation applications produce have to do with the translation of cultural aspects, abbreviations, proper names of cities and people, as well as the loss of coherence of some extracts due to the character limit imposed by some of the tools used. Finally, emphasis is placed on the need to open new lines of research considering proverbs and idioms with a wider range of data.

Keywords: *machine translation; adolescents' speech; COLT; LIC; mistranslation*

El presente artículo explora la traducción del habla de los adolescentes del inglés al español utilizando tres herramientas gratuitas de traducción en línea (*Apertium, Systran* y *Google Translate*) y dos corpus basados en comunicación oral (el *Bergen Corpus of London Teenage Language* y el *Linguistic Innovators Corpus*). Asimismo, los errores se clasifican en términos de precisión y fluidez, proporcionando una versión corregida de los mismos tras llevar a cabo un proceso de posedición con el objetivo de solucionar los problemas de traducción detectados. Los resultados muestran que los errores que estas aplicaciones generan tienen que ver con la traducción de aspectos culturales, abreviaturas, nombres propios de ciudades y personas, así como con la falta de coherencia debido al límite de caracteres impuesto por las herramientas utilizadas. Por último, se hace hincapié en la necesidad de abrir nuevas líneas de investigación considerando refranes y modismos con una muestra más amplia del corpus de los adolescentes analizados.

Palabras claves: *traducción automática; lenguaje juvenil; COLT; LIC; traducción errónea*

1. INTRODUCTION

Machine translation (MT) tools have evolved remarkably in recent years, improving the accuracy of the texts produced using different MT instruments. According to Hutchins and Somers (1992: 1), the idea that a machine could translate human language has always been considered an unreachable dream for humans. However, it is only in the twentieth century that more detailed research in this field began. Nonetheless, one limitation that MT presents since its beginning is the difficulty to manage figurative and creative aspects of language adequately, as is the case of adolescents' slang. In fact, as the authors stated, "the major obstacles to translating by computer are, as they have always been, not computational but linguistic. They are the problems of lexical ambiguity, syntactic complexity, vocabulary differences between languages, of elliptical and 'ungrammatical' constructions" (Hutchins & Somers, 1992: 2).

Several studies have focused on teenagers' language applied to different fields but, to our knowledge, there is a scarcity of research in regard with the way adolescents' language has been translated into Spanish by online translation tools. The main objective of this article is to conduct an analysis and comparative assessment of three translation tools, namely *Apertium*, *Systran* and *Google Translate*, with a focus on their potential errors in the translation of colloquial texts. This comparison will allow the reader to notice the progression of MT and to identify the specific challenges it continues to face. Hence, we set out to find answers to the following questions: a) what are the most frequent fluency and accuracy errors in each translation tool?; b) specifically, are there any mistranslations that are repeated in the three MT systems? If so, which ones are the most common?; and, c) is post-editing necessary? If so, what kind of post-editing does the translator need to apply?

The paper proceeds as follows. The first section provides a brief overview of the evolution of MT, highlighting the two approaches that have emerged over the years (*Rule-based translation* and *Corpus-based translation*). Section 3 is devoted to presenting the main characteristics of youth language, which are key to understand the difficulties the three MT systems used in the present research had to face. A report of the study conducted with main results found and interpretation of the same follows. This paper concludes by pointing out some lines for further research trying to overcome its main limitations.

2. APPROACHES TO MACHINE TRANSLATION

The well-known term Machine Translation (MT) refers to "computer systems responsible for producing translations from one natural language into another, with or without human assistance" (Hutchins & Somers, 1992: 3).

MT research emerged in the 1950s, marked by the advent of two systems designed to enhance translation quality (Hutchins & Somers, 1992: 148). These were *machine-aided human translation* (MAHT), which involved computer-based linguistic tools for checking spelling, grammar, and style (Hutchins & Somers, 1992: 149), and *human-aided machine translation* (HAMT), which relied on human assistance during translation, employing pre-editing and post-editing techniques (Hutchins & Somers, 1992: 151). Nowadays, with the widespread availability of the Internet, MT has achieved its primary objective of delivering high-quality translations (Brusasco, 2018: 4). Key contributors to this success are *Rule-based translation* (RBT) and *Corpus-based translation* (CBT) systems.

RBT is a system based on syntactic, semantic, and lexical rules (Costa-Jussà, Farrús, Mariño & Fonollosa, 2012: 248). It deals with analysing the sentences of the source text (ST), splitting them up and creating syntactic diagrams to reorder or modify sequences (Costa-Jussà et al., 2012: 249). This system relies on three methods: a) the direct approach, that can be

considered a word-for-word translation (Hutchins & Somers, 1992: 72); b) the Interlingua approach, in which the SL is transformed into an interlingua, that is, an abstract language independent representation; and c) the transfer approach that is based on analysis, transfer, and generation processes (Hutchins & Somers, 1992: 75).

In the last decade of the twentieth century, CBT arose. It is based on "large collections of (usually) translated and aligned documents" (Marshman, 2023: 3). It complies three different approaches: a) *example-based machine translation* (EBMT) that matches chunks of words within a database, identifies the correct translation corresponding to each fragment (Sharma, Diwakar, Singh, Singh, Kadry & Kim, 2023: 6) and, finally, assemblies them (Marshman, 2023: 3) to generate the target text (TT); b) *statistical machine translation* (SMT) that is based on mathematical parameters and probable occurrences, that is to say, a certain word is assigned several possible translations (Sharma et al. 2023: 9) and it is translator who selects them focusing on a "matter of taste" (Sharma et al. 2023: 9); and, c) *neural machine translation* (NMT) in which information processing is inspired by biological human brain networks (Dias Esqueda, 2021: 283). It is considered to produce a TT with a higher quality with respect to the abovementioned systems, as it not only allows the insertion of very long sequences but generates "a less predictable output" (Marshman, 2023: 3).

The two types of systems that have been just described (RBT and CBT) will be used to translate the adolescents' language, being it the main purpose of this paper. However, as will be shown in Section 5, the goal of a system capable of generating translations without the help of the human being is far from being real. Be that as it may, the quality of the output has always been a matter of concern since the first years of the MT, making this arise several strategies for evaluating the final text.

2.1 *MT quality assessment*

Translators had to follow a number of standards in order to ensure a good quality, such as the EN 15038 (Gouadec, 2010: 270), based on the idea that the quality of a translation had to satisfy both the provider and the translator themselves. Following this model, therefore, the translator needs to provide improvements in the ST, if necessary (Gouadec, 2010: 271). In recent years, there has been a change in the application of the standards mentioned by Gouadec (2010) since assessing the quality of a translation can be harder as there is not a single correct translation or a universal evaluation method to follow (Pym, 2020: 445). In this vein, there are two methods to assess the quality of a translation. The first one relies on human judgment, considering factors such as the time taken to read the translated text and the time devoted to post-editing. It categorizes errors into two distinct types: a) accuracy errors, which make the reader have difficulties when reading the TT since the words or expressions have been changed from their original text (Popović, 2020: 259); and b) fluency errors, those that emerge when the translated text presents grammatical errors and are considered not genuine by native speakers (Fernández-Torné & Matamala, 2021: 100). Additionally, the effort invested by human translators in post-editing can be classified into: a) *cognitive*, involving the mental processing behind editing decisions, which considers the choices the translator must make; b) *technical*, related to mechanical text changes; and c) *temporal*, representing the time needed for the task, including both cognitive and technical aspects (Vieira, 2020: 319).

The second possible method to evaluate a translation is the algorithm-based automatic one (Pym, 2020: 441). The advantage of these systems lies in the fact that they are "objective, faster, and allow to minimize human intervention" (Pym, 2020: 442). Therefore, evaluation or assessment of MT systems takes place either with human evaluations or with automatic metrics (Vela-Valido, 2021).

Nevertheless, we will follow the human evaluation method since the present article aims to be an exploratory study of how machines face the translation of the language of English-speaking adolescents into Spanish. We will also make reference to the error taxonomy classification carried out by Webster, Fonteyne, Tezcan, Macken and Daems (2020: 5), according to which fluency errors affect lexicon, grammar, syntax, and spelling while accuracy ones have to do with mistranslation, addition, omission, and untranslated words.

Once we have explored the origins and current trends in MT, we move to analyse the characteristics of teenage language and how it has been translated into different languages.

3. THE LANGUAGE OF ADOLESCENTS

Throughout the twentieth century, the world has faced technological advances which imply a change for language, especially for English that is the language used for global communication. These changes mainly affect the way young people speak, which is characterized by a wide use of slang. Eckert (2003: 114) defines *slang* as "a term used to refer to lexical innovation by delegitimized groups—among them adolescents—implying a qualitative difference from other kinds of lexical innovation".

The tendency towards new types of communication (i.e., text messaging, email, instant messaging, etc.) is fundamentally changing the types of contacts young people have on a daily basis (Tagliamonte, 2016: 2). New terms are learned that will then be extracted from the technological context and taken to everyday life (e.g., *LOL*—laughing out loud—or the abbreviations, characterized by violations of linguistic grammatical rules), which are typical of teenage language.

As for English-speaking teens, their special code includes a variety of interesting and distinctive features (Palacios Martínez, 2011: 1) as they belong to the young population of London. First of all, existing studies emphasize the high frequency that adolescents make of words belonging to vague language, used especially when the name of a person or thing is not remembered. Among the most common expressions, we can find *thingy*, *stuff like that*, and *loads of* (Palacios Martínez, 2011: 115).

Another frequent aspect of youth jargon, there is the high production of negatives. The expression *ain't* for instance, is the most common (Palacios Martínez, 2013: 213) and is used for all verbal forms and person numbers. Among the negative variants are also words such as *innit*, *wunnit* and *dunnit* (or *dunno*, more frequently), which have been introduced into youth discourse due to the presence of a wide variety of ethnic groups in London (Palacios Martínez, 2010: 10). It should also be noted the tendency of young English speakers to produce abbreviations. With respect to this, Palacios Martínez (2021: 6) defines it as a strategy "to reinforce their identity as a group, that is, to support the feeling that they all share the same code that binds them together". Among the best known are *sis, cuz, fam, diff*. It is also striking the recurrent use of vocatives that adolescents produce, such as *folks, guys, dude, mate, man* (Palacios Martínez, 2018: 34), to which the latter is also employed when referring to female speakers (Hall, 2020: 127).

To our knowledge, the linguistic aspects that have been just described have not been translated by a MT system alone. It is the human translator who reviews the output provided by the system, as the language of teenagers is dynamic and creative, and also presents gaps that cannot be solved without human intervention. Hence, human translators need to resort to translation techniques.

These techniques can be classified into literal (or direct) and oblique translation. The literal translation procedures are those that occur when there is an exact structural, lexical, even morphological equivalence between two languages (Molina & Hurtado Albir, 2002: 499).

Among them, we can find *borrowing, calque* and *literal translation.* Oblique translation, on the contrary, occurs when it is impossible to realize a word-for-word translation (Molina & Hurtado Albir, 2002: 499) thus, the translator needs to make use of *transposition, modulation* or *adaptation*. The former consists in the modification of the grammatical category without changing the meaning of the message (Vinay & Darbelnet, 1995: 36). Modulation, on the other hand, focuses on meaning; thus, it is a variation obtained by a change in the point of view (Vinay & Darbelnet, 1995: 36). Finally, adaptation is the technique used to adapt a SL expression or message to the TL culture. The translation of cultural issues, however, requires two other different techniques: *exoticizing*, which brings readers closer to the source culture while retaining original source language (SL) elements (Haywood, Thompson & Hervey, 2009: 71), and *domesticating* that makes the translation more accessible and familiar to the target culture, sometimes requiring additional information or adapting SL cultural elements to the target language (TL) (Haywood et al., 2009: 71).

4. METHOD

4.1 *Instruments and procedure*

The present study employs a corpus-driven approach in which data is examined "with no prior assumptions" (Lu, Ali & Ghani, 2021: 122). Subsequently, linguistic phenomena are identified based on these observations (Storjohann, 2005: 5), and, finally, a new theory is developed, subject to potential modification or improvement in the future (Lu et al. 2021: 122).

Hence, the study takes into consideration the language of adolescents gathered in two corpora: the COLT (*The Bergen Corpus of London Teenage Language*) and the LIC (*Linguistic Innovators Corpus*). The COLT is the first corpus on teenagers' language compiled by the University of Bergen in 1993, and which has approximately half a million words (Andersen & Stenström, 1996: 1). It consists of spontaneous conversations between young Londoners aged 13 to 17 from different areas. Participants recorded their conversations during their daily lives, whether with friends at home, on the streets, or at school, both inside and outside the classroom (Torgersen, Gabrielatos, Hoffmann & Fox, 2011: 99). On the other hand, the LIC contains data for the project *Linguistic Innovators: The English of Adolescents in London* (Torgersen et al. 2011: 98) carried out by Lancaster University and which dates from 2004 to 2007. It is based on interviews of 16- to 19-year-olds belonging to two London districts, Havering and Hackney. In addition, both collections are based on oral language, because examples of written language by young people can only be obtained by private texts.

As both corpora were formed by a huge number of words, it was necessary to facilitate reading and locate typical teenage expressions. To carry out this task, the Terminal has been used to analyse the texts. It is a computer application, usually with a black or white background, that allows the execution of commands and to make transformations on files (Baek, 2017). First, a command was inserted followed by the expressions to be analysed. Then, the Terminal has returned the results by highlighting them in red. The authors have selected 500 words from a conversation belonging to the same context to facilitate its use in online tools and not to lose dialogue coherence. The criteria for choosing the units to be translated focuses on the adolescence expressions (abbreviations, vocatives like *man* and *dude*, idioms) reported in Section 3. As a result, we have selected fragments that include a significant number of colloquial expressions, and these fragments serve as exemplary illustrations of how MT operates with informal language.

As for the number of words selected, the authors have chosen these specific fragments based on the character limitations imposed by the tools used. This task was employed to ensure

that the conversation remained coherent and uninterrupted, avoiding any potential disarray. Before translating the data, each corpus was edited by removing non-content elements such as hashtags, numbers and recording labels. As claimed by Rahm and Hai Do (2000: 1), editing is necessary because it takes care of removing errors and inconsistencies from the corpus in order to improve the quality of data.

As it was introduced in the previous section of this paper, *Apertium*, *Systran* and *Google Translate* have been chosen because each of them represents a different approach to MT, along with being a means to observe how MT has experienced significant growth over the years. In addition, all three provide a free service that everyone can easily access. *Apertium* is a rule-based system. It emerged with the intention of providing a platform to build MT systems for related (e.g., Spanish and Catalan) and non-related languages (e.g., Spanish-English) (Forcada, Tyers & Ramírez Sánchez, 2009: 1). Its goal is to generate a simpler translation, easier to post-edit so that linguistic complexity is kept as low as possible (Forcada, Ginestí-Rosell, Nordfalk, O'Regan, Ortiz-Rojas, Pérez-Ortiz, Sánchez-Martínez, Ramírez-Sánchez & Tyers, 2011: 19).

Regarding *Systran,* it is a hybrid system which was traditionally born as a rule-based system and, subsequently, it was combined with a corpus-based one. It contains extremely detailed linguistic rules and a large terminology database covering various domains (Yang & Croiset, 2009: 3). Finally, *Google Translate,* offers a neural translation system. Since 2016, *Google Translate* has adopted an AI system based on machine learning and neural networks (Tavosanis, 2019: 3) which leads users to translate whole sentences avoiding the introduction of one fragment at a time to ensure grammatical cohesion. It is also worth mentioning that both, *Systran* and *Google Translate,* present a 5,000 characters limit, while *Apertium* appears to be the only online translator, among the analysed, that allows a limitless number of characters.

As mentioned in the previous section, the translations will be evaluated taking into account the human method by classifying errors in terms of accuracy and fluency. Based on Webster et al.'s model (2020), accuracy errors are those that give rise to mistranslation during the translation process, which implies that some SL words could be omitted or left untranslated in the TT. Similarly, some elements could lead to incorrect or ambiguous translation due to a cultural reference, multi-word expressions or words semantically unrelated. Fluency errors affect text coherence; thus, they are related to grammar and syntax (such as word order, agreement, and verb tenses), lexicon (non-existing word in the TL) and orthography (mainly, punctuation). Finally, an alternative version will be provided as a means to observe how post-editing functions. First, we will insert the text in an online tool and, later, we will modify all those erroneous expressions or words so that the result would be as reliable as possible to Spanish colloquial language. To do this, we will rely on the abovementioned translation techniques: exoticizing, borrowing, calque, literal translation, transposition, modulation, and adaptation. We refer to Spanish colloquial language, as the Spanish used in daily communication and in informal situations, characterized by slang, spontaneity, linguistic economy and idiomatic expressions, that may be challenging to understand for individuals who are not acquainted with it (Garrido Rodríguez, 2000; Suárez González, 2022).

Section 5 analyses twelve examples from COLT and twelve from LIC to ensure the analysis is as evenly balanced as possible. Furthermore, the selected excerpts vary in length, ranging from 2 to 35 characters, as different types of teenage language such as slang, idioms, phrasal verbs, and abbreviations will be explored. The choice of this word range is also intended to visually represent the outcomes for each fragment in tables.

5. RESULTS

Subsequently, the fragments and their translations will be examined in each of the three tools. *Apertium* tends to generate accuracy errors that affect the comprehension of the context as well as the reading of the text. In fact, if some sentences were extracted from the context, it would be difficult, or even impossible, to understand what the speakers were talking about. This incomprehensibility is caused by the omission of certain words, as many phrases remain unfinished, e.g., the sentence *It's called having an advantage. As soon as we get back, I'm going home alright* (COLT) is translated as 'Es llamó habiendo una ventaja. Apenas volvemos, estoy yendo casa, bien'. The speaker, in this example, referred to his interlocutor's possibility to stay at home instead of going to school, and then to come back home during recess.

Hence, if we found this sentence in isolation, we could not understand the context due to omissions and incorrect verb tenses. Another challenge that may cause unintelligibility is untranslated words, which are typically abbreviations, such as *dunno* and *cos* (see Table 1).

Table 1: **Accuracy errors in Apertium**

SOURCE TEXT	APERTIUM	POST-EDITING
Anyway, I pretend to take ages in the toilet. (COLT)	En todo caso, finjo tomar edades en el lavabo.	De todos modos, finjo que llevo mucho tiempo en el baño.
Getting her stuff ready for work. (COLT)	Cogiendo su material a punto para obra.	Preparando sus cosas para ir a trabajar.
Well, it's so bloody boring here. (COLT)	Bien, es tan sangriento aburriendo aquí.	Bueno, es un puto coñazo estar aquí.
Oh, what a stink. (COLT)	Oh, lo que un hedor.	Oh, ¡qué cabrón!
For three year I went to a thingy, innit? Well, to a boarding school in (name of place) or something like that. I think it's a military school, innit? (LIC)	Para tres año fui a un thingy, innit? Bien, a una escuela de abordaje en (nombre de sitio) o a algo le #gustar aquello. Pienso que es una escuela militar, innit?	Durante tres años fui a eso, ¿no? Bueno, a un internado en (nombre del lugar) o algo así. Creo que es una escuela militar, ¿no?
Sue: What sort of music are you into? Talal: Rock! (LIC)	Demanda: ¿Qué clase de música eres a? Talal: Rock!	Sue: ¿Qué tipo de música te gusta? Talal: Rock!
East Ham! (LIC)	Jamón del Este!	East Ham!
I dunno cos what I realised down London ways the youth there, like… they start like drinking whatever smoking drugs and shit when they're like how old? (LIC)	Yo dunno cos me di cuenta abajo maneras de Londres, la juventud allí, gusta… empiezan gusta beber cualesquier drogas y mierda cuándo son gusta qué viejo?	No lo sé, porque de lo que me he dado cuenta, allí en Londres, es la manera en la que los jóvenes empiezan a beber cualquier cosa, fuman drogas y mierda, cuando tienen como… ¿cuántos años?

As Table 1 shows, *Apertium* tends to provide a literal translation, often producing a nonsensical version of the text. It also seems that this application is unable to distinguish proper names of a given city, as in the case of East Ham, and personal names that it mistakenly translates as a common noun, for instance, *Sue* is translated as *Demanda*.

Additionally, *Apertium* also presents some challenges when translating idioms. Indeed, it takes the adjective *boring* and translates it as a non-finite form ending in *-ing*. Therefore, to carry out the post-editing, the technique of cultural transplantation has been used so that a typical expression from the SL culture has been adapted to the TL culture (Haywood et al., 2009: 80). Finally, it is also worth mentioning the mistranslation of phrasal verbs. The system does not consider multi-word expressions; therefore, it translates *are you into* by taking into account only the first part of the phrase, and translating the preposition as if it were a movement.

As far as *Systran Translate* is concerned, it also produces more accuracy than fluency errors. This system deletes the names of some speakers, keeping only some of them. In the case of the COLT, for instance, it only maintains *Alex*.

*Table 2: **Accuracy errors in Systran***

SOURCE TEXT	SYSTRAN	POST-EDITING
Anyway, I pretend to take ages in the toilet. (COLT)	De todos modos, finjo que llevo años en el baño.	De todos modos, finjo que llevo mucho tiempo en el baño.
Alex: Getting her stuff ready for work. (COLT)	Preparando sus cosas para el trabajo.	Preparando sus cosas para ir a trabajar.
Well, it's so bloody boring here. (COLT)	Bueno, esto es muy aburrido.	Bueno, es un puto coñazo estar aquí.
Oh, what a stink. (COLT)	Oh, qué mal olor.	Oh, ¡qué cabrón!
For three year I went to a thingy, innit? Well, to a boarding school in (name of place) or something like that. I think it's a military school, innit? (LIC)	Durante tres años fui a una cosa, ¿no? Bueno, a un internado (nombre del lugar) o algo así. Creo que es una escuela militar, ¿no?	Durante tres años fui a eso, ¿no? Bueno, a un internado en (nombre del lugar) o algo así. Creo que es una escuela militar, ¿no?
Sue: What sort of music are you into? Talal: Rock! (LIC)	¿Qué tipo de música te gusta? Rock!	Sue: ¿Qué tipo de música te gusta? Talal: Rock!
East Ham! (LIC)	Jamón del Este!	East Ham!
I dunno cos what I realised down London ways the youth there, like… they start like drinking whatever smoking drugs and shit when they're like how old? (LIC)	No sé porque de lo que me di cuenta en Londres es de la manera en que los jóvenes allí, como empiezan a beber lo que sea, fumando drogas y mierda cuando son como de edad?	No lo sé, porque de lo que me he dado cuenta, allí en Londres, es la manera en la que los jóvenes empiezan a beber cualquier cosa, fuman drogas y mierda, cuando tienen como… ¿cuántos años?

As shown in Table 2, *Systran* has literally translated the English word *stink*. It usually refers to a bad smell, as it has been literally translated. However, it is also used to refer to a person's bad behaviour. As a matter of fact, the Spanish equivalent would be a swear word used by young people in a playful manner among themselves, with no intention of offense. Subsequently, the sentence 'De todos modos, finjo que llevo años en el baño' reveals a "confusion of senses", as named by Costa, Ling, Luís, Correia and Coheur (2015: 10) according to whom the sentence was translated with one of its possible meanings, but, in this context, the translation seems not to be precise. The error has been solved through post-editing ('De todos modos, finjo que llevo mucho tiempo en el baño') using the transposition technique, which consists of "replacing one part of the speech with another, without changing the meaning of the message" (Vinay & Darbelnet, 1995: 36). As shown in Table 2, it is surprising that this application is able to translate *innit*, *dunno* and *cos* correctly, as opposed to *Apertium* which had not been able to provide a translation in Spanish.

Google Translate also presents some mistranslations, but it is the free online tool among the three analysed in this study that produces the most understandable and fluent texts.

*Table 3: **Accuracy errors in Google Translate***

SOURCE TEXT	GOOGLE TRANSLATE	POST-EDITING
Anyway, I pretend to take ages in the toilet. (COLT)	Pretendo tomar años en el baño.	De todas formas, finjo pasar mucho tiempo en el baño.
Getting her stuff ready for work. (COLT)	Preparando sus cosas para el trabajo.	Preparando sus cosas para ir a trabajar.
Well, it's so bloody boring here. (COLT)	Bueno, es tan jodidamente aburrido aquí.	Bueno, es un puto coñazo estar aquí.
Oh, what a stink. (COLT)	Oh, qué hedor.	Oh, ¡qué cabrón!
For three year I went to a thingy, innit? Well, to a boarding school in (name of place) or something like that. I think it's a military school, innit? (LIC)	Durante tres años fui a una cosita, ¿no? Bueno, a un internado en (nombre del lugar) o algo así. Creo que es una escuela militar, ¿no?	Durante tres años fui a eso, ¿no? Bueno, a un internado en (nombre del lugar) o algo así. Creo que es una escuela militar, ¿no?
Sue: What sort of music are you into? Talal: Rock! (LIC)	Sue: ¿Qué tipo de música te gusta? Talal: Roca!	Sue: ¿Qué tipo de música te gusta? Talal: Rock!
East Ham! (LIC)	East Ham!	East Ham!
I dunno cos what I realised down London ways the youth there, like… they start like drinking whatever smoking drugs and shit when they're like how old? (LIC)	No sé porque me di cuenta de que en Londres los jóvenes allí, como … empiezan a beber lo que sea, fumar drogas y esa mierda cuanto tienen como ¿cuántos años?	No lo sé, porque de lo que me he dado cuenta, allí en Londres, es la manera en la que los jóvenes empiezan a beber cualquier cosa, fuman drogas y mierda, cuando tienen como… ¿cuántos años?

Table 3 shows that it presented problems as well as *Systran* when translating the same expressions. But *Google* translates more formally the word *stink* > 'hedor'. On the other hand, young English speakers use the expression *thingy* when they do not remember the name of something (Palacios Martínez, 2011: 115); instead, in Spanish, a demonstrative adjective is usually used rather than naming something as 'cosita'.

In regard with idioms, it can be observed that *Google Translate* produces a closer translation to the one in human post-editing, but always maintaining a formal register. Nonetheless, it still presents some challenges as in the case of *to take ages*, considering the expression as two separate phrases and, consequently, translates them as a verb and a noun. Finally, it should be noted that *Google Translate* literally translates the musical genre *rock* as a noun ('roca'). Furthermore, we can see the MT progress with personal names' translation, seeing that it has been the only one of the three services analysed to keep *East Ham* intact.

6. DISCUSSION

As for our first research question (what are the most frequent fluency and accuracy errors in each translation tool?), our data reveal that the three translator applications tend to render more accuracy than fluency errors. In regard with fluency errors, they mostly affect grammar, resulting in mistaken verb tenses, the use of the third person singular although the speaker uses the first, and the frequent alteration of word-order. These findings concur with Costa-Jussá and Farrús (2014: 163), who noted that semantic errors are more common than orthographic ones. They stated that when evaluating systems focusing on orthography, *Apertium* emerges as the leading performer. Indeed, our results seem to coincide with their study, as *Apertium* has demonstrated greater orthography accuracy in comparison to *Systran*. However, in contrast to *Google Translate*, both systems seem to generate a similar precision.

Costa-Jussá and Farrús (2014: 163) also suggest that in the case of semantic evaluations, *Google Translate* demonstrates excellence. The analysis we have covered agrees with this assertion. Consequently, accuracy is closely correlated with semantics and to some extent with syntax, but it does not include orthography and morphology, as these levels do not impact the overall comprehension of the translation (Costa-Jussá & Farrús, 2014: 164).

As for RQ2 (are there any mistranslations that are repeated in the three MT systems? If so, which ones are the most common?), our findings show the difficulties *Apertium* presents to translate proper names of cities or people, for instance, *East Ham* as 'Jamón del Este'. This outcome could be explained on the grounds of Hurskainen (2013: 1) assertion "there is a group of proper names, which need no translation. They can be transferred to the target language as such". This is done by the exoticizing technique, thanks to which the ST form is maintained since there are not equivalents in the TL, in addition to encouraging the TL speakers "to inform themselves about the SL culture and adapt their expectations (Haywood et al., 2009: 75). Additionally, Table 1 also illustrates that *Apertium* encounters difficulties when translating teenagers' vague language, such as *like, innit* and *thingy* (Palacios Martínez, 2011: 115). These are called untranslated errors, which, according to Costa et al. (2015: 8), occur because "the engine cannot find any translation candidate to a given source word, an option is to copy it to the translation output". Phrasal verbs seem to be a significant challenge in *Apertium* as it does not treat them as single units; instead, the first part is rendered as a verb and the second part as an adverb (Thiruumeni, Anand, Dhanalakshmi & Soman, 2011: 37).

Systran seems to produce relatively fewer errors compared to *Apertium,* which enhances text comprehension. The results indicated that it does not translate swear words and colloquial expressions (*what a stink* or *bloody boring*). In the same vein, Sharou and Specia (2022: 10) claimed that when the MT is dealing with insulting or offensive words, it tends to produce incorrect translations that are impossible to understand or with a different meaning than the SL. Our data seem to concur with this assertion. Furthermore, most of the errors in *Systran* are primarily attributed to the lack of punctuation marks, suggesting that it may face challenges in generating a coherent discourse. Nonetheless, the translated text is consistently comprehensible, as seen with the expression *to take ages.* Even though *Systran* offers a literal translation, it does not affect the understanding of the text. Moreover, both *Systran* and *Google Translate* effectively manage phrasal verbs, producing accurate translations.

Likewise, discourse coherence can be appreciated in *Google Translate,* although it often provides literal translations of many typical teenage expressions. This is the case of the genre-specific music name (rock) that, as suggested by Costa et al. (2015: 11), the term should remain unchanged as it does not constitute a semantic error, but the chosen word is not appropriate in the context, and because it also lacks an equivalent in Spanish.

In regard with RQ3 (is post-editing necessary? If so, what kind of post-editing does the translator need to apply?) the examples recently examined portray that these three online tools tend to produce more accuracy than fluency errors. Consequently, the translator needs to perform post-editing to guarantee that the translated content is both coherent and understandable to the intended audience in the TL.

In this vein, it can be noticed that all the systems of MT designed so far have an important drawback which is their limited world knowledge. The major obstacle is that it is currently crucial to include all the knowledge that might be needed to resolve all possible ambiguities (Hutchins & Somers, 1992: 93), because these tools are unable to understand certain expressions that would be easily interpreted by a human reader (Gaspari & Zanchetta, 2011: 63). In fact, as observed in the results section, none of the examined MT tools succeeded in providing an accurate translation for *what a stink*, even though the expression was placed in a specific context and was not presented in isolation.

In recent years, indeed, it has been concluded that the MT output must be useful and not perfect, as it needs to be considered a means to facilitate intercultural communication (Hutchins, 2003; Li, Graesser & Cai, 2014; Muftah, 2022). The improvements that MT has made in the language field have been rather slow mainly because of the difficulties that arose when translating cultural differences. Consequently, as Hutchins (2002: 18) also mentions, it is doubtful that any machine we can imagine at this time will be able to overcome them. Similarly, it is also dismissed the possibility that MT systems would replace the professional human translator (Gaspari & Hutchins, 2007; Zhao, 2022), but, despite this, they are regarded as a valuable resource because they enable humans to work faster in creating an initial draft and in the post-editing process (Gaspari & Hutchins, 2007; Leiva Rojo, 2018).

7. CONCLUSION

The present paper attempted to show what types of errors the three free online tools present when translating from the English to Spanish language, paying special attention to mistranslation. After having analysed the translations produced by *Apertium* (RBT), *Systran Translate* (CBT) and *Google Translate* (NT), we can conclude that accuracy errors are the most common, especially when they had to translate teenage cultural expressions into Spanish. Therefore, they generate a very literal translation that makes it difficult for TL speakers to understand the TT. This is mainly due to the ample world knowledge humans possess compared to MT tools. For this reason, the results produced by MT can be useful and understandable only if they are subsequently post-edited by a human translator, who applies diverse translation strategies.

Nevertheless, as we have seen in the present work, the character limit can be a restrictive factor. Both *Google Translate* and *Systran* allow the translation up to 5,000 characters, which means that they make it impossible to translate longer texts without splitting them into shorter excerpts. In fact, the text should be cut and inserted one fragment at a time, which can lead to a loss of coherence. As far as *Apertium* is concerned, this tool has no character limit, but our findings reveal it does not render a reliable translation product.

According to our results, another challenge that these three tools have to face with is their lack of precision since they translate words that should not be translated, such as abbreviations, specific terms or proper names. However, while *Apertium* and *Systran* have problems generating a correct result when the ST has no punctuation marks, *Google Translate* seems to be more accurate. Indeed, in most cases, it tends to place the correct punctuation in the output, allowing the exact understanding of the ST.

One limitation of the present research is that the analysis has been carried out taking into account free access tools, therefore we do not certainly know whether these limits are also repeated in premium versions, or whether they offer an improved service. Further research is needed to validate or refute this hypothesis.

As have been shown all through the paper, MT has undergone significant evolution over the years, and tools like *Apertium, Systran,* and *Google Translate* have played a prominent role in this evolution. Despite its progress, MT has not yet achieved the high-quality human translation has reached, primarily because human translators can understand context, cultural diversity, idiomatic expressions, and subtle linguistic nuances that machines find challenging to replicate accurately. Additionally, human translators possess the ability to interpret ambiguous content and provide translations that are not only linguistically correct but also culturally appropriate, considering the target audience's expectations. Hence, further research is needed with a wider amount of teenage language corpora and implementing the metric

evaluation method to quantity the degree of sufficiency and usability of these three online translation tools.

Finally, it would be interesting to open new lines of research on English-Spanish/Spanish-English MT, specifically concerning how these tools handle idioms or proverbs, due to the fact that they are linguistic aspects that frequently lack a direct equivalent in the TL and may lead to mistranslations.

ACKNOWLEDGMENTS

REFERENCES

Andersen, G. & Stenström, A. B. (1996). COLT: A progress report. *ICAME Journal,* 20, 133-136. Retrieved from https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=7fc6c98bc883d469a7d67b272afc770a7109b5d1

Baek, J. (2017, January). Introduction to terminal. In *Computing in Optimization and Statistics. Lecture 1*. Retrieved from */www.philchodrow.prof/cos_2017/1_terminal_and_git/Introduction%20to%20terminal.pdf*

Brusasco, P. (2018). La traduzione automatica. Un po' di storia: successi e qualche riflessione. *Tradurre. Pratiche teorie strumenti*, 14, online. Retrieved from https://rivistatradurre.it/la-traduzione-automatica/

Costa-Jussà, M., Farrús, M., Mariño, J. B. & Fonollosa, J. A. R. (2012). Study and comparison of rule-based and statistical Catalan-Spanish machine translation systems. *Computing and Informatics*, 31(2), 245-270. Retrieved from https://repositori.upf.edu/handle/10230/30925

Costa-Jussá, M. & Farrús, M. (2014). Towards human linguistic machine translation evaluation. *Digital Scholarship in the Humanities*, 30(2), 157-166. doi: 10.1093/llc/fqt065

Costa, Â., Ling, W., Luís, T., Correia, R. & Coheur, L. (2015). A linguistically motivated taxonomy for Machine Translation error analysis. *Machine Translation*, 29(2), 127-161. Retrieved from www.jstor.org/stable/44113801

Dias Esqueda, M. (2021). Machine translation and learning issues. *Trabalhos em Linguística Aplicada*, 60(1), online. doi: 10.1590/01031813932001520210212

Eckert, P. (2003). Language and adolescent peer groups. *Journal of Language and Social Psychology,* 22(1), 112-118. doi: 10.1177/0261927X02250063

Fernández-Torné, A. & Matamala, A. (2021). Human evaluation of three machine translation systems: From quality to attitudes by professional translators. *Vigo International Journal of Applied Linguistics,* 18, 97-121 doi: 10.35869/vial.v0i18.3366

Forcada, M. L., Tyers, F. & Ramírez Sánchez, G. (2009). The Apertium machine translation platform: Five years on. In J. A. Pérez-Ortiz, F. Sánchez-Martinez & F. M. Tyers. *Proceedings of the First International Workshop on Free/Open-Source Rule-Based Machine Translation,* (pp. 3-10). Alicante: University of Alicante. Retrieved from https://aclanthology.org/2009.freeopmt-1.3/

Forcada, M. L., Ginestí-Rosell, M., Nordfalk, J., O'Regan, J., Ortiz-Rojas, S., Pérez-Ortiz, J. A., Sánchez-Martínez, F., Ramírez-Sánchez, G. & Tyers, F. M. (2011). Apertium: A free/open-source platform for rule-based machine translation. *Machine Translation*, 25(2), 127-144. Retrieved from https://www.jstor.org/stable/41487458

Garrido Rodríguez, M. C. (2000). ¿Qué español coloquial enseñar en las clases de E/LE? In M. A. Martín Zorraquino & C. Díez Pelgrín (Eds.), *Actas del XI Congreso Internacional ASELE*. Zaragoza: University of Zaragoza. Retrieved from https://dialnet.unirioja.es/servlet/articulo?codigo=608273

Gaspari, F. & Hutchins, J. (2007). Online and free! Ten years of online machine translation: Origins, developments, current use and future prospects. In B. Maegaard (Ed.), *Proceedings of Machine Translation Summit XI: Papers*, Copenhagen: University of Copenhague. Retrieved from aclanthology.org/2007.mtsummit-papers.27/

Gaspari, F. & Zacchetta, E. (2011). Scrittura controllata per la traduzione automatica. In G. Bersani Berselli (Ed.), *Usare la Traduzione Automatica* (pp. 63-79). Bologna: Clueb.

Gouadec, D. (2010). Quality in translation. In Y. Gambier & L. Van Doorsalaer (Eds.), *Handbook of Translation Studies (Vol. 1)* (pp. 270-275). Amsterdam: John Benjamins.

Hall, D. (2020). The impersonal gets personal. *Nat Lang Linguist Theory,* 38, 117-150. doi: 10.1007/s11049-019-09447-w

Haywood, L., Thompson, M. & Hervey, S. (2009). Genre. Text type and purpose. In L. M. Haywood, M. Thompson & S. Hervey (Eds.), *Thinking Spanish Translation: A Course in Translation Method: Spanish to English* (2nd edition) (pp.47-67). London and New York: Routledge.

Hurskainen, A. (2013). Handling proper names in Machine Translation. *Technical Reports in Language Technology Report,* 12, online. Retrieved from https://www.academia.edu/98195885/Handling_proper_names_in_Machine_Translation

Hutchins, J. & Somers, H. (1992). *An Introduction to Machine Translation.* London: Academic Press Limited.

Hutchins, J. (2002). Two precursors of machine translation: Artsrouni and Trojanskij. Online. Retrieved from https://aclanthology.org/www.mt-archive.info/IJT-2004-Hutchins.pdf

Hutchins, J. (2003, March). *Machine translation and computer-based translation tools: What's available and how it's used.* Valladolid, University of Valladolid. Retrieved from https://www3.uji.es/~aferna/course4/MT%20process.pdf

Leiva Rojo, J. (2018). Aspects of human translation: The current situation and an emerging trend. *Hermēneus. Revista de traducción e interpretación,* 20, 257-294. doi: 10.24197/her.20.2018.257-294

Li, H., Graesser, A.C. & Cai, Z. (2014). Comparison of Google translation with human translation. In W. Eberle & C. Boonthum-Denecke (Eds.), *Proceedings of the Twenty-Seventh International Florida Artificial Intelligence Research Society Conference* (pp. 190-195). Florida: Pensacola Beach. Retrieved from https://cdn.aaai.org/ocs/7864/7864-36722-1-PB.pdf

Lu, R., Ali, A. M. & Ghani, C. A. A. (2021). A comparative study of corpus-based and corpus-driven approaches. *Linguistics International Journal*, 15(2), 119-132. Retrieved from https://connect.academics.education/index.php/lij/article/view/157

Marshman, E. (2023). *Outside The black box: Machine Translation and the importance of what we don't know.* Canada: University of Ottawa. Retrieved from https://dhsiteorg.files.wordpress.com/2023/02/marshman_dht2023.pdf

Molina, L. & Hurtado Albir, A. (2002). Translation techniques revisited: A dynamic and functionalist approach. *Meta XLVII*, 4, 499-500. doi: 10.7202/008033ar

Muftah, M. (2022). Machine vs Human Translation: a new reality or a threat to professional Arabic-English translators. *PSU Research Review*, online. doi: 10.1108/PRR-02-2022-0024

Palacios Martínez, I. M. (2010). The expression of negation in British teenagers' language: A preliminary study. *Journal of English Linguistics*, 39(1), 4-35. doi: 10.1177/0075424210366905

Palacios Martínez, I. M. (2011). The language of British teenagers: A preliminary study of its main grammatical features. *Atlantis*, 33(1), 105-126. Retrieved from https://minerva.usc.es/xmlui/bitstream/handle/10347/21543/1/2011_atlantis_palacios_language.pdf

Palacios Martínez, I. M. (2013). Non-standard negation in Modern English: A corpus-based study of four salient features. *ES Review. Spanish Journal of English Studies*, (34), 211-226. Retrieved from https://revistas.uva.es/index.php/esreview/article/view/2183

Palacios Martínez, I. M. (2018). 'Help me move to that, blood', A corpus-based study of the syntax and pragmatics of vocatives in the language of British teenagers. *Journal of Pragmatics,* 130(4), 33-50. doi: 10.1016/j.pragma.2018.04.001

Palacios Martínez I. M. (2021). Recent changes in London English: An overview of the main lexical, grammar and discourse features of Multicultural London English (MLE). *Complutense Journal of English Studies*, 29, 1-20. doi: 10.5209/cjes.77504

Popović, M. (2020). Relations between comprehensibility and adequacy errors in machine translation output. In R. Fernández & T. Linzen (Eds.), *Proceedings of the 24th Conference on Computational Natural Language Learning* (pp. 256-264). Online version: The Association for Computational Linguistics. Retrieved from https://aclanthology.org/2020.conll-1.19/

Pym, A. (2020). Quality. In M. O'Hagan (Ed.), *The Routledge Handbook of Translation and Technology* (pp. 437-449). New York: Routledge.

Rahm, E. & Hai Do, H. (2000). Data cleaning: Problems and current approaches. *IEEE Data Eng. Bull,* 23, 3-13. Retrieved from https://dbs.uni-leipzig.de/research/publications/data-cleaning-problems-and-current-approaches

Sharma, S., Diwakar, M., Singh, P., Singh, V., Kadry, S. & Kim, J. (2023). Machine translation systems based on classical-statistical-deep learning approaches. *Electronics*, 12(7), 1716. doi: 10.3390/electronics12071716

Sharou, K. A. & Specia, L. (2022). A taxonomy and study of critical errors in Machine Translation. In H. Moniz, L. Macken. A. Rufener, L. Barrault, M. R. Costa-jussà, C. Declercq, M. Koponen, E. Kemp, S. Pilos, M. L. Forcada, C. Scarton, J. Van den Bogaert, J. Daems, A. Tezcan, B. Vanroy & M. Fonteyne (Eds.), *Proceedings of the 23rd Annual Conference of the European Association for Machine Translation* (pp. 171-180). Ghent: Ghent University. Retrieved from https://aclanthology.org/2022.eamt-1.20/

Storjohann, P. (2005). Corpus-driven vs. corpus-based approach to the study of relational patterns. In University of Birmingham (Ed.), *Proceedings of the Corpus Linguistics Conference 2005* (pp. online). Birmingham: University of Birmingham. Retrieved from https://ids-pub.bsz-bw.de/frontdoor/deliver/index/docId/5006/file/Storjohann_Corpus_driven_vs_corpus_based_approach_to_the_study_of_relational_patterns_2005.pdf

Suárez González, I. (2022). Una aproximación a la anticortesía y su explotación didáctica en el aula de ELE. *Foro de profesores de E/LE*, 18, 199-229. doi: 10.7203/foroele.18.25311

Tagliamonte, S. A. (2016). *Teen Talk. The Language of Adolescents*. Cambridge: University of Cambridge.

Tavosanis, M. (2019). Valutazione umana di Google Traduttore e DeepL per le traduzioni di testi giornalistici dall'inglese verso l'italiano. In R. Bernardi, R. Navigli & G. Semeraro (Eds.), *CLiC-it 2019. Proceedings of the Sixth Italian Conference on Computational Linguistics* (pp. 1-7). Retrieved from https://ceur-ws.org/Vol-2481/paper70.pdf

Thiruumeni, P.G., Anand, K., Dhanalakshmi, V. & Soman, K.P. (2011). An approach to handle idioms and phrasal verbs in English-Tamil Machine Translation system. *International Journal of Computer Applications,* 26, 36-41. doi: 10.5120/3139-4328

Torgersen, E. N., Gabrielatos, C., Hoffmann, S. & Fox, S. (2011). A corpus-based study of pragmatic markers in London English. *Corpus Linguistics and Linguistic Theory*, 7(1), 93-118. doi: 10.1515/cllt.2011.005

Vela-Valido, J. (2021). Translation quality management in the AI age. New technologies to perform translation quality management operations. *Revista Tradumàtica. Tecnologies de la Traducció*, 19, 93-111. doi: 10.5565/rev/tradumatica.285

Vieira, L. N. (2020). Post-editing of machine translation. In M. O'Hagan (Ed.), *The Routledge Handbook of Translation and Technology* (pp. 319-331). New York: Routledge.

Vinay, J. P. & Darbelnet, J. (1995). *Comparative Stylistics of French and English: A Methodology for Translation.* Amsterdam: John Benjamins.

Webster, R., Fonteyne, M., Tezcan, A., Macken, L. & Daems, J. (2020). Gutenberg goes neural: Comparing features of Dutch human translations with Raw Neural Machine Translation outputs in a corpus of English literary classics. *Informatics,* 7(32), online. doi: 10.3390/informatics7030032

Yang, J. & Croiset, T. (2009). *SYSTRAN Chinese-English and English-Chinese Hybrid Machine Translation Systems.* Retrieved from aclanthology.org/www.mt-archive.info/CWMT-2009-Yang.pdf

Zhao, L. (2022). The relationship between Machine Translation and Human Translation under the influence of Artificial Intelligence Machine Translation. *Mobile Information Systems*, 5, 1-8. doi: 10.1155/2022/9121636