UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Università degli Studi di Padova

Dipartimento di Studi Linguistici e Letterari

Dipartimento di Ingegneria Informatica

Corso di Laurea Magistrale in
Lingue Moderne per la Comunicazione e la Cooperazione Internazionale
Classe LM-38

Tesi di Laurea

"Spare me your medical mumbo-jumbo": A comparison among neural machine translation applications in the medical domain

Relatore
Prof. Giorgio Maria Di Nunzio
Correlatrice
Prof.ssa Maria Teresa Musacchio

Laureando
Lorenza Gattini
n° matr.1171859 / LMLCC

Anno Accademico 2019 / 2020

Abstract

In this thesis we discuss the role of machine translation, focusing especially on neural machine translation applications. We make a comparison among Google Translate and DeepL. through an error-analysis present in the translated texts. We have chosen three specialized and three popular medical articles concerning sexually transmitted diseases. Throughout the thesis we demonstrate how developed machine translation have become thanks to the implementation of word embedding, one of the main feature of neural network. Starting from an overview of the evolution of English and Italian medical language, focusing on the discipline of terminology, we then explore the functioning of neural machine translation, outlining the theoretical concepts of those who laid the foundation of machine translation: Warren Weaver and Claude Shannon. We try to understand the linguistic concept behind them taking up the idea of pivotal philosophers of language – Ferdinand de Saussure, Ludwig Wittgenstein and J. R. Firth. Indeed we explore how ideas of these thinkers contribute to and underpin the mechanism that allow neural machine translation applications to distinguishes words' meaning thanks to context information. Contextual information and distributional semantics are two theories that help us in this task, giving us the tools to define the concept of meaning from a statistical point of view. Then it follows the main chapter of the thesis, the comparative analysis between Google Translate and DeepL. The quality analysis of machine translations outcome is based on DQF/MQM harmonised integrated error typology. Finally, in addition to a micro-analysis of the texts, we conduct a register analysis to understand if and how machine translations can adapt the speech according to the addressees.

Table of Contents

Introduction

The use of machine translation to aid translation workflow is a phenomenon that has become increasingly common in recent years. The question of the future of the role of specialist translators is an increasingly debated issue among experts in the field, undermined as it is by the continuous development of IT and the risk of seeing deeply restructured the very nature of their job. These changes, due to the disruptive development that in recent decades have marked the field of computer science and computational science, are only a small part of an issue that entails the very structure of work in our societies, whose restructuring requires long-term strategic and political changes. The research that we present in the following pages aims to contribute to the reasoning on some of these nodes starting from a precise and micro-logical analysis. Indeed, we believe that my knowledge about the linguistic questions and my ignorance about the highly technical matters concerning machine translation do not represent necessarily a detriment: productive analysis and abstract reflections always need, as their necessary precondition, a precise and meticulous investigations on aspects that could appear at first sight as minimal, even negligible. It is in this spirit that the thesis proposes a qualitative comparative analysis of two machine translation applications that have incorporated the most recent innovations in the field of computational semantics into their algorithms: Google Translate and DeepL. Both software have become the protagonists of a true Copernican revolution in the field of automated translation, moving from a mode of operation based on the storage of terminological data to the use of the neural network architectures. Neural machine translation deeply changes the very philosophy on which machine translation is based, being inseparable from a different theory of the nature of language. A theory that, as we will see in the first introductory chapter, incorporates many of the suggestions and reflections elaborated in the field of the humanities: it is no longer about a mere relationship between signifier and signified. Through the vectors constituting the neural networks the idea of meaning that is continuously built up by virtue of the relationships between signifiers, contexts, and usages makes its way. The scale of the challenge is impressive. What are the margins of success? In order to answer, in an evidently partial way, such a question, we have

thought to proceed inductively, analyzing both machine translation applications actual functioning, shortcomings and errors, rather than entering into a technical reflection on the basis of assumptions that are difficult to demonstrate given my knowledge in IT questions. In my thesis, therefore, I will analyze the two machine translation applications, analyzing two corpora of texts of different nature, both of which have sexually transmitted diseases as their object. The first corpus will in fact be formed by specialized papers, with strongly standardized terms, while the second will be made up of popular articles, which obey to a different kind of rules. Following the very logic of the neural network, we should have found translations capable of changing as the specific social fields in which the texts act (vocabulary, audience, reference context) change. As we will see, if to a large extent this happens with relative success, the results are far from perfect. The analysis of inaccuracies, as well as misunderstandings, has allowed me to deduce the specific difficulties and blind spots that characterize the neural network system. While some are easy to improve, and I am sure it is a matter of time for them to be overcome, others involve complex reasoning and distinctions that are not easy to resolve. The road of comparison between Google and DeepL is useful to show how the errors are similar and recurrent and therefore not attributable to a single software. DeepL has proven to be incredibly accurate in the translation of specialized medical texts, while Google Translate provides better results in terms of popular texts. The reasons for this difference will be investigated in the course of the thesis. On the other hand, it is certainly indicative that in the very period of the writing of the work, between its start and its end date, DeepL has considerably improved the quality of its functioning, demonstrating to possess the ability to adapt and improve itself, imitating the plasticity that, according to some philosophers, would be the only ability really attributable to human nature. In conclusion, leaving aside the catastrophic tones that often accompany the discussion of these issues, the thesis constitutes a constructive and critic contribution to the perspectives of machine translation. As can be seen from the analysis of the two corpora, rather than a suppression of the role of translators, we are faced with a necessary transformation in which technology can become an important tool to facilitate his more mechanical work, so that translators can devote their intellectual skills and theoretical knowledge to deliver the most accurate and precise

outcomes. Precisely for these reasons I believe that contributing to the development and improvement of these software is pivotal, not only because the needs of the processes of economic globalization put us in front of stressful situations difficult to avoid, but above all because a careful, critical and conscious use of these tools can allow a more rapid dissemination of high quality outcomes, that can lead us towards a "democratization" of knowledge and communion of intellects. A democratization that often seems unreachable because of – to borrow a famous quote by Homer Simpson – the “mumbo-jumbo” discourses that makes communication between lay people and specialists impossible, especially in the medical domain. Although the existence of a specialized language belonging to a scientific community is necessary in order to render communication effective, it is essential that those belonging to such community are able to adapt the register according to their addressees. All this, keeping in mind that, as summarized by G. C. Spivak (1993), in translation "one betrays oneself out of love".

1. AN OVERVIEW OF MEDICAL LANGUAGE

1.1 A diachronic perspective of medical English and Italian language

As Leon McMorrow tells us (McMorrow 1998), medicine has always had a long history of writing. All the great civilization had the tendency to record and organize medical findings through systems which allows to craft transcripts of medical research. In the western world, the cradle of medical language in the V century B.C. is Greece, whose brilliant scientists had not a technical terminology to name their discoveries, so they had to draw on common words, enriching them with specific meaning through analogical mechanisms or exploiting the word formation processes, something that a language as flexible as Greek allowed to do (Mapelli 2014). Greek civilization passed its heritage first to the Roman Empire and then to the Medieval Europe. It played a predominant role for centuries, so much so that its traces are still discernible today. Greek physicians developed and spread their knowledge based on a new approach, the naturalist method, which focused on the description of disease and anatomy, leaving aside the religious-magical orientation. Of course, all this was possible thanks to the huge hegemony exercised by Greece in all fields of knowledge, first of all the philosophy which encompassed several disciplines. In an immense empire as the Greek one was, with its learning centers dispersed in Europe and Asia Minor, writing represented an essential medium to communicate and register knowledge and findings. With the move of the political and cultural center to Rome, Greek scholars continued to be held in high regard. Greek civilization was definitely eradicate with the Muslim conquest in the VII century, and medical Greek corpus was progressively translated into the local languages of Eastern Roman and Byzantine Empire. So, many of the Greek works became available in Latin via Arabic intermediate translation. One of the first to be translated was Galen, whose writings had an important influence upon Islamic culture. The first major translator of Galen was Hunayn ibn Ishaq, who during the Abbasid era worked with a group of translator to translate most of the works of Greek corpus into Arabic and Syriac. Galen's writings were then translated from Arabic to Latin between 1000 A.D and 1200 A.D. by Muslims, Jewish, and Christian scholars.

One of the most important figure was Constantine the African, who studied medicine in Egypt and became a monk in the monastery of Monte Cassino and translated medical works from Arabic, including the ancient medicine of Hippocrates and Galen adapted by Arabic physicians. Arabic language represented a link to Greek scientific and philosophical thought, but then was progressively replaced by Latin in the period between 1250 and 1500, when Greeks works were translated directly into this language. After the later medieval period, Late Latin became the medium of communication in the most important European universities. It assimilated Greek and Arabic terminology thanks to the flexibility of its morphological structure which allows procedures such as prefixation and suffixation. Nonetheless by 1800 Latin as cultural lingua franca came to an end because of the advance of local languages, however all medical Western languages resemble each other, by virtue of the their shared Greek-Latin core.

With the demise of Roman Empire, Britain became an open area attractive for new invasions. Small isolated of migrating German people began to invade Britain. These tribes – Jutes, Angles, and Saxons – pushed the Celtic population present on the isles into the regions of Wales and Scotland. As McMorrow puts it "languages develop in isolation […]. But they also develop by contact, as seen in some political conquests" (McMorrow 1998). Indeed, through the contact of tribes' dialects born a unique language used as lingua franca, which allows communication all over the country. This language was called Englisc, better known as Old English. Englisc means pertaining to the Angles, one of the tribe which conquered part of Great Britain. It is not to be regarded as monolithic entity, since it emerged over a long period of time during which contacts between different dialects influence each other. As for Celtic languages, linguistics agree that it was not an influential feature in the development of Englisc lexicon, however it is thought, especially by a school of thought known as Brittonicism, that it may have influenced on developments in English syntax in the post-Old English period, especially in the shift from syntheticism to analyticism (German, 2001). Clearly also several loanwords from Latin were present in the Englisc, because of the contacts occurred during the battles to take over the control of the isles. During the IX century, Englisc-language literature began to flourish, the best known work of which is the Anglo-Saxon Chronicle.

On the European continent France was invaded by Norsemen, who adopted Old French and, when they conquered Britain in 1066, they brought along French. So French spread along the isles, notwithstanding the attempts of King Alfred of increasing the education of his nation by having many Latin texts translated into Englisc. French influence upon English not only pertained to its grammar, but also to its lexicon and pronunciation. Therefore by this contact arouse a hybrid language, Norman English, an amalgam of Anglo-Saxon English and French, which is the basis of modern English. So for three centuries coexisted three different languages in Britain – a Norman-speaking governing class; native Anlgo-Saxon subordinates; and Latin used for religious and educational purposes. We have to wait substantial events such as the loss of Norman territory in 1204, the Hundred Years War, and the split from the Roman Catholic Church under Henry VIII to see the development of Middle English and the gradually the demise of French. Of course, its influence in the evolution of Middle English was very strong: French words came into use in the legal, religious, administrative, political, military, and artistic terminologies, but less in the medical lexicon. However French played a pivotal role as a medium for penetration of Latin and Romance medical terms into English (Dzuganova 2002).

In the meantime, Greek influence of medical terminology lost its power and retained its position only in the East-Roman Empire, in Constantinople, until the Turks conquered it in 1453 (Dzuganova 2002). The only resource for technical written expression in Western Europe remained Latin, but it was confined to church schools. The come back of Greek as a the scientific language par excellence occurred during the Humanism. Finally, in the late Middle Ages local medical knowledge was translated in their dialects, although Latin was retained priority in universities. With the advent of Middle English it was possible to translate the medical lore. However the need to choose an appropriate register that would suit the medical discourse arose. Indeed, as mentioned before, French was been absorbed in some specific domain, Latin was used mostly for educational purposes and an Anglo-Saxon terminology existed for certain disciplines, such as anatomy, physiology and many diseases. One of the figures who solved this dilemma was an unknown translator who translated the Compendium medicinae written in Latin by Gilbertus Anglicus. This translation is remarkable for its

preference for Anglo-Saxon medical terminology, although some traces of French are also present. Latin loanwords were utilized only when the Anglo-Saxon ones were not befitting or precise enough. From this point on followed the era of the national medical languages, which replaced Latin as vehicles for international communication. The so-called national medical languages had much in common since several medical terms were derived from specialized medical Latin (Wulff 2004). In some German languages anatomical terms and disease names are often imported directly with their correct Latin endings, while in the Latin languages Latin terms were naturalized according to the norms of each language. As for English, that is a Germanic language influenced as we have seen above by Norman domination, has half of its terminology of Romance origin. Indeed, English medical terms tend to follow the Romance pattern except in placing adjective before the noun (Wulff 2004).

However, medicine keeps developing and it is necessary to name new diseases, new treatments and diagnostic tools. In this regard English has become the prominent medium of communication of scientific community. Indeed, physicians all over the world write technical papers in English, something that contributes to craft the technical language. As McMorrow explains to us (McMorrow 1998)

> It is axiomatic that dominance in knowledge, customs or technology has major repercussion upon language relationships. What is seen as superior tends to flow into what is seen as inferior; one may view the process in terms of either push (imposition) or pull (borrowing). Whoever leads the field gets to create the words that capture the emerging concept and products. (McMorrow 1998)

Given U.S.A. dominant role in several fields – from technology to medicine and pharmacy – English has been exported from and imported into many languages. Therefore, it is obviously easier for other linguistic domains to borrow the foreign terminology, than to deliver their own languages with proper expressions. As history teaches us, “nativism in language development is probably a pipe dream” (McMorrow 1998). However, as we have seen, medical language has a particular history which has lent it a heterogeneous and layered style, given the influence of so many languages, from Greek and Latin to French and English.

We would like to point out how non-English translators as well aid to keep building medical language. Indeed, the proficiency of speaking and writing in another language is not the same for everyone and a non-English mother tongue probably will never be able to get know the language as a native-speaking. This conduct, of which not everybody are aware of, contributes to the creation of new terms or expressions in English. There are several papers which concerns how English has been modifying because it is spoken all over the world. Since nowadays the work of translation can be automatized even by people who have not expertise in this field – for instance physicians – not English languages' terminology and specific expressions are created by them with the aid of machines translation. We will see in the chapter concerning the error analysis how some Italian terms seems a translation made by non-humans. In other words, there exists a process consisting of translations made into English by not mother tongue English medical specialists and retranslations made into other languages, both conducted by the aid of machines translation. So, the advent and development of technology and artificial intelligence in machine learning abets the evolution of medical language, helping with the finding of new technical terms which fit for a given description or definition of a newly born concept. Indeed, the conventional status of a specialized language cannot be fixed for ever, but it is subject to changes that naturally occur. As we have seen above in this chapter, languages change over time because of a series of events which influence it and create new "systems" updated with what Wittgenstein would have called orders (Wittgenstein 1953). Its metaphor of the city could help us to explain better how we see that fascinating medium of communication that is language

> Our language can be seen as an ancient city: a maze of little streets and squares, of old and new houses, and of houses with additions from various periods; and this surrounded by a multitude of boroughs with straight regular streets and uniform houses. (Wittgenstein 1953)

For Wittgenstein the suburbs of this city represent the languages of games that can be seen as languages for specific purpose, that have their own rules and terminology.

As for Italian medical language, we believe that it is subject to the same changes of the English due to the improvement of machine translation. Many terms are

translated by physicians who do not know how the job of translation works and often they introduce new terms and technical expressions basing just on their knowledge. So, usually new terms come into use by the translation or the borrowing of the dominant medicine language – English. However, Italian medical language has long history as well. Its roots go back to Greek and obviously Latin. Being a Romance language, Italian medical terminology has many terms whose root has Greek origin; so much that one of the classic thing that everybody tell to young students who want to study medicine is knowing Ancient Greek would be a great tool to understand better the subject. It is indeed quite useful, since from the etymology of several Italian medical terms one can discovers its meaning. Italian medical terminology terms are created through some morphological processes such as prefixation, suffixation, circumfixation, infixation, modification, and reduplication. There are also several compound words created through processes like disjunction or juxtaposition. Finally there are many eponymous, especially concerning diseases' name. One of the differences that we noted doing the comparative analysis is that English has more popular words used to convey some concept more clearly to lay people. This is something that misses in Italian, maybe because the Latin and Greek influence plays still an important role. In the analysis we will see how gonorrhea in English has several popular synonyms unlike Italian that has more specialized terms to name the disease. This is probably due to the great direct influence that Latin, Greek and other Romance languages as Spanish and French had on Italian, which had a less tortuous evolution than English which, as we have seen above, have a really winding pattern from the philological point of view.

1.2 A focus on the discipline of terminology

It is not so easy defining the concept of terminology. It could seems a paradox since the discipline of terminology is born to avoid misunderstanding and create a shared system of language in a given field. However, terminology has many definitions. A differentiation is always made because this word refers to three concepts: terminology as the study of terms, terminology as the practical aspect of doing terminographical work, and terminology as a set of specialized terms (Brenes 2017). According to the ISO 1087-1: 2000, terminology is the science studying the structure, formation,

development, usage and management of terminologies in various subject fields. It is considered a branch of linguistics, and it can be considered related to lexicography, but with a focus on concepts in special domain. The origins of terminology can be traced back to the 18th century, when the researches in chemistry by Lavoisier and Berthollet or in botany and zoology by Linné exemplify the interest that the naming of scientific concepts has always had for the specialists (Cabré 1992). The subsequent internationalization of sciences in 19th century made manifest the need of a set of rules for formulating terms and for their respective disciplines. As we noted above, the fast development of technologies required the naming of new concepts and the agreement of the terms used by a given scientific community. However, it is only from the 1950s that linguists began to pay attention to terminology, laying the foundations of a discipline concerning the principle which governs all human languages seen as tools of communication. A great contribution for the birth of this discipline is due to scientists and technicians, who already during the 1920s got interested in terminology, such as the Austrian E. Wüster, who is considered the father of terminology (Cabré 1999). Indeed, Wüster was interested in methods of compilation and standardization of terms, ideas expressed in his pivotal work *The machine tool* (1968), which can be considered one of the first termbase ever. The Austrian linguist intuited that the rapid development of technologies gave birth to a new knowledge that needed precise organization in order to be communicated. The terminological theory arose and the work to standardize the multiplicity of languages born together with the improvement and birth of new disciplines was carried out during the 1930s by Austrian, Soviet, and Czech scholars. Terminology began to be defined as an interdisciplinary but autonomous subject, interested in the classification of concept systems and the organization of knowledge. It is multidisciplinary because it is concerned with the relationships between a given subject, for instance medicine, and linguistics. A particular focus is on the assignation of terms to concepts, in order to craft an ambiguous and clear way of communication within a given specialized community. This is the main difference with lexicography: while terminology move from the concept to term, that is it tries to name concepts with a specific and unequivocal noun, following an onomasiological process, lexicography move from the word to the concept, following a semasiological process. So,

terminology and lexicography differ in the way they deal with the methodology. Substantially, terminology's aim is mapping objects in the real world with the concepts they represent. They deal with their object of study in different ways. Lexicology studies just the word and it is not interested in meaning, unless it is related to the word. In terminology instead the meaning, that is the concept, is of pivotal importance, it is prior to the denomination: its main aim is name concept. Plus, lexicology is linked to grammar, word in dictionaries are described with respect to their use in context. On the other hand, terminology studies terms on their own account, regardless their inflections or syntax. Another important difference is that while linguists consider words from a diachronic or a synchronic point of view, terms are concerned only with synchronic aspects (Cabré 1999).

Terminology turns to be really useful in translation studies. Indeed, one of the purpose of translators is mediating communication between specialists. One of the great Wüster's intuitions was considering how important was the link between terminology and computer science. Indeed, computer science offers the opportunity to store and retrieve information and to order conceptual systems. Nowadays we take for granted computers to store terminological database, but actually it was a thought of huge importance, which led to the conceptualization and creation of termbase. A termbase is a database consisting of concept-oriented terminological entries and related information, usually in multilingual format. In this way the work of translators specialized in a given subject is aided by a fast and an ease tool by which they can retrieve some terms. A termbase sheet includes the term usage in accordance with ISO 0241-1; the subject field where the term is used; an example of context; a source confirming the validity of the entry; and a date, given that, how we have seen above, language changes relentlessly and so is terminology. Termbase sheets are of pivotal importance for translators. They facilitate translation workflow, speeding it up. Indeed over the years, terminology has become a field of study on which several scholars have investigated. Multilingual terminological activities have become more and more important for translators, whose main purpose is facilitating communication between speakers of different language. Since translation implies being familiar with the subject of the source text, translators experts in given technical fields need to have knowledge of the subject matter they are

translating. Plus, they not only have to convey the content of the source text, but also render it in an appropriate manner for the target text reader. For instance, in the case of medical specialized texts, the addressee is a specialist reader, who knows well the matter of the papers he is reading. For this reason, the translation has to comply with precise register patterns and terminological expressions, or in the words of Wittgenstein, it has to consist of a specific order. We will see in the last subject, how both DeepL and Google Translate are able to respect the rules of medical specialized language. We believe that one of the reason of their correct behavior is due to the language pair we have chosen to do the analysis: translation from English to Italian turn to be more easier for machine translations to execute, since – as we have already said – English is the lingua franca of practically every sector of knowledge, so the set of corpora on which machines translation can be trained on are countless. Plus, English plays a pivotal role in the dissemination of medical news and discoveries all around the world and it is also the language of communication in the field of computer engineering – so, it is a natural consequence that machines translation are skilled in translating tasks when English is involved. So, termbase turn to be really useful in translating unusual language pair that machines translation like DeepL and Google Translate cannot yet perform at their best. However, we do not know if these kind of machines translation are equipped with terminological database. This is an issue that we will explore in more detail in the final chapters.

The discourse is different when we deal with the translation of popular articles. Indeed although journalism language is a specialized language, it is governed by principles that are looser than the medical languages' ones. Here too the purpose is conveying a message in an effective and concise manner (Mapelli 2014). An unavoidable duty of journalists, who are the specialists in this field, is communicating current and correct information in a comprehensible way. However, as far as articles on medical subjects are concerned, writers are journalists, and therefore experts in the field of media communication, but they are nor doctors or physicians. Medical news are conveyed following the register's rules governing journalism style, but often the technical content can leave specialists readers disoriented for its inaccuracy. Other times, even if journalists try to simplify specialized questions to lay readers, their

knowledge on the matter lead them to write in a too complicated manner which compromises the understanding. Indeed, the most difficult thing of writing about specialized subjects like medicine or diseases is rendering the topics accessible for everybody. As we said in the introduction, we all remember how Homer Simpson could not understand his doctor when he was told he needed a heart surgery. The speech gets complicated when we transpose this discourse into machine translation. They do not possess the awareness to understand the difference between different speeches and their addressees. Nevertheless, the can rely on the language structure to choose the most suitable manner to convey the message. In our analysis we will see how machines translations have worst outcomes when translating popular article compared to specialized papers. One of our hypotheses is that this is due to the unpredictability of the journalism discourse, which is very similar to everyday language. While specialized medical language has a fixed syntactic structure, journalism language is more variable. As we will see in the next chapter, neural machine translation is based on probabilistic estimations, which turns to be really effective when translating sequences belonging to specialized texts, but they give worst outcomes when translating popular article. Obviously a precise syntactic pattern, with few subordinate clauses and a strict order is easier to translate for a machine. Reordering clauses is something that machine has recently started doing, thanks to the implementation of new neural architectures, but it is not an easy task. Plus, for many years machine translation have been trained on specific subject for some reasons in our opinion: first because before technical corpora were easier to find; then, restricting the training set to a particular matter could help investigators to understand better the functioning of machine translation models; finally, as we have seen above, there has always been a close connection between specialized discourse and computer science.

In this chapter we have outlined the history of the evolution of English medical language and how it has been influenced by several languages. Its origins can be traced back to ancient Greek up to the contributions of the Normans. Then, we wanted to dedicate a part to the discipline of terminology and specialized language, describing its birth and its developments and its link with machine translation. In any case we leave

open the question of whether machine translations have termbase or not. We will go into this in more detail in the following chapters.

2. THE TURN OF MACHINE TRANSLATION: FROM STATISTICAL TO NEURAL MACHINE TRANSLATION

2.1 The origins of Machine Translation

Warren Weaver is one of the pioneers of Machine Translation, who mentioned the possibility of computers to translate documents between different languages. Weaver, an American mathematician, in the Translation Memorandum proposes four ideas to go beyond the limits of a word-for-word approach: a turning point in the history of Translation Studies. Interested in the science of cryptography, he realized it was possible to decode a message in an unknown language. Weaver supposed that "there are certain invariant properties, which are, again not precisely but to some useful degree, common to all language" (Weaver 1949): there is the idea of the linguistic universals, a pattern that occurs systematically across natural language. This thought, together with the position of Weaver as director of the Applied Mathematics Panel, led him to consider using the computer to translate texts written in different natural language. Besides, Weaver was also concerned about the problem of incommunicability between peoples speaking different languages, which prompted him to write to Norbert Wiener of the Massachusetts Institute of Technology. In this letter Weaver suggested the idea of designing a computer which would translate different languages, "even if it would translate only scientific material […] and even if it did produce an inelegant (but intelligible) result" (Weaver 1949).

After a discouraging response from Wiener, Weaver, thanks to the help of others scholars, tried to overcome the word-for-word translation and the problem of multiple meanings taking into examination the immediate context of a word. He considered that statistical semantic could be very helpful in creating a machine able to translate from one language to another, with the contribution of cryptography. Of course, he was aware that the results could not be perfect and anticipated, in a certain sense, the need of post-editing during the process of translation. Also the work of Claude Shannon, of the Bell

Telephone Laboratories, was considered of huge utility in theorizing a machine which could translate a text. In fact, the paper written by Shannon *A Mathematical Theory of Communication* (Shannon 1948), represents a landmark in the Statistical Machine Translation. In this paper it can be traced the idea of n-gram language model: indeed the main question that Shannon poses is "given a sequence of letters, what is the likelihood of the next letter?". This interrogative can be extended to words or even sentences. According to Shannon the answer to this question can be derived from the probability distribution, a mathematical function that provides the probabilities of occurrence of different possible outcomes in a given experiment. We could easily say that Shannon lays the foundations of the field of natural language processing. Within this field fall machine translation and machine learning. Natural language processing is the task of predicting the next word in a text giving the previous word.

2.2 Natural language processing

Natural language processing is a field of machine learning which enables machines to read, understand, and derive meaning from human languages. In other words natural language processing investigates how machine handle natural human language text or speech. The idea behind this subject is imitating how human brains understand and process language and meaning. One of the first and most famous program which manages to fulfills such tasks was the STUDENT program developed by Daniel Bobrow in 1964. As a human brain, NLP needs a lexicon of the language, a parser, and grammar rules to break sentences into representations. The system of course needs also theory from semantics to train the machine to understand meaning. In order to reach this goal semantic parsers are used to convert natural language texts into formal meaning representations. To train a NLP model it is necessary to represent a given word: to do that there exist different manners. One of these is the distributed word representations. This kind of representation allows to represent words as feature vectors – a vector which contains information describing an object's important characteristics. In this way they can reveal semantic or syntactic dependencies and the vector encodes many linguistic regularities and patterns (Mikolov et al. 2013). This method takes up the distributional hypothesis (Salhgren 2008) which assumes that there is a correlation

between distributional similarity and meaning similarity, something that allows us to utilize the former to predict the latter. In other words "words that are similar in meaning occur in similar contexts" (Rubenstein and Goodenough 1965). But what is intended with meaning in this case? Well, according to Sahlgren, who refers to the Saussurian structuralist linguistic, to understand the meaning of a word the only viable route is investigating the syntagmatic and paradigmatic relations between words. Before explaining these concepts we need to introduce the notion of meaning theorized by Saussure. According to the Swiss linguist, linguistic meaning does not exist in a vacuum, it is not an entity that exists independently. Rather, they are dependent on other linguistic signs within their language system to determine what they are. So, meaning can be understood by what it is not, by its differences to other linguistic signs. As Saussure puts it "Concepts are purely differential and defined not by their positive content but negatively by their relations with the other terms of the system. Their most precise characteristic is in being what they other are not" (Saussure 1916). Relations between linguistic sign can be syntagmatic and paradigmatic. As for the former, words, or rather signifiers, are similar to chunks chained together in a sequence. They follow a linear nature. Therefore, syntagmatic relations concern positioning, and relate entities that co-occur in a text (Sahlgren 2008). Paradigmatic relations instead concern substitution, and connect entities that do not co-occur in the text through a relation in absentia. They are substitutional relations, which means that linguistic entities have a paradigmatic relation when the choice of one excludes the choice of another (Sahlgren 2008). A paradigm is thus a set of substitutable entities. Syntagmatic and paradigmatic relations are represented as orthogonal axes in a 2-dimensional space.

If we shift this notion of meaning in the field of NLP and the distributed representation, we assume that the Saussurian notion of meaning is incredibly useful in the NLP field, especially as for the distributed representation of words since it consist of the essential and more suitable features to define meaning in such domain of study. Indeed, the concept of meaning based on differential relations with a sentence's other meanings permits to represent it in an effective and fit manner for the purpose of NLP, which need a mathematical writing to train an algorithm to think as a human brain.

Thinking of meaning in an abstract and metaphysical way would not have made all this possible, since calculators treat words as atomic and discrete signs.

2.3 Statistical Machine Translation

As we have seen, the history of machine translation goes back over 60 years. The main idea behind statistical machine translation derives from information theory. This theory was developed by Claude Shannon in his paper *A Mathematical Theory of Communication* in which information means a set of possible messages that must be sent over a noisy channel to a receiver whose task is reconstructing the message with low probability of error in spite of this channel. This is one of the most important paper which theorizes the huge benefit that probability theory could generate in statistical machine translation. Even if the idea could appear difficult to understand, especially without a mathematical background, we have to admit that probabilities are used daily "when we have to deal with events with uncertain outcomes" (Koehn 2010). A probability distribution is a function that maps possible outcomes to values between 0 and 1. In the field of translation this concept can be utilized to predict a possible target word equivalent to a source word, that is probability distribution. Declined in different functions, probability distribution provides a set of methods that allows to calculate more complex distributions. One of the major contribution that led to the application of probabilistic models in machine translation theory is distributional semantics. This is a research area according to which semantic similarity between linguistic items can be inferred by their distributional hypothesis in large samples of language data. In other word, to cite the famous quote of Firth, "you shall know a word by the company it keeps". It suggests that more semantically similar two words are, the more distributionally similar they will be, that is they will tend to occur in similar linguistic contexts. We must open a parenthesis on the concept of model, or rather of language model. Language model measures the fluency of the output, affecting word order, choice, and other decisions. From the mathematical point of view, they assign each sentence a probability indicating how likely a sentence is to occur in a text. There are several language model. One of the most utilized statistical language model is the n-gram model which is a contiguous sequence of n items from a given sample of text. The

n-gram model can be used to estimate the probability of the n-word of an n-gram given the previous n-1. In other words, it is possible to calculate the value representing a given word based on the word that precedes it. This reminds us the Saussurian notion of meaning above mentioned. There exist different methods to estimate a lexical translation probability distribution. One of these is the maximum likelihood estimation, a function that given a source text word returns a probability, for each choice of the target text, that indicates how likely the translation is (Koehn 2010). However, this method turns out to be computationally expensive. An effective way to estimate a probable translation comes from the Markov property, which holds that only a limited number of previous words affect the probability of the next word. Nowadays it is proved that this assumption is wrong, however limited data restrict the collection of reliable statistics to short histories – that is the collection of previous words.

The initial statistical models for machine translation are based on words as atomic units that may be translated, inserted, dropped, and reordered (Koehn 2010). Originally, according to statistical machine translation methods language was broken up into sentences, which are strings of words. Distribution words in language does not follow a systematic pattern, that is that some occur more frequently than others. Categorization can be done depending on part of speech or words meaning. Plus, language are not the same from the morphological point of view – they can belong to synthetic or analytical typology. A pivotal concept in statistical machine translation is parallel corpora, text collections paired with a translation into another language. In other words, source and target texts are aligned in order to proceed with the translation. There are different kinds of alignment, word or sentence alignment. One of the major problem of this technique is that not every words have a corresponding equivalent in another language. For instance, the auxiliary verb “do” does not have a translation in Italian, so it cannot be aligned.

Since, according to some linguistic theories, language has a hierarchical structure of sentence, clause, phrases and words, there are different types of statistical machine translation. In word-based models the task of word alignment reveals itself to be very complicate – for reasons that we have explained above – however is a fundamental task to be executed. In this model words represents the unit of translation. The model stems from the work on statistical machine translation by the IBM Candide project in the late

1980's. In word-based statistical translation the outcome is produced trough the lexical translation probability distribution, that is function which estimate the most likely translation for a source word. The most used type of estimation is the maximum likelihood estimation, that maximizes the likelihood of the data. Lexical translation probabilities, together with the alignment function, are the basis of the IBM model, that is a generative model – it breaks up the process of generating the data into smaller steps, modeling them with probability distributions, and combining the steps into a coherent story (Koehn 2009). A step further in statistical machine translation is represented by phrase-based models, where the unit of translation is phrase indeed. The purpose is overcoming the limitations of word-based models – such as polysemy or synonymousness – translating whole sequence of word of different length. In phrase-based model one source sentence is segmented into phrases. Then, each phrase is translated into the target sentence and finally phrases are reordered, since word order is not the same in all languages. The distance-based reordering model is the method by which reordering is handled. When we talk about phrase in phrase-based model, we do not refer to the syntactical concept of phrase, but to a grouping of words, whose number can variate according to necessities. The fundamental data structure in this model is a table of phrase pairs with associated scores which may come from probability distribution. To acquire these table there are different ways. One could be aligning words between each sentence pairs of the parallel corpus and then extract phrase pairs that are consistent with this word alignment. When we talk about consistency, we mean that phrases are mapped one-to-one. Finally source sequences are decoded into target sequences. Phrase-based model turn out to be a huge advancement in statistical machine translation, in that translating word groups instead of single words avoid ambiguities. Plus, as the training corpora grows larger, we can learn longer and longer useful phrases and translations can be memorized so that outcomes become increasingly accurate.

2.4 Neural machine translation

Over the past few decades there was the emergence and development of neural machine translation (NMT), which pushed the performance of machine translationn to new heights, modeling entire sequences in a single integrated model (Cheng 2019). This

represents one of the major difference with SMT, which uses, as we saw above, separately engineered subcomponents. SMT's main constraints are represented by the fact that the history of a given word to translate – that is the number of the previous words – augments with increasing context. But, in order to obtain a training set robust enough to estimate a meaning's word according to the context, a huge amount of data would be needed. Something that would costs too much from a computational point of view. The tremendous development of artificial intelligence and deep learning led scholars to propose a new paradigm of machine translation. Thanks to models that we are going to explain soon, NMT consider the context, that is the entire sequence of words, in order to translate (Wolk et al. 2015).

The main departure from SMT is the use of word embeddings, that is vector representation for words. Basically, words or phrases are mapped to vectors of real numbers. These vectors are in the so called vector space, which ideally represents the text context. In linguistics word embedding are one of the main topic of distributional semantics. Indeed word embeddings are based on the idea that contextual information alone constitutes a viable representation of linguistics items. This idea takes up the structuralist linguistic notions developed by Ludwig Wittgenestein and John Firth. It is a concept developed in the field of distributional semantics, whose one of the underpinning notion is that the meaning of a word can be inferred by the their occurrence in the same context. In other words "word co-occurrence statistics extracted from text corpora can provide a basis for semantic representations" (Lenci 2015). Therefore the main difference with SMT is that words are represented as vectors, that thanks to distributional semantics characteristics enable to including semantic features into a unique entity. As we have seen above, meaning has proved to be one of the greatest challenges in artificial intelligence. Through word embeddings, neural networks manage to learn concept representations directly from data, without human intervention – such as language model, reordering, or alignment – taking into account the context of a sentence (Skelac and Jandrić 2020). When we say that this concept takes up the work of philosophers of language like Wittgenstein, we refers to his belief that "only the preposition has sense; only in the context of a proposition has a name meaning" (Wittgenstein 1921). So, words have no meaning in isolation. Wittgenstein was

influenced by Frege's work, according to which "only in a proposition have the words meaning". Frege maintains that since a word standing alone can refer to multiple entities, have multiple meaning, its reference cannot be determined outside a specific context. In general linguistics, the importance of context for establishing meaning is developed by J.R. Firth, who holds that the meaning of a word is always contextual (Firth 1957). In the late 1920s. Wittgenstein's view on the context of use of words changed dramatically. In his new understanding, words have a definable meaning only within a system of propositions independent from the rest of the language. He described his intuitions in the its late period, when he compared the word as a tool – "One cannot guess how a word functions. One has to look at its use, and learn from that" or "The meaning of a word is in its use in the language" (Wittgenstein 1953). Language consists of small and overlapping systems, each constituted by its own rules that prescribe the use of its terms. It can be thought as a specific language domain system, which has its terminology and suitable expressions. In a certain sense we could consider the Austrian philosopher as the pioneer of terminology and specialized language. Indeed, with the introduction of the notion of language games, Wittgenstein lays the foundations of the terminology, in that he considers them as "forms of language, complete in themselves, but easily imagined as evolving in new and more complex ones" (Skelac and Jandrić 2020). The theoretical contribution given by Wittgenstein in word embedding is precisely this, thinking of language as a multiple set of configurations of languages shared by a community and rule-governed.

Firth takes up Saussurian notion of sign, expanding it – according to him signs are yes dependent from system and their meaning changes with the context in which they are used (Skelac and Jandrić 2020). The pivotal concept of Firth is the conceptual situation which involves the participants, the object of the discourse, and the effects caused by a given exchange. Besides, Firth investigates the use of collocations - "the mere word accompaniment, the other word material in which the word is embedded" (Firth 1952). The meaning of a given word can be inferred by the words close to it. This Firthian notion is fundamental in the word embedding models, since a given model maps vectors with words' meaning taking into account their collocations.

From a technical point of view, word embedding turn to be very effective in the building of neural models. One of the most famous neural model is the one developed by Mikolov (Mikolov et al. 2013) Word2vec, constituted by two architecture neural models: Continuous Bag-of-Words and Continuous Skip-Gram. The great advantage of this model is that has fewer computational costs than SMT and it allows the model to consider larger contexts. The architecture Continuous Bag-of-Words maximized a given word's probability analyzing its context. It assures better quality translation for frequent words. The architecture Continuous Skip-Gram works basically works the other way around: on the basis of a given word, it predicts the word's context. It turns to be very effective in the translation of rare words and sentences. Taking up the example that Skelac & Jandrić do in their paper, we can see that the word "beer" and "wine" are more conceptually similar than the couple of word "beer" and "cat", because they belong to the same context. This is the idea behind Word2vec and it is basically the pivotal principle of the distributional hypothesis above mentioned. Word2vec manages to recognize such similarities, because its main aim is that of creating a system that emulates human brain. So, obviously this model, analyzing the words of the example above, would give similar values for corresponding vectors of the words "beer" and "wine", and a totally different corresponding vector for the word "cat".

Of course Wittgenstein and Firth's theories are not conceptually identical to Word2vec. For instance, Word2vec offers a restricted view of what context is and what its importance is for communication effectiveness. In Word2vec context is restricted to to neighboring words only. Neighboring longer sequences are not considered. Plus, the Wittgensteinian referential meaning, according to which there are some words that pick out something in the extra-linguistic reality, is something that eludes Word2vec. Vectors can only track collocations and analyze inter-linguistic connections between words. Besides, we have to keep in mind that for Wittgenstein words have meaning within a context, but contrary to Firth he does not identify context with meaning. So, a suitable mathematician representation of this concept would be a function that to every language game (in which the word is used) ascribes the meaning the word has in that particular game: a set of rules governing the use of the word in game, or better in context (Skelac and Jandrić 2020). Instead, Word2vec does not distinguishes referential meaning from

meaning in context, its vectors do not map contexts in which the word is used to meanings the word has in each of them, but merge them. So a vector has its particular and distinguished meaning in a specific vector space, but if the same word finds itself in another context, and so in another vector space, would have a different value.

Of course, these are nuances that we do not except neural models are able to recognize and we maintain that developments reached so far by NLP and neural models are of inestimable value, even if they still cannot investigate deeper the issues relating to meaning.

2.5 DeepL and Google Translate architectures

Google was one of the first company to launch its own machine translation service. One of the first release was based on statistical phrase-based machine translation. While at first it seemed a useless tool which translated overly literally short sentences giving non-sense – often hilarious – outcomes, over the years it incredibly improved. In 2016 Google announced the Google Neural Machine Translation (GNMT) system, which utilizes advanced techniques in order to achieve outstanding improvements for machine translation quality (Wu et al. 2016). GNMT overcomes some of the restrictions of phrase-based machine translation, such as expensive computational costs both in training and in translation inference. Neural machines translation represent a great improvement in machine translation because they are able to learn directly the mapping from input text to associated output text. Its architecture consists of two recurrent neural networks (RNNs), one to be employed in the input sequence and one to generate the translated output text. In other words, the network encodes the source sequence as a list of vectors, where each vector represents the meaning of all the words read so far. Then, the decoder generates each target sequence one word at a time. It translates one word at a time in a strict left-to-right or right-to-left order. In order to do so, the decoder "pays attention" to a weighted distribution over the encoded source vectors most relevant to generate the target words – that is, through a probabilistic estimation the decoder manages to understand what are the source text's words most useful to predict the meaning of their neighbors so that it can generates a target output. This is the renowned attention mechanism, which I am going to discuss further in the

final chapter. Thanks to this model the neural machine translation can have an overview of the entire sequence. One of the other advantages of the attention mechanism is that it reuses the hidden layers corresponding to the source sequence, whose weighted media is used as input together with the compressed vector representation. It means that thanks to attention mechanism the model can go back and see again some strings of sequence that can be useful to understand better some context nuances (Vaswani et al. 2017). Another characteristic of RNNs are long short-term memories (LSTMs) – given a sequence of words they predict the probability of each word given the previous. A mechanism hugely enhanced by the attention mechanism (Pascale).

Despite DeepL keeps jealously secrets its architectures and algorithms, it is thought that one of its strengths is convolutional neural networks (CNNs). Unlike RNNs, CNNs can process information hierarchically, a method that allows them to look for non-linear relationships in data. This means that a CNN can easily grasp contextually meaning and translate it accordingly (Gehring et al. 2017). Another great model that probably DeepL utilizes is the multi-hop attention capability (Gehring et al. 2017; Iida et al. 2019) developed by Facebook engineering. This model mimics the way humans translate sentences, returning to them multiple times to double-check meaning. Multi-hop translation mechanism allows networks to look repeatedly at the sentence and make choices about what to translate first – a first look might point CNN to a verb; a second look to an associated subject (Pascale).

It almost seems that human translator are destined to disappear, however meaning nuances still are not understand by machine translation applications. They can capture the context of the words meaning through probabilistic estimations, but unlike humans they cannot really know the linguistic, pragmatic, and semantic reasons behind each word's choice.

3. A COMPARATIVE ERROR ANALYSIS BETWEEN DEEPL AND GOOGLE TRANSLATE

3.1 Translation quality assessment

Translation Quality Assessment (TQA) has been a much debated subject in translation studies and in particularly in machine translation (MT). Indeed, the exponential evolution of machine translation has resulted in the need for a standardization of the evaluation metrics for a MT outcome. This represents a difficult task, since there are a lot of variables coming into play when we are talking about TQA. Dissimilarity can be given whether an evaluation takes place as a part of a production process or a research study (Castilho et al. 2018). As for research, the aim is to get a measure that "can show a demonstrable change in quality, most usually an improvement, from previous work or between different translation processes" (Castilho et al. 2018). This represents the reason behind our machine services comparison, by which we are trying to understand how far has developed and improved neural machine models and the differences existing between Google Translate (GT) and DeepL. However, it is difficult to separate industry and research, since they affect themselves mutually. As a matter of fact, quality assessment research focus primarily on the underlying translation theory that one holds with, so that "different views of translation lead to different concepts of translation quality, and hence different ways of assessing it" (House 2015). The equivalence, Skopos, the functionalist and the perspective approach diverge from each other, and the pursue of a unique method of evaluation in this field seems yet difficult to find.

As for the TQA in translation industry, it is predominately the error-based one, where errors found in the target text are classified and weighted according to their severity. There are several error typology to assess translation projects: one of these is the Localisation Industry Standard Association (LISA) QA Model, containing a list of types of errors categorised as 'minor', 'major', or 'critical'. However, it faced important limitations. First, low inter-annotator agreement means that reviewers are not

interchangeable: two reviewers might disagree as to how important an error was, or even if it was an error. Besides, models that have been developed with specific text types in mind might not be appropriated for different text types (Lommel 2018). Another error-based assessing model is the one developed by the Translation Automation User Society (TAUS), a translation industry think tank, the Dynamic Quality Framework (DQF). It is a collection of several approaches to the issue of assessing an MT outcome. The first version of the DQF Error Typology had six main error types:

- Accuracy, problems related to the transfer of meaning from source to target text.
- Linguistic, problems related to the language of the target text. The idea behind this metric resembles Fluency.
- Terminology, problems related to the use of domain approved glossary.
- Style, problems related to the text type specific style.
- Layout, problems related to non-textual aspects of the content, such as formatting and length.

Then, there is another group who works on TQA, the EU-funded QTLaunchPad project, led by the German Research Centre for Artificial intelligence (DFKI). It developed the Multidimensional Quality Metrics (MQM), which takes up principles and ideas previously promoted by LISA. In a certain sense MQM adopts the functionalist approach that quality is defined by how well a text meets its communicative purpose. MQM provides a set of metrics to assess a translated content and the output of translation systems. It does not assess the translation processes or projects. With "translated content" is intended the text, the graphics, and any other content which may be translated or adapted for multiple locales. The most common metrics that can be faced in TQA of translated texts are

- Accuracy
- Fluency
- Design
- Style
- Locale Convention
- Style

- Terminology
- Verity

When we talk about accuracy and fluency, we refer to two of the most important evaluation metrics together with terminology. Accuracy, also called fidelity or adequacy, is defined as the extent to which the translation transfers the meaning of the source-language unit into the target. Fluency focuses on the target text and is defined as the extent to which the translation follows the rules and the norms of the target-language, regardless of the source text (Casthilo et al. 2018). These are two metrics that we are going to follow analyzing our articles.

As for the assessment framework which we are going to use, we have opted for the DQF/MQM integrated error typology. So, it has less features than the MQM hierarchy and is focused more tightly on terminology and style. Thanks to its smaller size is easier to understand and utilize.

Another reference point precious for our assessment is the work of Maja Popović, whose paper has helped us to understand better and set our texts error-evaluation (Popović 2018). We have decided to perform it manually, focusing on some DQF/MQM metrics, that is adequacy, fluency and overall quality and terminology. The purpose of the assessment is to see which types of error DeepL and GT carry out, how they affect the overall translation and, especially in the specialized texts, if the models are able to recognize and translate technical term. Then, by the comparison of the outcomes of the two neural machine translation we will see which one performs better and why. Some of the metrics that we will use more are terminology, mistranslation and word order errors. Terminology and mistranslation are two similar concepts, but they are actually different. In fact mistranslation refers to the erroneous translation of a general noun, verb, adverb, and so on, while a terminology error has to do with the wrong rendering of a specific term pertaining in this case to the medical domain.

We have chosen several popular and specialized texts whose macro-area of interest is immunology and sexually transmitted diseases. With regard to the specialized texts, we analyze translations of medical papers' abstracts. In order to narrow these subject matter, we have decided to focus on three illness: human papillomavirus

infection (HPV), syphilis and gonorrhea. We present an average of six articles for each disease, among these three are popular articles and three are specialized articles.

We have chosen three abstracts' article on immunization focusing mainly on HPV, syphilis and gonorrhea. The abstract have been searched on the free search engine Pubmed. As for the popular texts, the material comprises three newspapers' article (New York Time, The Guardian, and the Washington Post). We have decided to assess specialized and popular articles to represent different text types and see if adequacy varies depending on text types. As stated above, the criteria normally used to analyze the quality of a translation are adequacy, fluency, and overall quality (Popović 2018). We will assess our texts taking as reference point the general error taxonomy draught by M. Popović as well as the DQF/MQM integrated error typology (Lommel 2018). The latter is particularly useful for our aim since it promote two category which we will focus on, that is terminology and style; the former allow us to study if machine translation possess a sufficient and appropriate terminological database to translate specialized medical texts. The latter will let us to observe how the machine translations' outputs change according to different registers.

3.2 Specialized medical texts' analysis

If we look at the specialized abstract about gonorrhea, HPV and syphilis translated by DeepL, we can notice that syntax is generally respected. However we see some awkward outputs like in the article "Effectiveness of a group B outer membrane vescicle meningcoccal vaccine against gonorrhea in New Zealand: a retrospective case-control study".

Let us begin from the abstract title: in the DeepL's output we observe a major omission.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
Effectiveness of a group B outer membrane vesicle meningococcal vaccine against gonorrhoea in New Zealand: a retrospective	Efficacia di un vaccino contro la gonorrea in Nuova Zelanda: uno studio retrospettivo caso-controllo.	Efficacia di un vaccino meningococcico sulla vescicola della membrana esterna di gruppo B contro la gonorrea in Nuova

case-control study.		Zelanda: uno studio retrospettivo caso-controllo.

Nonetheless, the general meaning of the source text's title is preserved and conveyed in an effective manner to the reader even if the technical definition of the vaccine is not translated. At the other hand, Google's translation is more literal and translates all the title, but the message is a bit confused. Besides fails to translate the technical name of the vaccine, as we will see soon in the beginning of the abstract.

SOURCE TEXT	DEEPL
Effective vaccine development has been unsuccessful, but surveillance data suggest that outer membrane vesicle meningococcal group B vaccines affect the incidence of gonorrhoea.	Lo sviluppo efficace del vaccino non ha avuto successo, ma i dati di sorveglianza suggeriscono che i vaccini di meningococco di gruppo B della membrana esterna vescicola vescicola meningococco influenzano l'incidenza della gonorrea.

In the DeepL's output we notice how the Italian word order is not respected, so that the real meaning of the sentence is totally compromised: in fact, we would expected to read "lo sviluppo di un vaccino efficace" because the position of "efficace" next to "sviluppo" ends to connote this substantive, making the sentence meaningless. The next clause shows an unusual word order as well: if we shift our focus to lexis level, we can see a needless addition of the noun "vescicola" and the output is totally meaningless due to an incorrect word order because the machine translation fails to recognize some specific terms such as "outer membrane vesicle meningococcal group B vaccines" that should be translated with "vaccini contenenti vescicole della membrana esterna del meningococco B" (Gabutti, G. 2014). In fact, this terminological expression is compounded by many medical terms: "outer membrane vesicle", for example, is one of the term of this expression, often written in technical essay with the acronym OMV.

If instead we look at the same example translated by Google Translate, we can observe that, even if it is still incorrect nonetheless it apparently seems to have an appropriate syntactic structure:

SOURCE TEXT	GOOGLE TRANSLATE
Effective vaccine development has been unsuccessful, but surveillance data suggest that outer membrane vesicle meningococcal group B vaccines affect the incidence of gonorrhoea.	Lo sviluppo efficace del vaccino non ha avuto esito positivo, ma i dati di sorveglianza suggeriscono che i vaccini del gruppo B meningococcico della vescicola della membrana esterna influenzano l'incidenza della gonorrea.

First of all, in the Google Translate's translation we do not have the useless repetition of "vescicola"; furthermore it recognizes "meningococcal" as attribute of vaccine, which DeepL does not. Google Translate grasps other terminological nuances: for example the verb phrase "has been unsuccessful" in this context assume a terminological value that GT conveys with "non ha avuto esito positivo": here, "esito positivo" is in fact a term that connote the entire expression in a terminological manner. Conversely, DeepL translates quite literally this terminological expression with "non ha avuto successo".
So, if we can notice an apparently respect of the fluency metrics, looking more closely at the translations, we can observe how some terminological nuances are better rendered by Google Translate. In fact, in the next sentences as well Google Translate expresses into Italian the meaning of some verb phrase through the use of what we consider terms in this specialized context.

SOURCE TEXT	DEEPL TRANSLATION	GOOGLE TRANSLATION
We did a retrospective case-control study of patients at sexual health clinics aged 15-30 years who were born between Jan 1, 1984, and	Abbiamo fatto uno studio retrospettivo caso-controllo di pazienti in cliniche di salute sessuale di 15-30 anni che sono nati tra il 1	Abbiamo condotto uno studio retrospettivo caso-controllo su pazienti di cliniche di salute sessuale di età compresa tra 15 e 30

Dec 31, 1998, eligible to receive MeNZB, and diagnosed with gonorrhoea or chlamydia, or both.	gennaio 1984, e il 31 dicembre 1998, idonei a ricevere MeNZB, e con diagnosi di gonorrea o clamidia, o entrambi.	anni nati tra il 1° gennaio 1984 e il 31 dicembre 1998, idonei a ricevere MeNZB e con diagnosi di gonorrea o clamidia, o entrambi.

Here, the verb "to do" in the clause "we did a retrospective case-control study [...]" is translated as "abbiamo fatto" by DeepL and as "abbiamo condotto" by Google Translate: while the first choice is more generic and does not distance itself from the literal and more common sense of the verb "to do", the second one utilizes a verb more specific and suitable in the medical context. In fact, for "condurre uno studio" there are 1.510 results on texts pertaining to specialized technical domain on Google Scholar, while for "fare uno studio" there are 1.340 results and the most of these are in specialized, however pertaining to literary domain.

As to regard to medical specialized terms, both machine translation fail in translating them. For example both DeepL and Google Translate place the attribute "retrospettivo" in the wrong position in the phrase "studio retrospettivo caso-controllo": in fact, the correct term in Italian is "studio caso-controllo retrospettivo", where the attribute "retrospective" could be omitted since case-control study are usually retrospective. Besides, inserting an attribute between "studio" and "caso-controllo" would not be possible because creates a discontinuity right within the term. So, the only accepted solution is collocating the attribute after the term. However this represents a minimal mistake since if we google it we can find pertinent information whether it is written one way or the other. For instance on the Eupati website, a European project aimed to teach patients to understand the research and development on drugs, we find "studio retrospettivo caso-controllo", but being a European project it could have been translated using a machine translation. Plus, this website is intended to popularize medical and pharmacological terms, so a colloquial or inaccurate style can be accepted. Of course, all of this opens up an interesting and unsolved issue about terminology and machine translation, that is, how strict are the principles governing specialized languages and terminology and in which way are they changing due to the machine translations' improvement. In fact, some terms translated by machine translation can spread out and

become an accepted term by the community. We will see how all the specialized texts we have chosen, even if inaccurate, are comprehensible.

Another expression that machine translation fails to translate is “sexual health clinics”. Here the problem is that there is not an exact equivalent place in Italy such as sexual health clinics probably due to the fact the in New Zealand there is a different healthcare system. A suitable equivalent in Italian could be “dipartimento di venerologia”, which is the Italian name discipline that studies issues such as sexually transmitted diseases.

In this sentence the vaccine’s name is written with an acronym “MeNZB”, where NZ stands for New Zeland, the state where the vaccine has been developed and “MenB” for meningococcal disease caused by serogroup B. The problem with the acronym in this case is that usually in Italian one should writes the substantive “vaccino” before it, so the correct translation in our opinion should be “abbiamo condotto uno studio caso-controllo retrospettivo su pazienti […] eleggibili per il vaccino MeNZB”. Eligible here should be translated as “eleggibile” as we can see in almost 1.500 results on Google News. Here we would change also “idonei a ricevere”, used by both translation machine: in fact “eleggibile” is the correct term in medical language as we can observe on other Italian specialized medical text.

The next sentence has some terminology errors:

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
Demographic data, sexual health clinic data, and National Immunisation Register data were linked via patients' unique personal identifier.	I dati demografici, i dati della clinica di salute sessuale e i dati del Registro Nazionale delle Immunizzazioni sono stati collegati attraverso l'identificatore personale unico dei pazienti.	I dati demografici, i dati delle cliniche di salute sessuale e i dati del registro nazionale delle vaccinazioni sono stati collegati tramite l'identificatore personale univoco dei pazienti.

First of all we have in both outputs the expression “dati demografici” that is a literal translation of what should be “indici demografici”. As we said in the previous chapters, a specialized medical text language pertains to different domains, having not

only medical terms but also those of other disciplines, such as statistics. So we consider “demographic data” a terminology error.

We have another terminological error with “National Immunization Register”, which both machine translation translates literally as “Registro Nazionale delle Immunizzazioni” (DeepL) and “registro nazionale delle vaccinazioni” (Google Translate); the only difference is the rendering of “immunization”. However, the Italian equivalent should be, in our opinion, “anagrafe nazionale vaccini” without capitalizing them like DeepL does. Both registers record vaccinations given to people of all ages in the country. In this case, the Italian term “anagrafe” is more appropriate than “registro” since it is the one more used in public administration. Actually, the greekism “anagrafe” suit better given that derives from *ἀναγράφω* that means “write upon”. Instead, “registro” comes from Old French *registre* and from Medieval Latin *registrum,* alteration of Late Latin *regesta*, neuter pluaral of *regestus*, past participle of *regerere* that means literally “to bring back (information)”. So, seen the origin of the noun, we consider that “anagrafe” is the correct equivalent term of “register” in this case, despite the noun “registro” shows a similar graphic and phonetic form.

A further terminological error is the rendering of “patients’ unique personal identifier”. Again, both machine translations fails to translate the term, rendering it literally: DeepL translates it with “identificatore personale unico dei pazienti” and Google Translate with “identificatore personale univoco dei pazienti”. Google Translate’s output creates a bit of confusion given that in Italian there is the term “codice identificativo univoco”, that is a drug track trail number necessary to avoid risks of fraud and counterfeiting. Actually, the best equivalent of “patients’ unique personal identifier” would be “tessera sanitaria”, a card that grants the holder the right to obtain health services throughout the European Union.

Let us look at the last part of the abstract’s background

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
For primary analysis, cases were confirmed by laboratory isolation or detection of Neisseria gonorrhoeae only from a	Per l'analisi primaria, i casi sono stati confermati dall'isolamento di laboratorio o dal rilevamento di Neisseria	Per l'analisi primaria, i casi sono stati confermati dall'isolamento di laboratorio o dal rilevamento di Neisseria

clinical specimen, and controls were individuals with a positive chlamydia test only. We estimated odds ratios (ORs) comparing disease outcomes in vaccinated versus unvaccinated participants via multivariable logistic regression. Vaccine effectiveness was calculated as 100x(1-OR).	gonorrhoeae solo da un campione clinico, e i controlli sono stati effettuati solo su individui con un test di clamidia positivo. Abbiamo stimato i rapporti di probabilità (OR) confrontando i risultati della malattia in partecipanti vaccinati con quelli non vaccinati attraverso una regressione logistica multivariabile. L'efficacia del vaccino è stata calcolata come 100×(1-OR).	gonorrhoeae solo da un campione clinico e i controlli erano soggetti con un solo test positivo per la clamidia. Abbiamo stimato gli odds ratio (OR) confrontando gli esiti della malattia nei partecipanti vaccinati rispetto a quelli non vaccinati mediante regressione logistica multivariabile. L'efficacia del vaccino è stata calcolata come 100 × (1-OR).

The outputs made by the two machine translation are very different. It seems that Google Translate manages to convey the source text's meaning in a more correct and appropriate way than DeepL, which carries out a very literal translation. At first, we have focused ourselves on the terminological expression "analisi primaria", which sounds really weird. However, after a research on Google we have found that is a medical term that express the first steps of a trial, as we can see in the report by EMA (European Medical Agency) about Flixabi, a medicine for Chron's disease.

Both outputs have the same errors in the first sentence; for example, "[…] sono stati confermati dall'isolamento di laboratorio o dal rilevemento di Neisseria gonorrhoeae [...]": the translation of "laboratory" as "di laboratorio" is not necessary since in Italian both sample processing steps are termed without the specification "di laboratorio", as we can see on some microbiology handbook (Lanciotti 2012). Plus, the machine translation make a syntactic error, failing to recognize that "laboratory" is an attribute both of "isolation" and of "detection". Besides, both software translate incorrectly "detection" with "rilevamento", while the equivalent term in this case is "identificazione". Instead, they correctly leave untranslated the term Neisseria gonorrhoeae: in fact it is the technical denomination of the bacterium that causes the gonorrhea. In English it is colloquially called the clap; there are few theories about the

origin of this expression: according to one of these, the origin of this word dates back to the Middle French noun *clapier* that means “brothel”. There are synonyms of Neisseria gonorrhoeae in Italian as well: one is gonorrea, that is phonologically similar to the technical term but it differs graphically. The other synonym is blenorragia. Both nouns come from Ancient Greek and they are compound nouns: gonorrhea is compounded by *gònos*, that means “seed”, and *reo* that means “to pour”; blenorragia is compounded by *blénna*, that means “mucus”, and *rhḗgnymi*, that means “to spurt”. This term is less common than “gonorrea”. Both nouns refer to the disease’s main symptom affecting men. We have to clarify that in Italian these nouns are used not only colloquially, but are also employed in medical contexts, so colloquial and specialized term in this case correspond.

We do not agree with the translation of “only” in “[…] only from a clinical specimen [...]”, which both machines translate with “solo”. We would have translated it with “soltanto”, because we think that in this context it is more effective and gives more emphasis to the idea of exclusiveness.

In the final part of the first sentence it seems that Google Translate manages to give a more adequate output than DeepL. First, DeepL is less fluent with regard to the punctuation, maintaining the source text’s punctuation. Besides, it seems to mistranslate “controlli” with its ordinary usage of “supervision”, not understanding that in this specific case “controlli” is a technical term that defines individuals who are not affected by the disease[1]. This error lead to an addition of “sono stati effettuati”, that DeepL presumably does because the noun “controlli” is usually co-occurs with the verb “effettuare”. It almost seems that DeepL, in order to carry out the translation of this clause, that does not make sense due to the mistranslation of the medical term “controlli”, adds the most probable verb that would occur with it. We could hypothesize that DeepL in this case exploits its training dataset to find a befitting translation of the clause. In fact, if we assessed the clause considering only its fluency, we could state that it is correct. However, it does not convey the meaning of the source text as much

1 In case-control studies, cases are individuals who are affected by the disease and controls are those who are not. These groups are compared in order to highlight their exposure to some causal attributes suspected to be linked to the disease, so as to determine a connection between the disease and the tested exposure (Lanciotti 2012).

precisely as Google Translate does. Google Translate's translation of "positive chlamydia test" fits better than the one made by DeepL: in our opinion sound better "test positivo per la clamidia" than "test di clamidia postivo".

In the last clause of this sentence there are not many translation's issues. DeepL translate "odds ratio" with "rapporti di probabilità", that it is correct. However, also the untranslated term "odds ratio" of Goolge Translate's output is fine, given that in the Italian statistical terminology this terminological expression has come into use as well. As for "multivariable logistic regression", both DeepL and Google Translate translate it with "regressione logistica multivariabile": it is not totally incorrect but the most fitting terminological expression in our opinion would be "modello di regressione logistica multivariata". (https://www.medinews.it/news,14676)

As for fluency, we can observe how Google Translate's output is not as befitting as DeepL's output, because of some slight imperfections: the use of "nei" could appear more appropriate than the use of "in" in the prepositional phrase "[…] in vaccinated versus unvaccinated participants […]", but in a specialized text's abstract it may seems more appropriate "in", maybe due to a kind of syntactic borrowing from English that represents the lingua franca of the medical domain.

Now, let us analyze the last part of the abstract.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
FINDINGS: 11 of 24 clinics nationally provided records. There were 14 730 cases and controls for analyses: 1241 incidences of gonorrhoea, 12 487 incidences of chlamydia, and 1002 incidences of co-infection. Vaccinated individuals were significantly less likely to be cases than controls (511 [41%] vs 6424 [51%];	RILEVAZIONI: 11 di 24 cliniche a livello nazionale fornito record. Ci sono stati 14 730 casi e controlli per le analisi: 1241 casi di gonorrea, 12 487 casi di clamidia e 1002 casi di coinfezione. Gli individui vaccinati avevano una probabilità significativamente inferiore di essere casi rispetto ai controlli (511 [41%] vs 6424 [51%]; corretto OR 0-	RISULTATI: 11 di 24 cliniche hanno fornito documenti nazionali. Ci sono stati 14 730 casi e controlli per le analisi: 1241 incidenze di gonorrea, 12 487 incidenti di clamidia e 1002 episodi di coinfezione. Gli individui vaccinati avevano significativamente meno probabilità di essere casi rispetto ai controlli (511 [41%] contro 6424

adjusted OR 0·69 [95% CI 0·61-0·79]; p<0·0001). Estimate vaccine effectiveness of MeNZB against gonorrhoea after adjustment for ethnicity, deprivation, geographical area, and sex was 31% (95% CI 21-39). INTERPRETATION: Exposure to MeNZB was associated with reduced rates of gonorrhoea diagnosis, the first time a vaccine has shown any protection against gonorrhoea. These results provide a proof of principle that can inform prospective vaccine development not only for gonorrhoea but also for meningococcal vaccines.	69 [95% CI 0-61-0-0-79]; p<0-0001). L'efficacia stimata del vaccino MeNZB contro la gonorrea dopo l'aggiustamento per etnia, privazione, area geografica e sesso è stata del 31% (95% CI 21-39). INTERPRETAZIONE: L'esposizione al MeNZB è stata associata a tassi ridotti di diagnosi di gonorrea, la prima volta che un vaccino ha mostrato una protezione contro la gonorrea. Questi risultati forniscono una prova di principio che può informare lo sviluppo futuro del vaccino non solo per la gonorrea ma anche per i vaccini meningococcici.	[51%]; aggiustato O · 69 [IC 95% 0 · 61-0 · 79]; p <0 · 0001). Stimare l'efficacia del vaccino di MeNZB contro la gonorrea dopo aggiustamento per etnia, privazione, area geografica e sesso è stata del 31% (IC 95% 21-39). INTERPRETAZIONE: L'esposizione a MeNZB è stata associata a tassi ridotti di diagnosi di gonorrea, la prima volta che un vaccino ha mostrato protezione contro la gonorrea. Questi risultati forniscono una prova di principio che può informare lo sviluppo di vaccini potenziali non solo per la gonorrea ma anche per i vaccini meningococcici.

We can easily see how the first clause is translated literally by both translation services: “out of” in Italian is “su”, so the correct translation should be “undici cliniche su ventiquattro”. Besides, DeepL commits a mistranslation of “incidences”, that is translated as “casi”: however, since the abstract is about case-control studies, translating “incidence” as “casi” creates a misleading effect to the reader which can cause a communication breakdown. Plus, the equivalent term of “incidence” exists in Italian, that is “incidenza”. So in this case we should talk of terminological error instead of mistranslation, two concepts whose difference is defined by blurred boundaries, as we have seen in the introduction of this chapter. Conversely, Google Translate uses the appropriate term pertaining to the epidemiological domain.

DeepL translates correctly “adjusted OR” in the next sentence as “corretto OR”, even if maybe it would be more appropriate to put the adjective after the noun, thus translating it as “OR corretto” (searching this terminological expression written in inverted commas on Google we obtained 362 results). However, it seems to be more appropriate Google Translate’s translation as “aggiustato OR”, (in fact there are about 1.380 results on Google), although here too it would have been more appropriate placing the adjective after the noun. An Italian pulmonologist conducting researches on asthma we talked to answered that in his papers he uses “OR corretto”, but he admits that “OR aggiustato” is suitable as well and that there is not a unique terminological expression. This is in accordance with what we said previously about how the Italian medical terminology is not anymore fixed by an official standardization due to the hegemony of English language in the medical domain.

In the next clause there is a mistranslation of “estimate” in Google Translate’s output: it translates it as “estimare”. This is a case of misselection, that is “a deviation in the morphological aspect of language. It can be a problem at word class-level and at verbal level”. (Costa et al. 2015) Or it could be seen as a part of speech error, very important in inflected languages, “where the big variability of the open word classes poses a difficult problem for machine translation” (Koponen 2010). Anyway this error is part of the morphological error class. (Popović 2018). Here an adjective is needed instead of a verb, for example, but Google Translate is not able to produce the correct form of a word, although the translation of the base form is correct (Koponen 2010). This is due the fact that the word “estimate” in English can be both an adjective and a verb and in this case Google Translate produces an infinitive verb in Italian, that is “estimare”. At the other hand, DeepL produces the correct form of the word, translating it as “stimata”. Google Translate’s output sounds awkward also because of the telegraphic aspect of this sentence given by the omission of the article – that is needed in Italian – in front of “aggiustamento”. With regard to the translation of “adjustment” made by both machine translation as “aggiustamento” is valid what we have said above about the adjective “adjusted”.

In the last part of the article there are few elements to analyze. In terms of fluency both outputs seems correct, however there is a punctuation error that renders them a

little bit awkward. In “Exposure to MeNZB was associated with reduced rates of gonorrhoea diagnosis, the first time a vaccine has shown any protection against gonorrhoea.” we would have substitute comma with a colon and add the verb “to be”. So, the correct translation in our opinion would be “L'esposizione al vaccino MeNZB è stata associata a tassi ridotti di diagnosi di gonorrea: è la prima volta che un vaccino ha mostrato una protezione contro la gonorrea.”. Plus, both machine translation translate equally the sentence, except for some elements such as the translation of “prospective vaccine development”: Google Translate renders it as “lo sviluppo di vaccini potenziali”, understanding that “prospective” is an adjective linked to the noun “vaccine”, but choosing an unfitting translation for it, “potenziale”, that does not represent the correct output in this case. At the other hand, DeepL sounds better in term of fluency but links erroneously “prospective” to “development”. As for the last sentence, we have dwelt upon the translated term “prova di principio”, that seemed too much literal due to the similarity with the source text term “proof of principle”. After a search on Google Scholar, we have seen that this is the correct equivalent term, together with the synonym “prova di concetto”, whose English counterpart is “proof of concept” (that is also on the terminological database IATE). However, there are few results on Google Scholar of papers containing these terms, perhaps because of the scarcity of medical papers written in Italian about this topic compared to those written in English.

Both DeepL and Google Translate translate the verb “inform” as “informa” that does not fit well in this context because it is used as a transitive verb, but the machine translation render it in its intransitive meaning. In fact, as a transitive verb “to inform” can mean “to give essence to something”, “to form”, “to inspire”, so the fitting translation, smoothing the output as much as possible, would be “[…] una prova di principio che può favorire lo sviluppo di un eventuale vaccino […]”. Of course, machine translation is not able to make distinctions between certain nuances of language yet.

The next article’s abstract is about syphilis, another STD for which doctors are still doing researches for a vaccine. As a whole both outputs seems respect the metrics of accuracy, however there are some fluency mistake to analyze.

Let us look at the title of the abstract:

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
Current status of syphilis vaccine development: need, challenges, prospects.	Stato attuale dello sviluppo del vaccino contro la sifilide: necessità, sfide, prospettive.	Stato attuale dello sviluppo del vaccino contro la sifilide: necessità, sfide, prospettive.

The outputs are identical and there are not mistakes that we need to highlight. The translations are quite literal, but being a technical paper's title it can be accepted as correct also in Italian that is less suitable to a telegraphic style than English.

The first sentence shows some terminological errors:

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
Syphilis is a multistage disease caused by the invasive spirochete Treponema pallidum subsp. pallidum. Despite inexpensive and effective antibiotic therapy, syphilis remains a prevalent disease in developing countries and has re-emerged as a public health threat in developed nations.	La sifilide è una malattia multistadio causata dalla spirocheta invasiva Treponema pallidum subsp. pallidum. Nonostante una terapia antibiotica poco costosa ed efficace, la sifilide rimane una malattia prevalente nei paesi in via di sviluppo ed è riemersa come una minaccia per la salute pubblica nei paesi sviluppati.	La sifilide è una malattia a più stadi causata dallo spirochete invasivo Treponema pallidum subsp. pallidum. Nonostante la terapia antibiotica economica ed efficace, la sifilide rimane una malattia prevalente nei paesi in via di sviluppo ed è riemersa come una minaccia per la salute pubblica nei paesi sviluppati.

The term "multistage disease" is translated by DeepL as "malattia multistadio", that is not used in the Italian medical language. In fact, "multistadio" is a literal translation that has not an only word equivalent in Italian. The term "multistadio" in Italian it is an adjective used in the engineering domain. It is more correct the translation made by Google Translate, "una malattia a più stadi", that uses more words to render the source text adjective qualifying the noun "disease". Reading several article on Google about this disease, we have found that the best translation would be "si verifica in diversi stadi" or "si sviluppa in diversi stadi". Nonetheless, the translation made by DeepL could fill the gap existing for this term, borrowing a term belonging to the electronic

and engineering domain. We propose also “multifase” as suitable equivalent term of “multistage”.

Another terminological problem is represented by the denomination of the bacterium causing syphilis, that is “invasive spirochete Treponema pallidum subsp. Pallidum”: it is a technical name derived from Latin that both DeepL and Google Translate fail to translate it accurately. The correct term in Italian is “treponema pallidum sottospecie pallidum della spirocheta invasiva” (Goldman-Cecil Medicina Interna). DeepL trnaslates it as “spirocheta invasiva Treponema pallidum subsp. Pallidum” keeping the source text order so that it results pretty literal. Google Translate does quite the same and fails to recognize the gender of the noun “spirochete”, translating it as male form instead of female form. Both machine translation preserve the abbreviation of “subspecie”, a borrowing from Latin, even if in Italian, being a Romance language, there is the equivalent term “sottospecie”.

In the next sentence there is only a slight error of the adjective “prevalent”:

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
Despite inexpensive and effective antibiotic therapy, syphilis remains a prevalent disease in developing countries and has re-emerged as a public health threat in developed nations.	Nonostante una terapia antibiotica poco costosa ed efficace, la sifilide rimane una malattia prevalente nei paesi in via di sviluppo ed è riemersa come una minaccia per la salute pubblica nei paesi sviluppati.	Nonostante la terapia antibiotica economica ed efficace, la sifilide rimane una malattia prevalente nei paesi in via di sviluppo ed è riemersa come una minaccia per la salute pubblica nei paesi sviluppati.

Both DeepL and Google Translate translate it as “prevalente” whereas its equivalent is “diffusa”: it seems that both machine translation do not recognize a false friend and let themselves be misled by the graphemic similarity. The same error occurs in the next sentence with the adjective “significant”:

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
In addition to the medical burden imparted by infectious syphilis,	Oltre all'onere medico derivante dalla sifilide infettiva, la sifilide	Oltre al carico medico impartito dalla sifilide infettiva, la sifilide

congenital syphilis is considered the most significant infectious disease affecting fetuses and newborns worldwide, and individuals afflicted with syphilis have an enhanced risk for HIV transmission and acquisition.	congenita è considerata la malattia infettiva più significativa che colpisce feti e neonati in tutto il mondo, e gli individui affetti da sifilide presentano un rischio maggiore di trasmissione e acquisizione del virus HIV.	congenita è considerata la malattia infettiva più significativa che colpisce feti e neonati in tutto il mondo e le persone affette da sifilide hanno un rischio maggiore di trasmissione e acquisizione dell'HIV.

Even if one of its meaning is “significativo”, this translation is a bit unfitting in this context. The most appropriate translation in this case is “grave”.

In this part we can observe an inaccuracy in the translation of “burden” made by DeepL, that renders it as “onere”. Here too, even if it is a corresponding translation, the rendering made by GT, “carico”, is more suitable in this case. However, GT carries out another error similar to the previous one made with “prevalent”. In fact, it translates “imparted” as “impartito”, an error due to the similarity of the two words. At the other hand, DeepL translates it correctly as “derivante”.

We have another false friend error in the sentence:

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
Although challenges associated with T. pallidum research have impeded understanding of this pathogen, the existence of a relevant animal model has enabled insight into the correlates of disease protection.	Sebbene le sfide associate alla ricerca su T. pallidum abbiano impedito la comprensione di questo agente patogeno, l'esistenza di un modello animale rilevante ha permesso di comprendere i correlati della protezione della malattia.	Sebbene le sfide associate alla ricerca su T. pallidum abbiano impedito la comprensione di questo patogeno, l'esistenza di un modello animale rilevante ha permesso di comprendere meglio i correlati della protezione della malattia.

Here too the verb “to impede” is translated by both machine translation as “impedire” while it would be more appropriate to translate it as “ostacolare”. Plus, we would have

translated “[…] with T. pallidum research […]” with “[…] sul Treponema pallidum […]”, avoiding to abbreviate the term Treponema, that is unusual in Italian, and putting the preposition “sul”, being implied the term “bacterium”.

In this part we observe an omission made by GT in “[…] understanding of this pathogen […]”, since it does not put the term “agente” as DeepL does. Also the term “microrganismo” would have been suitable.

We can see a further false friend error in “relevant animal model”, where “relevant” is translated by both DeepL and Google Translate as “rilevante”, whereas the correct translation in our opinion would be “adeguato”. It almost seems that machine translation choose the output most phonologically and graphically similar to the one of the source text. The term “animal model” is correctly translated by both machine translation[2]. There is an imprecision in the translation of the technical term “correlates of disease protection”, that both DeepL and Google Translate translate as “correlati della protezione della malattia”. This translation is not totally incorrect, it is just a bit inaccurate. In fact the correct Italian equivalent is “correlati di protezione per la malattia”[3]: both machines translate literally the prepositions of this terminological expression, maintaining its correct meaning. They are terminological errors, however the expression is correct from a grammar point of view.

In the following sentence there are several terms very specific, thus hard to be translated by the machine translation.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
Complete protection against infection has been achieved in the animal model using an extended immunization regimen of γ-irradiated T. pallidum, demonstrating the importance of treponemal	Una protezione completa contro le infezioni è stata ottenuta nel modello animale utilizzando un regime di immunizzazione estesa di T. pallidum γ-irradiato, dimostrando	La protezione completa contro le infezioni è stata raggiunta nel modello animale usando un regime di immunizzazione esteso di T. pallidum γ-irradiato, dimostrando l'importanza

2 Animal models are used to study the development and progression of diseases and to test new treatments before they are given to humans. (https://www.cancer.gov/publications/dictionaries/cancer-terms/def/animal-model)

3 A synonym of this terminological expression is “correlato sierologico di protezione per la malattia”. The correlates of disease protection is really important because it states the clinical effectiveness of a vaccine based on its immunogenicity. (http://www.fimmg.org/index.php?action=pages&m=view&p=2178&lang=it)

surface components in generation of protective immunity and the feasibility of syphilis vaccine development.	l'importanza delle componenti superficiali treponemiche nella generazione dell'immunità protettiva e la fattibilità dello sviluppo del vaccino contro la sifilide.	dei componenti della superficie del treponemico nella generazione dell'immunità protettiva e la fattibilità dello sviluppo del vaccino contro la sifilide.

The first mistranslation that we can observe is made by GT that chooses the unfitted equivalent of the verb “achieved” in this context. In fact the past participle “raggiunta” is unsuitable in this case, while the past participle chosen by DeepL, “ottenuta”, is more suitable. Plus, given that there is not a definite article before “complete protection”, it is more correct collocating an indefinite article in Italian as DeepL does and avoiding to put the definite article “la”.

We had some problems to investigate the correctness of the terms in the phrase “an extended immunization regimen of γ- irradiated T. pallidum” that DeepL and Google Translate translate very similar: searching on Google Books we have seen that the noun “regimen” can be omitted in Italian, however translating it as “regime” is not an error. The translation of “extended” as “esteso” is incorrect as it is an adjective collocated next to “immunization” very often and in Italian the befitting equivalent is “immunizzazione prolungata”. The term “γ-irradiated” is translated literally by both machine translation. However, it does not surprise us since this term is very specific and our hypothesis is that there are few or even no examples of this term in the training dataset for the language pair English-Italian. In our opinion the correct translation should be “treponema pallidum irradiato con raggi γ”. It is interesting to notice that DeepL links the adjective “extended” to the noun “immunization”, while Google Translate links it to the noun “regimen”.

In the phrase “treponemal surface components” DeepL translates erroneously the adjective “surface” as “superficiali”, that in this case in Italian sounds weird. We believe that translating it with “della superficie” represents a better option, as Google Translate does. For the same reason we think that the adjective “treponemal” should be translated as “del treponema” and not as “treponemiche”. Nonetheless, DeepL manages to convey

the meaning of the sentence, even if it is not fully correct from the fluency point of view. Google Translate's output of this phrase is more befitting, except for the adjective "treponemico", which preceded by the preposition "del" is incorrect from a fluency point of view in Italian. Both machine translation translate correctly "protective immunity" with the equivalent Italian term "immunità protettiva".

In the last section of the abstract there are some errors to analyze:

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
Syphilis is a prime candidate for development of a successful vaccine due to the (1) research community's accumulated knowledge of immune correlates of protection; (2) existence of a relevant animal model that enables effective pre-clinical analyses; (3) universal penicillin susceptibility of T. pallidum which enhances the attractiveness of clinical vaccine trials; and (4) significant public health benefit a vaccine would have on reduction of infectious/congenital syphilis and HIV rates. Critical personnel, research and market gaps need to be addressed before the goal of a syphilis vaccine can be realized, including recruitment of additional researchers to the T. pallidum research field with a proportional increase in research funding, attainment of a definitive	La sifilide è un candidato principale per lo sviluppo di un vaccino di successo grazie (1) alle conoscenze accumulate dalla comunità di ricerca sui correlati immunitari di protezione; (2) all'esistenza di un modello animale rilevante che consente analisi pre-cliniche efficaci; (3) alla sensibilità universale alla penicillina di T. pallidum che aumenta l'attrattiva degli studi clinici di vaccino; e (4) ai significativi benefici per la salute pubblica che un vaccino avrebbe sulla riduzione della sifilide infettiva/congenitale e dei tassi di HIV. Il personale critico, la ricerca e le lacune del mercato devono essere affrontate prima che l'obiettivo di un vaccino contro la sifilide possa essere realizzato, incluso il reclutamento di ulteriori ricercatori nel campo di	La sifilide è la prima candidata per lo sviluppo di un vaccino di successo grazie alla (1) conoscenza accumulata dalla comunità di ricerca sui correlati immunitari di protezione; (2) esistenza di un modello animale rilevante che consenta efficaci analisi precliniche; (3) suscettibilità universale alla penicillina di T. pallidum che aumenta l'attrattiva degli studi clinici sui vaccini; e (4) significativi benefici per la salute pubblica che un vaccino avrebbe sulla riduzione della sifilide infettiva / congenita e dei tassi di HIV. Il personale critico, la ricerca e le lacune del mercato devono essere affrontati prima che possa essere realizzato l'obiettivo di un vaccino contro la sifilide, incluso il reclutamento di ulteriori ricercatori nel campo di ricerca del T. pallidum con

understanding of correlates of protection in humans, and engagement of industry/funding partnerships for syphilis vaccine production.	ricerca di T. pallidum con un aumento proporzionale dei finanziamenti alla ricerca, il raggiungimento di una comprensione definitiva dei correlati di protezione nell'uomo, e l'impegno di partnership industriali e di finanziamento per la produzione di vaccino contro la sifilide.	un aumento proporzionale del finanziamento della ricerca, il raggiungimento di una comprensione definitiva dei correlati di protezione nell'uomo e impegno dell'industria / finanziamento di partenariati per la produzione di vaccini contro la sifilide.

The translation of the adjective “prime” is incorrect in both outputs: a more proper rendering in this context is “ideale”, because it conveys better the meaning of being the best candidate possessing the features listed below needed to develop a vaccine. The translation of “successful” is incorrect as well in this context, a suitable translation here is “efficace”. However, they are not errors that impede the understanding of the sentence. As of the term “pre-clinical analyses” both models translate it as “analisi pre-cliniche” that is not incorrct, however another suitable translation can be “studi pre-clinici”. The noun “attractiveness” here is a term, meaning the ability to attract more medical researchers in the syphilis vaccine trial, and it must be translated as “attrattività”.

The last part has a huge syntactic error that we are going to see in a minute. In the first phrase both models link the adjective “critical” to the noun “personnel”. Actually, that adjective should be link with the subject “gaps”, together with its adjuncts “personnel”, “research” and “market”. So, the term “market gaps” is translated incorrectly in this case as “lacune del mercato” (the preposition “di” would be more befitting). In fact, we think that the noun “gaps” is the only subject of the verb “need to be addressed”, and that also the nouns “personnel” and “research” are adjuncts of the noun “gaps”, but the identification of the terminological expression “market gaps” made by both models leads them to mistranslate the entire sentence. So the correct translation of it should be “Prima che l’obiettivo di un vaccino contro la sifilide possa essere

realizzato, devono essere affrontate le lacune critiche in materia di personale, di ricerca e di mercato, […]".

Focusing again on the mistranslation issues we signal the erroneous translation of the preposition "including" as "incluso", while in this case is more appropriate "compreso"; as for the expression "industry/funding partnerships", we had some trouble to find a suitable translation, since both machine translation outputs are incorrect. We made several researches on Google, including translating again this part with DeepL and GT and we discovered an interesting thing: DeepL's output is completely changed, improving considerably. Its output of this sentence is practically perfect:

> Prima di realizzare l'obiettivo di un vaccino contro la sifilide è necessario colmare le lacune critiche a livello di personale, ricerca e mercato, compreso il reclutamento di ulteriori ricercatori nel campo della ricerca sul T. pallidum, con un aumento proporzionale dei finanziamenti per la ricerca, il raggiungimento di una comprensione definitiva dei correlati della protezione nell'uomo e l'impegno di partnership industriali e di finanziamento per la produzione di vaccini contro la sifilide.

The same cannot be said for Google Translate, whose output remains practically unchanged (we have to mention the improvement of the translation of "γ-irradiated T. pallidum" as "T. pallidum irradiato con γ").

But how could DeepL improves in so few months? We are going to discuss this issue in the conclusions of the present chapter.

Collectively, DeepL makes a better output than Google Translate: the latter commits some morphological errors as in "devono essere affrontati" where the verb agrees only with the first subject, while it should agree with the last one; or in the last part, where the noun "impegno" is without the article.

The next article's abstract is about another sexually transmitted disease, HPV, that is human papillomavirus. From a fluency and an accuracy point of view, this abstract is more imperfect than the one about syphilis or the one about gonorrhea; the reason may lie in the fact that this abstract was automatically translated months before than the other two. We have seen, with the syphilis abstract analysis, that neural machine models can improve themselves in a very short period of time. In fact, if we translate again this text with DeepL, we can observe great improvements both from the accuracy and the

fluency point of view. GT's output remains unaltered in this case as well. As we said above we will discuss this issue soon.

Let us look the terminological and fluency matters present in this text. The main topic is human papilloma virus, also known with the acronym HPV, that in Italian has not been translated. At a glance, it seems that Google Translate's output is more correct from the fluency point of view.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
Human papillomavirus vaccination: the population impact We currently have the knowledge and experience to prevent much of human papillomavirus (HPV)-related disease burden globally. In many countries where prophylactic HPV vaccination programs have been adopted as highly effective public health programs with good vaccine coverage, we are already seeing, in real-world settings, reduction of vaccine-related HPV-type infections, genital warts and cervical pre-cancers with potential reductions in vulvar, vaginal and anal pre-cancers. Moreover, we are seeing a change in cervical screening paradigms, as HPV-based screening programs now have strong evidence to support their	Vaccinazione del papillomavirus umano: l'impatto sulla popolazione Attualmente abbiamo le conoscenze e l'esperienza per prevenire gran parte di papillomavirus umano (HPV) - carico di malattia correlato a livello globale. In molti paesi in cui sono stati attuati programmi di vaccinazione profilattica contro l'HPV adottato come programmi di sanità pubblica altamente efficaci con un buon vaccino di copertura, stiamo già assistendo, in contesti reali, ad una riduzione di vaccino-correlate infezioni di tipo HPV, verruche genitali e cervicali pre-infezioni di tipo HPVtumori con potenziale riduzione dei pre-cancri vulvare, vaginale e anale. Inoltre, stiamo assistendo ad un cambiamento nei paradigmi di screening cervicale, come ad esempio	Vaccinazione contro il papillomavirus umano: impatto sulla popolazione Al momento disponiamo delle conoscenze e dell'esperienza per prevenire gran parte del carico di malattie legate al papillomavirus umano (HPV) a livello globale. In molti paesi in cui i programmi profilattici di vaccinazione contro l'HPV sono stati adottati come programmi di salute pubblica altamente efficaci con una buona copertura vaccinale, stiamo già assistendo, in contesti reali, alla riduzione delle infezioni di tipo HPV correlate al vaccino, alle verruche genitali e ai pre-tumori cervicali con potenziali riduzioni dei pre-tumori vulvari, vaginali e anali. Inoltre, stiamo assistendo a un cambiamento nei paradigmi dello screening

use as more sensitive ways to detect underlying cervical abnormalities, as compared with conventional cervical cytology.	I programmi di screening basati sull'HPV hanno ora forti evidenze a sostegno di quanto segue il loro uso come modi più sensibili per rilevare il sottostante cervicale anomalie, rispetto alla citologia cervicale convenzionale.	cervicale, poiché i programmi di screening basati sull'HPV hanno ora forti prove a supporto del loro uso come modi più sensibili per rilevare le anomalie cervicali sottostanti, rispetto alla citologia cervicale convenzionale.

The first sentence is translated incorrectly by DeepL: it does not manage to translate the terminological expression presenting itself as a compound noun “human papilloma (HPV)-related disease”. The problem is that in Italian compound nouns are not separated by a hyphen; plus, compound nouns are less frequent in Romance languages than in Germanic languages and they usually consist of two or three elements. Here, the source text’s compound noun is made up of several elements, so that DeepL is not able to translate it properly, maintaining even the hyphen. GT manages to render the expression properly as “carico di malattie legate al papilloma virus (HPV) a livello globale”. In the next sentence the term “prophylactic vaccination” is translated as “vaccinazione profilattica” by DeepL: it is not completely incorrect, but an equivalent term in Italian exists, that is, “vaccinoprofilassi”. However, it does not respect the number agreement between verb and noun, translating “adopted” as “adottato” when the subject “programs” is plural. GT instead links incorrectly the adjective “prophylactic” to the noun “programs”, syntactically failing to give a proper output and not recognizing the term. We have another terminological error made by DeepL with “vaccine coverage” translated as “vaccino di copertura” and not as “copertura vaccinale”, term recognized by GT. In the final part of this sentence we have again some lexical issues correlated to the compound noun question. The part “reduction of vaccine-related HPV-type infections, genital warts and cervical pre-cancers with potential reductions in vulvar, vaginal and anal pre-cancers” is completely mistranslated by DeepL, that again does not recognize a compound noun, “vaccine-related”, and translates the term “pre-cancer” with “pre-cancri”, when the correct term is “fase pre-cancerosa”. Collectively, the

sentence is totally erroneous from the accuracy point of view, failing to convey the meaning of the source text. We assist to an overly literal translation in the phrase “riduzione di vaccino-correlate infezioni di tipo HPV, verruche genitali e cervicali pre-infezioni di tipo HPVtumori con potenziale riduzione dei pre-cancri vulvare, vaginale e anale”: DeepL maintains all the hyphens of the compound nouns, a sign that is almost absent in the Italian language; the word order is totally wrong and there is an awkward addition of the expression “infezioni di tipo HPV” between “pre” and “cancers”. This error could be attributed to the fact that the terms “verruche genitali” and “verruche cervicali” in several papers are often associated to HPV-type infections. So, taking up the distributional semantic and compositionality theories, representing the starting point of machine translation theory, we could imagine that in a hypothetical vector space, the vectors representing each of these terms usually co-occur very close to each others. GT instead manages to translate better the sentence, outputting a result correct both from the fluency and from the adequacy point of view.

In the next sentence, both machine translation utilize the term “screening”, an Anglicism meaning a paradigm of diagnostic investgations. Here too, GT’s outcome is more accurate than the one made by DeepL. Again, in DeepL’s outcome we can observe an addition of the expression “di quanto segue”, maybe for the reason that the prepositional phrase “a sostegno” is usually followed by the prepositional phrase “di quanto segue”. As we said above, it seems that DeepL makes a mistake because of the usual co-occurrences present in the Italian langauge, especially in technical papers. We have a mistranslation in the phrase “underlying cervical abnormalities” where the adjective underlying is connected incorrectly to the adjective “cervical” and not to the noun “abnormalities”. Both machine translation place the adjective “conventional” at the end of the sentence, however in our opinion placing it before the terminological expression “citologia cervicale” would be more befitting.

Let us see the final part of the abstract

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
This article describes the impact of prophylactic vaccination on these	Questo articolo descrive l'impatto della vaccinazione profilattica su di essi	Questo articolo descrive l'impatto della vaccinazione profilattica su questi esiti e

outcomes and in settings where these vaccines have been implemented in national immunisation programs. Given the successes seen to date and the availability of essential tools, there has been a global push to ensure that every woman has access to effective cervical screening and every girl has the opportunity for primary prevention through vaccination. A gender-neutral approach by offering vaccination to young boys has also been adopted by some countries and is worthy of consideration given that HPV-related cancers also affect males. Furthermore, vaccination of young boys has the advantage of reducing the risk of HPV transmission to sexual partners, lowering the infectious pool of HPV in the general population and ultimately HPV-related diseases for both genders. Therefore, it is appropriate that all countries consider and promote national guidelines and programs to prevent HPV-related diseases.	risultati e nei contesti in cui questi vaccini sono stati implementati in programmi nazionali di vaccinazione. Visti i successi visti fino ad oggi, e la disponibilità di strumenti essenziali, c'è stata una spinta globale per garantire che ogni donna ha accesso ad uno screening cervicale efficace e ogni ragazza ha l'opportunità di prevenzione primaria attraverso la vaccinazione. A l'approccio neutrale dal punto di vista del genere, offrendo la vaccinazione ai giovani ragazzi, ha anche è stata adottata da alcuni paesi e merita di essere presa in considerazione, tenuto conto del fatto che che i tumori legati all'HPV colpiscono anche i maschi. Inoltre, la vaccinazione di giovani ragazzi ha il vantaggio di ridurre il rischio di trasmissione dell'HPV ai partner sessuali, abbassando il pool infettivo di HPV in generale e, in ultima analisi, le malattie legate all'HPV per entrambi i sessi. Pertanto, è opportuno che tutti i paesi considerino e promuovano linee guida e programmi nazionali per la prevenzione delle malattie legate all'HPV.	in contesti in cui questi vaccini sono stati implementati nei programmi nazionali di immunizzazione. Dati i successi finora raggiunti e la disponibilità di strumenti essenziali, c'è stata una spinta globale per garantire che ogni donna abbia accesso a uno screening cervicale efficace e ogni ragazza abbia l'opportunità di prevenzione primaria attraverso la vaccinazione. Un approccio neutrale dal punto di vista del genere, offrendo la vaccinazione ai giovani ragazzi, è stato adottato anche da alcuni paesi ed è degno di considerazione dato che i tumori correlati all'HPV colpiscono anche i maschi. Inoltre, la vaccinazione dei giovani ragazzi ha il vantaggio di ridurre il rischio di trasmissione dell'HPV ai partner sessuali, abbassando il pool infettivo di HPV nella popolazione generale e, in definitiva, le malattie correlate all'HPV per entrambi i sessi. Pertanto, è opportuno che tutti i paesi considerino e promuovano linee guida e programmi nazionali per prevenire le malattie

		correlate all'HPV.

In the first sentence of DeepL's outcome we can see a mistranslation of the adjcetive "these" as "essi", that is a personal pronoun. This represents an awkward error, given the robustness of DeepL translation model. In the following sentence there is a repetition of the past participle "visti", as "given" and "seen" are translated both with the same word. GT instead translates "given" as "dati" and avoids the annoying repetition of the same word. Plus Google Translate manages to gather the subjunctive mood of the verb "has access" that is translated as "abbia accesso", while DeepL maintains the indicative form. The term "primary prevention" is translated with the corresponding equivalent "prevenzione primaria", that is basic form of prevention focusing on the implementation of conducts aimed to avoid the onset of the disease. In the next sentence DeepL make some clumsy errors. There is the untranslated indefinite article "a" and what appears as a double translation of the verb "has been adopted" that it translates as "ha è stata adottata", that is completely nonsense in Italian. It is like DeepL makes an overly literal translation of the verb, translating one word at a time. In the last part the term "pool" is untranslated by both machine translation, however anglicisms are used very often in specialized medical papers and it represent the correct equivalent in Italian. DeepL makes an omission of the noun "population", translating only "in the general" as "in generale"; Google Translate instead translates correctly the sentence placing the adjective "general" after the noun "population", so as to contrast the "general population" with the "infectious pool".

Overall, this abstract is translated better by Google Translate than by DeepL. The latter actually gives an awful outcome, almost literal, that does not resemble to the ones about gonorrhea and syphilis, maybe because of the fact that it has been translated months ago. Probably, in these few months DeepL has managed to improve its translation model, reaching unimaginable results. However we will come back to this issue in the chapter conclusions.

3.3 Popular medical texts' analysis

Let us now discuss the analysis of popular medical texts. We have chosen some paper article from the most popular newspaper: New York Times, Washington Post, and the Guardian. Of course the addressee is different: is a lay reader, who does not have the knowledge to understand clearly the specialized medical language. The addresser as well does not pertain to the specialized domain, but often is a journalist who tries to explain in the most efficient manner the topic to the reader. So the language of these texts change completely. We have less terms, less terminological expressions, and we except to find a simpler style, more narrative and rambling and less synthetic and bald. From a syntactic point of view, popular texts show a complex structure, with more subordinate clauses and asides. Evidently the machine translation will have more difficulties to translate these texts, because is less predictable than technical papers and there are fewer sample to which they can be trained.

The first article we will analyze is about gonorrhea and its characteristic of being drug-resistant. It is a New York Time's article and as we said above is addressed to a lay reader who does not have a specialized knowledge about this subject.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
Agency Urges Change in Antibiotics for Gonorrhea The rates of drug-resistant gonorrhea in the United States have increased so greatly in the last five years that doctors should now treat the infection with a different class of antibiotics, the last line of defense for the sexually transmitted disease, officials said yesterday. The percentage of drug-resistant gonorrhea cases among heterosexual men	Agenzia sollecita un cambiamento negli antibiotici per la gonorrea I tassi di gonorrea resistente ai farmaci negli Stati Uniti sono aumentati così tanto negli ultimi cinque anni che i medici dovrebbero ora trattare l'infezione con una diversa classe di antibiotici, l'ultima linea di difesa per la malattia a trasmissione sessuale, funzionari ha detto ieri. la percentuale di casi di gonorrea resistente ai	L'agenzia sollecita il cambiamento degli antibiotici per la gonorrea I tassi di gonorrea resistente ai farmaci negli Stati Uniti sono aumentati così tanto negli ultimi cinque anni che i medici dovrebbero ora curare l'infezione con una diversa classe di antibiotici, l'ultima linea di difesa per la malattia a trasmissione sessuale, hanno detto ieri i funzionari. La percentuale di casi di gonorrea resistenti ai

jumped, to 6.7 percent in 2006 compared with 0.6 percent in 2001, officials from the Centers for Disease Control and Prevention said. Standard monitoring of gonorrhea cases is conducted among men who go to S.T.D. clinics. New data from such sites in 26 cities show that rates of drug-resistant gonorrhea among heterosexual men at the clinics last year reached 26 percent in Philadelphia and more than 20 percent in Honolulu and four areas in California, Long Beach, Orange County, San Diego and San Francisco. Among gay men at the clinics, the rates of the bacterial infection jumped, to 38 percent in the first half of 2006 from 1.6 percent in 2001. For 14 years, most cases of gonorrhea have been treated with a class of antibiotics known as fluoroquinolones.	farmaci tra gli uomini eterosessuali saltato, al 6,7 per cento nel 2006 rispetto allo 0,6 per cento nel 2001, funzionari dei centri per il controllo delle malattie e prevenzione ha detto. Il monitoraggio standard dei casi di gonorrea è condotto tra gli uomini che vanno alle cliniche S.T.D. I nuovi dati provenienti da tali siti in 26 città mostrano che i tassi di gonorrea resistente ai farmaci tra gli uomini eterosessuali presso le cliniche l'anno scorso ha raggiunto il 26 per cento a Philadelphia e più del 20 per cento a Honolulu e quattro aree in California, Long Beach, Orange County, San Diego e San Francisco. Tra gli uomini gay delle cliniche, i tassi di infezione batterica sono balzati al 38% nella prima metà del 2006, dall'1,6% del 2001. Per 14 anni, la maggior parte dei casi di gonorrea sono stati trattati con una classe di antibiotici noti come fluorochinoloni.	farmaci tra gli uomini eterosessuali è balzata al 6,7% nel 2006 rispetto allo 0,6% nel 2001, secondo quanto riferito dai funzionari dei Centri per il controllo e la prevenzione delle malattie. Il monitoraggio standard dei casi di gonorrea viene condotto tra gli uomini che vanno a S.T.D. cliniche. Nuovi dati da tali siti in 26 città mostrano che i tassi di gonorrea resistente ai farmaci tra gli uomini eterosessuali nelle cliniche lo scorso anno hanno raggiunto il 26% a Filadelfia e oltre il 20% a Honolulu e quattro aree in California, Long Beach, Contea di Orange, San Diego e San Francisco. Tra i gay nelle cliniche, i tassi di infezione batterica sono balzati al 38 percento nella prima metà del 2006 dall'1,6 percento nel 2001. Per 14 anni, la maggior parte dei casi di gonorrea sono stati trattati con una classe di antibiotici noti come fluorochinoloni.

Let us start analyzing the title. Both machine translation render it quite similar, except for the omission of the definite article before the noun “agency” in DeepL outcome. As for the translation of the noun “agency” as “agenzia” we consider it correct, even if a human translator would definitely specify which agency is – in this case it is the Center

for Disease Control and Prevention, whose goal is to protect public health and safety in the US and internationally. DeepL, unlike Google Translate, translates literally the expression “officials said yesterday” putting the subject before the verb, that in Italian sounds a bit awkward in the journalism language and especially in reporting expressions. Besides, there is an agreement error between the subject “funzionari” and the verb “ha detto”. In the next sentence DeepL commits an error in the translation of the verb “jumped”, translating it as a past participle and not as a past simple, which Google Translate does instead, translating the verb as “è balzato” - giving also to the verb a figurative meaning without choosing the first acceptation of it. In the last part of this sentence again DeepL fails to translate “officials from the Center for Disease Control and Prevention said”, redoing the same error of the previous sentence and rendering it as “funzionari dei centri per il controllo delle malattie e prevenzione ha detto”. Plus, mistranslates the agency denomination “Centers for Disease Control and Prevention”, giving an overly literal translation. GT too mistranlates it, and capitalizes “Center”, that does not make much sense if the other nouns of the agency name are not capitalized as well. The correct translation is “centri per la prevenzione e il controllo delle malattie”, putting “prevenzione” before “controllo delle malattie” so that it sound more fluent. In the following sentence there is a slight error in the translation of “S.T.D. clinincs”: DeepL translates it as “cliniche S.T.D.”, collocating correctly the acronym after the noun, but without translating the acronym; Google Translate translates it as “S.T.D. cliniche” that sounds a bit awkward since it puts the noun “cliniche” after the acronym “S.T.D.”. Both machine models do not translate the acronym, notwithstanding there is one in Italian, MTS (malattie trasmissibili sessualmente) or ITS (infezioni trasmissibili sessualmente). Then there is the overly literal translation of “sites” that both DeepL and GT render it as “siti”, when a more befitting translation is “posti”. In the next sentence there are not relevant errors, except for the literal translation made by DeepL of “gay men” that it is translated as “uomini gay”, however the translation of “men” could be omitted, since in the previous part the article was talking about the monitoring of gonorrhea cases among men. Plus, even if it is a borrowing from English widely used in Italian, it seems a little bit strange that both machine translation do not translate “gay” as “omosessuale”, in fact in a newspaper article a reader would except to

read the latter and not the former. The clause “the rates of bacterial infections jumped to 38 percent […] from 1.6 percent in 2001” is translated better by DeepL than by GT, which translates “in 2001” with “nel 2001”, while we consider more correct the outcome of DeepL which translates it with “del 2001”. It sounds much more better in Italian. The term “fluoroquinolones” is well translated by both machine translation with the corresponding equivalent “fluorochinoloni”.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
Now, officials at the center are urging doctors to prescribe drugs in the cephalosporin class.	Ora, i funzionari del centro stanno esortando i medici a prescrivere farmaci nella classe delle cefalosporine.	Ora, i funzionari del centro stanno esortando i medici a prescrivere farmaci nella classe delle cefalosporine.
No new antibiotics for gonorrhea are in the pipeline, officials of the centers told reporters by telephone.	Non sono in cantiere nuovi antibiotici per la gonorrea, hanno detto i funzionari dei centri ai giornalisti per telefono.	Non ci sono nuovi antibiotici per la gonorrea in cantiere, i funzionari dei centri hanno detto ai giornalisti per telefono.
“Now we are down to one class of drugs,” said Dr. Gail Bolan, an expert in sexually transmitted diseases at the California Department of Health Services.	"Ora siamo scesi a una classe di farmaci", ha detto il dottor Gail Bolan, esperto di malattie sessualmente trasmissibili presso il California Department of Health Services.	"Ora siamo passati a una classe di farmaci", ha affermato il dott. Gail Bolan, esperto di malattie a trasmissione sessuale presso il Dipartimento dei servizi sanitari della California.
“That’s a very perilous situation to be in.”	"Questa è una situazione molto pericolosa in cui trovarsi".	"È una situazione molto pericolosa in cui trovarsi."
Dr. Bolan is a spokeswoman for the Infectious Diseases Society of America, a professional organization.	Bolan è portavoce della Infectious Diseases Society of America, un'organizzazione professionale.	Il dottor Bolan è un portavoce della Infectious Diseases Society of America, un'organizzazione professionale.
Health officials are also concerned about extremely drug-resistant tuberculosis and a number of other microbes like Pseudomonas aeruginosa, Klebsiella penumoniae and Acinetobacter species that	I funzionari sanitari sono anche preoccupati per la tubercolosi estremamente resistente ai farmaci e un certo numero di altri microbi come Pseudomonas aeruginosa, Klebsiella	I funzionari sanitari sono anche preoccupati per la tubercolosi estremamente resistente ai farmaci e una serie di altri microbi come Pseudomonas aeruginosa, Klebsiella penumoniae e Acinetobacter che sono

are resistant to most antibiotics. The United States has an estimated 700,000 new cases of gonorrhea a year, occurring among sexually active people of both genders at all ages. It is the second most commonly reported infectious disease, behind chlamydia, another sexually transmitted disease. After a substantial decline from 1975 to 1997, the gonorrhea rates had leveled off in recent years. Action was taken yesterday because the level of resistance has exceeded the standard of 5 percent set by the centers and the World Health Organization. Although the centers' recommendations are not binding, physicians generally follow them. "We are running out of options," said Dr. John M. Douglas Jr., who directs the division of sexually transmitted diseases prevention at the centers. Cephalosporins, like their cousin penicillin, thwart bacteria by damaging a microbe's cell wall, not by attacking DNA as the fluoroquinolones do, Dr. Douglas said.	penumoniae e specie Acinetobacter che sono resistenti alla maggior parte degli antibiotici. Gli Stati Uniti hanno stimato 700.000 nuovi casi di gonorrea all'anno, che si verificano tra le persone sessualmente attive di entrambi i generi a tutte le età. È la seconda malattia infettiva più comunemente segnalata, dopo la clamidia, un'altra malattia a trasmissione sessuale. Dopo un calo sostanziale dal 1975 al 1997, il tasso di gonorrea si è stabilizzato negli ultimi anni. L'azione è stata intrapresa ieri perché il livello di resistenza ha superato lo standard del 5% fissato dai centri e dall'Organizzazione Mondiale della Sanità. Anche se le raccomandazioni dei centri non sono vincolanti, i medici generalmente le seguono. "Stiamo esaurendo le opzioni", ha detto il Dr. John M. Douglas Jr., che dirige la divisione della prevenzione delle malattie sessualmente trasmissibili nei centri. Le cefalosporine, come la loro cugina penicillina,	resistenti alla maggior parte degli antibiotici. Gli Stati Uniti hanno stimato circa 700.000 nuovi casi di gonorrea all'anno, che si verificano tra le persone sessualmente attive di entrambi i sessi a tutte le età. È la seconda malattia infettiva più comunemente segnalata, dietro la clamidia, un'altra malattia a trasmissione sessuale. Dopo un sostanziale declino dal 1975 al 1997, i tassi di gonorrea si erano stabilizzati negli ultimi anni. È stata intrapresa un'azione ieri perché il livello di resistenza ha superato lo standard del 5 percento stabilito dai centri e dall'Organizzazione mondiale della sanità. Sebbene le raccomandazioni dei centri non siano vincolanti, i medici generalmente le seguono. "Stiamo esaurendo le opzioni", ha dichiarato il Dr. John M. Douglas Jr., che dirige la divisione di prevenzione delle malattie a trasmissione sessuale nei centri. Le cefalosporine, come la loro cugina penicillina,

Gonorrhea has not shown resistance to the cephalosporins, which were first marketed in the United States in the 1980s, Dr. Douglas said.	contrastano i batteri danneggiando la parete cellulare di un microbo, non attaccando il DNA come fanno i fluorochinoloni, ha detto il Dr. Douglas. La gonorrea non ha mostrato resistenza alle cefalosporine, che sono stati commercializzati per la prima volta negli Stati Uniti negli anni '80, ha detto il dottor Douglas.	contrastano i batteri danneggiando la parete cellulare di un microbo, non attaccando il DNA come fanno i fluorochinoloni, afferma Douglas. La gonorrea non ha mostrato resistenza alle cefalosporine, che sono state commercializzate per la prima volta negli Stati Uniti negli anni '80, ha affermato il dott. Douglas.

In the first sentence of the second part of the article, both machine translation translate with the corresponding Italian equivalent the term “cephalosporin”. However the translation of “in the cephalosporin class” with “nella classe delle cefalosporine” sounds a bit strange; the preposition “in” should be translated in this case with the preposition “delle”. In the next sentence GT commits the same error DeepL has done in the previous part, when it translates the clause “officials of the center told reporters by telephone” with a literal translation “i funzionari dei centri hanno detto ai giornalisti per telefono”. For a newspaper article it is not an appropriate style, and in Italian in general it sounds odd. There is an overly translation of the verb “to be down” made by DeepL with “siamo scesi”. GT instead chooses an appropriate meaning, that is “siamo passati”. DeepL does not translate “California Department of Health Services”, unlike Google Translate: we believe that in a newspaper article it is necessary to translate organizations’ and services’ name to make everything as clear as possible to the reader. In the next sentence instead both machine translation do not translate “Infectious Diseases Society of America”: here too we would have translated it or maybe we would have put in brackets an Italian translation. The term “professional organization” is translated in both outcomes as “organizzazione professionale”, but we think that “società tra professionisti” suits better. In fact, after a research on Google we have chosen it as the best equivalent for this expression, even if there is not a perfect Italian equivalent. A professional organization does not have a unique definition, and changes

its characteristics according to the purpose they are established for. Therefore, the translation made by both machine translation is not incorrect, as long as it allow addressees to understand clearly what they are reading. The terms of the bacteria – Pseudomonas Aeruginosa, Klebsiella Penumoniae, Acinetobacter - are translated correctly by both machine translation. Google Translate gives a better outcome since it does not translate the term “specie”, as DeepL does. In the next sentence Google Translate gives a better translation of the clause “The United States had an estimated 700,000 new cases […]” because it adds the adverb “circa” given in the source text by the indefinite article “an”. However, in the following sentence makes a mistranslation with the preposition “behind”, that is translated as “dietro”, while in this case it should be translated as “dopo”. In the following sentence too “decline” is translated by Google Translated as “declino”, maybe because of the phonological and graphical similarity, but more proper translation in this case is “calo”. DeepL instead translates it correctly in our opinion.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
Gonorrhea has not shown resistance to the cephalosporins, which were first marketed in the United States in the 1980s, Dr. Douglas said. Now “increased vigilance is essential,” he said, because resistance could still develop at any time, particularly with increased usage. The disease centers say doctors should now prescribe ceftriaxone, sold as Rocephin, which is injected once into a muscle. The centers also recommend the one-time use of cefixime, or Suprax,	La gonorrea non ha mostrato resistenza alle cefalosporine, che sono stati commercializzati per la prima volta negli Stati Uniti negli anni '80, ha detto il dottor Douglas. Ora "una maggiore vigilanza è essenziale", ha detto, perché la resistenza potrebbe ancora svilupparsi in qualsiasi momento, in particolare con l'aumento dell'uso. I centri di malattia dicono che i medici dovrebbero ora prescrivere il ceftriaxone, venduto come rocefina, che viene iniettato una volta in un muscolo.	La gonorrea non ha mostrato resistenza alle cefalosporine, che sono state commercializzate per la prima volta negli Stati Uniti negli anni '80, ha affermato il dott. Douglas. Ora "una maggiore vigilanza è essenziale", ha detto, perché la resistenza potrebbe ancora svilupparsi in qualsiasi momento, in particolare con un maggiore utilizzo. I centri di malattia affermano che i medici dovrebbero ora prescrivere ceftriaxone, venduto come Rocephin, che viene iniettato una volta in un

but tablets of cefixime are not available in the recommended 400-milligram dose. These drugs are meant to substitute for the three currently recommended fluoroquinolones, ciprofloxacin, or Cipro; ofloxacin, or Floxin; and levofloxacin, or Levaquin. For patients allergic to cephalosphins, the centers recommend one injection of spectinomycin, a drug not available in the United States. Over the years, gonorrhea has become resistant to a number of antibiotic classes starting with sulfa, then penicillin and the tetracyclines before fluoroquinolones. The disease centers have gradually cautioned against using fluoroquinolones because of the emergence of resistance in different regions. In 2000, the centers recommended against fluoroquinolones for any patient who acquired gonorrhea in Hawaii, other Pacific Islands and Asia. The agency extended the recommendation to California in 2002. In 2004, the centers	I centri raccomandano anche l'uso una tantum di cefixime, o Suprax, ma le compresse di cefixime non sono disponibili nella dose raccomandata di 400 milligrammi. Questi farmaci sono destinati a sostituire i tre fluorochinoloni, ciprofloxacina, o Cipro; ofloxacina, o Floxin; e levofloxacina, o Levaquin. Per i pazienti allergici alle cefalosfine, i centri raccomandano un'iniezione di spectinomicina, un farmaco non disponibile negli Stati Uniti. Nel corso degli anni, la gonorrea è diventata resistente ad un certo numero di classi di antibiotici a partire dal sulfano, poi la penicillina e le tetracicline prima dei fluorochinoloni. I centri di malattia hanno gradualmente messo in guardia contro l'uso di fluorochinoloni a causa della comparsa di resistenza in diverse regioni. Nel 2000, i centri hanno raccomandato contro i fluorochinoloni per tutti i pazienti che hanno contratto la gonorrea alle Hawaii, in altre isole del Pacifico e in Asia.	muscolo. I centri raccomandano anche l'uso singolo di cefixime o Suprax, ma le compresse di cefixime non sono disponibili nella dose raccomandata di 400 milligrammi. Questi farmaci dovrebbero sostituire i tre fluorochinoloni attualmente raccomandati, ciprofloxacina o Cipro; ofloxacina o floxina; e levofloxacina o Levaquin. Per i pazienti allergici alle cefalosfine, i centri raccomandano un'iniezione di spectinomicina, un farmaco non disponibile negli Stati Uniti. Nel corso degli anni, la gonorrea è diventata resistente a numerose classi di antibiotici a partire dal sulfa, quindi dalla penicillina e dalle tetracicline prima dei fluorochinoloni. I centri di malattia hanno gradualmente messo in guardia contro l'uso di fluorochinoloni a causa della comparsa di resistenza in diverse regioni. Nel 2000, i centri hanno raccomandato contro i fluorochinoloni per tutti i pazienti che hanno acquisito la gonorrea nelle

recommended that fluoroquinolones not be used among gay men with gonorrhea. In 2005, Britain recommended against using fluoroquinolones for gonorrhea because of a resistance problem there. The centers do not plan a letter to doctors on the recommendations. They are relying on news reports and state and local health departments to spread the information.	L'agenzia ha esteso la raccomandazione alla California nel 2002. Nel 2004, i centri hanno raccomandato che i fluorochinoloni non siano usati tra gli uomini gay con gonorrea. Nel 2005, la Gran Bretagna ha raccomandato di non utilizzare i fluorochinoloni per la gonorrea a causa di un problema di resistenza. I centri non pianificano una lettera ai medici sulle raccomandazioni. Si affidano ai notiziari e ai dipartimenti sanitari statali e locali per diffondere le informazioni.	Hawaii, in altre isole del Pacifico e in Asia. L'agenzia ha esteso la raccomandazione alla California nel 2002. Nel 2004, i centri hanno raccomandato di non utilizzare fluorochinoloni tra i gay con gonorrea. Nel 2005, la Gran Bretagna ha raccomandato di non usare fluorochinoloni per la gonorrea a causa di un problema di resistenza lì. I centri non pianificano una lettera ai medici sulle raccomandazioni. Si basano su notizie e dipartimenti sanitari statali e locali per diffondere le informazioni.

In the first sentence, we can observe an agreement error made by DeepL between the noun “cefalosporine” and the verb in the next relative clause “commercializzati”. In both machine translation outcomes there is an error at the level of word order: indeed, it would be correct to collocate the final clause “Dr. Douglas said” at the beginning of the sentence in Italian. Another option would be to put Dr. Douglas’ assertion between in quotes, a usual style in Italian journalism. In the next sentence, DeepL is able to convey the meaning of the source text, however there are some fluency mistakes, for example the phrase “in particolare con l’aumento dell’uso”, which is grammatically correct but it risks to be unclear the reference to the cephalosporins if a possessive adjective is not added before the noun “usage”. A best translation would be “in particolare con l’aumento del loro utilizzo”. We can say the same as for GT’s outcome, where an addition of the possessive adjective “loro” would be useful. However, translating “increased” with “maggiore” does not represent a correct choice because this adjective is already used in the beginning of the sentence, thereby creating an avoidable

cacophony. The translation of “disease centers” as “centri di malattia” is obviously correct, however in Italian is more common to hear of “centri di malattie”, translating “disease” as a plural. Both machine translation translate literally “recommend against” as “raccomandare contro”, while the correct equivalent of this verb is “sconsigliare”. In one of the following sentence we can observe an error committed by DeepL: it translates the that-clause “that fluoroquinolones not be used” as a finite clause - “hanno raccomandato che i fluorichinoloni non siano usati […]”, but utilizing a finite clause after the verb “raccomandare” is unusual in this context in Italian nowadays, even if not totally incorrect. The best translation is the one made by Google Translate which translates the that-clause with a non-finite clause - “hanno raccomandato di non utilizzare”. It seems that the error made by DeepL is caused by the presence of the relative pronoun “that”, given that in the next sentence it does not repeat the same mistake.

The second newspaper article we are going to analyze is about congenital syphilis. It is a quite long article from the Washington Post that explains how congenital syphilis is spreading because of an insufficient prevention. Let us see the first part of the article.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
‘It’s a scandal’: Spiking congenital syphilis rates put babies in grave danger Health experts are pushing for better education, testing and access to health care, and some states with the most cases of congenital syphilis are beginning to respond to the crisis. “Every case of congenital syphilis that we see is an absolute, shameful failure of public health because it is completely preventable,” said Khalil Ghanem, deputy director of education at	E' uno scandalo": I tassi di sifilide congenita in aumento mettono i bambini in grave pericolo Gli esperti sanitari stanno spingendo per una migliore educazione, test e accesso all'assistenza sanitaria, e alcuni stati con la maggior parte dei casi di sifilide congenita stanno cominciando a rispondere alla crisi. "ogni caso di sifilide congenita che vediamo è un assoluto, vergognoso fallimento della salute	È uno scandalo: i picchi di sifilide congenita mettono in grave pericolo i bambini Gli esperti sanitari stanno spingendo per una migliore istruzione, test e accesso all'assistenza sanitaria e alcuni stati con il maggior numero di casi di sifilide congenita stanno iniziando a rispondere alla crisi. "Ogni caso di sifilide congenita che vediamo è un fallimento assoluto e vergognoso della salute pubblica perché è completamente

Johns Hopkins Bayview Medical Center. “We in the United States, a very wealthy country, should not have the rates of congenital syphilis that we have. It’s a scandal.” Curable sexually transmitted diseases, in general, are on the rise around the world.	pubblica perché è completamente prevenibile", ha detto Khalil Ghanem, vice direttore dell'istruzione presso il Johns Hopkins Bayview Medical Center. Noi negli Stati Uniti, un paese molto ricco, non dovremmo avere i tassi di sifilide congenita che abbiamo". È uno scandalo". Le malattie sessualmente trasmissibili curabili, in generale, sono in aumento in tutto il mondo.	prevenibile", ha affermato Khalil Ghanem, vicedirettore dell'istruzione presso il Johns Hopkins Bayview Medical Center. "Noi negli Stati Uniti, un paese molto ricco, non dovremmo avere i tassi di sifilide congenita che abbiamo. È uno scandalo ". Le malattie curabili a trasmissione sessuale, in generale, sono in aumento in tutto il mondo.

The title is translated correctly by both machine translation, however maybe Google Translate’s output, by the use of the noun “picchi”, renders better the idea of something that is increasing inexorably. In the first sentence the translation of the noun “testing” as “test” is correct, even if a better translation would be “esami”. However, as we have said, English borrowings are spreading throughout Italian medical terminology. The verb “to push” is translated by both machine translation as “spingere per”: it is not an error, but maybe another meaning more befitting is “insistere su”; however, we are aware that neural machine translation has not yet achieved such a level of ability. In the last sentence of this part, we notice a slight word order error made by Google Translate, that is “curable sexually transmitted diseases” as “malattie curabili a trasmissioni sessuali”, but collocating the adjective “curabili” in that position in a certain sense breaks the term “malattia a trasmissione sessuale”, which is not an appropriate thing to do. The best thing is doing as DeepL and placing the adjective after the term.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
In the United States, congenital syphilis, which had been declining from	Negli Stati Uniti, la sifilide congenita, in calo dal 2008 al 2012, è in forte e costante	Negli Stati Uniti, la sifilide congenita, in calo dal 2008 al 2012, è in forte e costante

2008 to 2012, has been sharply and steadily rising since 2013. By 2017, there were more than 900 reported cases, causing more than 60 stillbirths and more than a dozen infant deaths, according to the Centers for Disease Control and Prevention. Congenital syphilis is easily preventable when pregnant women are properly screened for syphilis during their pregnancies and given antibiotics to treat it. “The trouble is, we’re just not preventing it,” said Virginia Bowen, an epidemiologist in the CDC’s Division of STD Prevention. Bowen said an increasing number of pregnant women are either not being tested and treated early enough in their pregnancies, or they are acquiring the STD after they have already been screened for it. And those with untreated syphilis may pass it on to their babies.	aumento dal 2013. entro il 2017, ci sono stati più di 900 casi segnalati, causando più di 60 nati morti e più di una dozzina di decessi infantili, secondo i centri per il controllo delle malattie e la prevenzione. La sifilide congenita è facilmente prevenibile quando le donne incinte sono adeguatamente controllate per la ricerca della sifilide durante la gravidanza e ricevono antibiotici per curarla. "Il problema è che non lo stiamo solo impedendo", ha detto Virginia Bowen, un epidemiologo della divisione del CDC per la prevenzione delle malattie sessualmente trasmissibili. Bowen ha detto che un numero crescente di donne in gravidanza o non vengono testate e trattate abbastanza presto nelle loro gravidanze, o stanno acquisendo la STD dopo che sono già stati sottoposti a screening per esso. E quelli con la sifilide non trattata possono trasmetterla ai loro bambini.	aumento dal 2013. Entro il 2017, ci sono stati più di 900 casi segnalati, causando più di 60 morti morti e più di una dozzina di morti infantili, secondo i Centri per il controllo e la prevenzione delle malattie. La sifilide congenita è facilmente prevenibile quando le donne in gravidanza sono adeguatamente sottoposte a screening per la sifilide durante la gravidanza e ricevono antibiotici per curarla. "Il problema è che non lo stiamo semplicemente impedendo", ha dichiarato Virginia Bowen, epidemiologa della Divisione di prevenzione delle malattie sessualmente trasmissibili del CDC. Bowen ha affermato che un numero crescente di donne in gravidanza non viene sottoposta a test e cure abbastanza presto durante la gravidanza, oppure sta acquisendo la MST dopo che è già stata sottoposta a screening. E quelli con sifilide non trattata possono trasmetterlo ai loro bambini.

In the first sentence of this article's part we can notice an awkward and inexplicable repetition made by Google Translate - "causando più di 60 morti morti", a translation of "stillbirths". We cannot understand what has led GT to make a mistake like that. DeepL instead translates the term correctly. In the following sentence, the differences between DeepL's and Google Translate's outputs are really interesting: while Google Translate makes a translation using a borrowing from English medical language – see the use of the terminological expression "sottoposte a screening", whose term "screening" has by now become part of the Italian medical terminology – DeepL translates the sentence using mainly an Italian terminology. In this manner, the terminological expression "to be scanned" is rendered as "essere controllate per", that does not represent an error, however is an unusual expression in the medical domain, where the borrowing of the term "to scan" is the equivalent more utilized. In the next sentence there a word order issue in both machine translation outcomes and an incorrect translation of the verb "to prevent": they translate it as "evitare", but in this case suits better the meaning "prevenire". Also the translation of "just" as "solo" in DeepL's outcome is slight incorrect; it is better the rendering of Google Translate, that is "semplicemente". Both DeepL and Google Translate mistranslate direct object it as masculine, instead it should be translated as feminine since it referes to the congenital syphilis, cited above. As for the word order error, the position of the adverb "just" in both outcomes is inappropriate and sounds a bit weird: it would be more suitable positing it before the verb. Therefore, a better translation could be "Il problema è che semplicemente non la stiamo prevenendo". In the next sentence, DeepL leaves untranslated the acronym term "STD", even if there a corresponding equivalent in Italian that is the one use by GT's, "MST". In the last sentence of this section, again we have an inflection error in the translation of "those" as a masculine demonstrative pronoun "quelli", while it should be feminine since it refers to pregnant mothers.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
"There's a systems failure somewhere," Bowen said. "It's either at the health-care level or public health level, but somehow we are	"C'è un guasto ai sistemi da qualche parte", ha detto Bowen. "È a livello sanitario o a livello di sanità pubblica,	"C'è un errore di sistema da qualche parte", ha detto Bowen. "È a livello di assistenza sanitaria o di sanità

failing these women.”	ma in qualche modo stiamo fallendo queste donne".	pubblica, ma in qualche modo stiamo fallendo queste donne".

In this section, we can see a mistranslation made by DeepL, that is “guasto ai sistemi”; the correct translation is the one made by Google Translate, “errore di sistema”. In the following sentence as well, DeepL translates “health-care level” as “livello sanitario”, but a better rendering is the one made by Google Translate, “assistenza sanitaria”. In this sentence both machine translation mistranslate the verb “to fail” as “fallire”, that seems a pretty literal translation. In this case would be more appropriate to translate it as “deludere”.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
“I wish I could say there’s a smoking gun. There isn’t. There are barriers in women’s lives” that keep women from getting prenatal care, increasing their risk for untreated infection, Bowen said. The highest rate of congenital syphilis cases has been in Louisiana, where the rate is four times the national average. The state reported 59 cases in 2017, a 23 percent increase from the previous year, with the highest numbers recorded in the Lafayette, Monroe and Shreveport areas, according to data from the state’s health department. The majority of the Louisiana mothers, 85	"Vorrei poter dire che c'è una pistola fumante. Non c'è. Ci sono barriere nella vita delle donne" che impediscono alle donne di ricevere cure prenatali, aumentando il rischio di infezioni non trattate, ha detto Bowen. Il tasso più alto di casi di sifilide congenita è stato in Louisiana, dove il tasso è quattro volte la media nazionale. lo stato ha riferito 59 casi nel 2017, un aumento del 23 per cento rispetto all'anno precedente, con i numeri più alti registrati nelle aree di Lafayette, Monroe e Shreveport, secondo i dati del dipartimento della salute dello stato. La maggior parte delle	"Vorrei poter dire che c'è una pistola fumante. Non c'è. Esistono barriere nella vita delle donne "che impediscono alle donne di ricevere cure prenatali, aumentando il rischio di infezione non trattata", ha affermato Bowen. Il tasso più elevato di casi di sifilide congenita è stato in Louisiana, dove il tasso è quattro volte la media nazionale. Lo stato ha segnalato 59 casi nel 2017, con un aumento del 23% rispetto all'anno precedente, con il numero più alto registrato nelle aree di Lafayette, Monroe e Shreveport, secondo i dati del dipartimento sanitario dello stato. La maggior parte delle

percent, who had babies with congenital syphilis were black, and more than half of them were younger than 25, the data shows.	madri della Louisiana, 85 per cento, che hanno avuto bambini con sifilide congenita erano nere, e più della metà di loro erano di età inferiore ai 25 anni, i dati mostrano.	madri della Louisiana, l'85%, che hanno avuto bambini con sifilide congenita erano neri e più della metà di loro aveva meno di 25 anni, i dati mostrano.

In the first sentence of this abstract's section, we can observe a mistranslation of an idiomatic phrase in both outcomes, namely "smoking gun". This phrase is translated literally as "pistola fumante", but in English it means something that represents an evidence or proof. Therefore, one of the proper translation could be "prova incofutabile", or better in this case "un preciso responsabile", since who is speaking is trying to say that he wishes there is someone or something to blame for spiking congenital syphilis, but unfortunately there is not. We know that it would be a too sophisticated rendering for machine translation, however translating an idiomatic phrase literally represents a massive error. In the following part it is performed another literal translation that causes an avoidable repetition of "women": now, in English, when the word at issue acts as possessive, the repetition is allowed; in Italian instead it is considered an error, and it is something that must be avoid. So, the better solution in this case is to replace the noun "women" in the relative clause with the personal pronoun "loro" acting as an indirect object. In the following sentence there is a slight mistranslation with the verb "to be", where the verb "has been" is translated as "è stato"; it would be more suitable translating it as "c'è stato". In the final sentence of this section, we see a word order error, since the clause "i dati mostrano", placed in the final part of the sentence in the source text, must be placed at beginning followed by the relative pronoun "che", turning in this way the main clause in a relative clause. It is something that neural machine translation can perform, as we have seen in the specialized medical abstract, where the order of a sentence was often changed in the output so as to respect the Italian syntax.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
Although most of the	Anche se la maggior parte	Sebbene la maggior parte

mothers received prenatal care at least two months leading up to delivery, 14 percent were not screened for syphilis in a timely manner.	delle madri ha ricevuto cure prenatali almeno due mesi prima del parto, il 14% non è stato sottoposto tempestivamente a screening per la sifilide.	delle madri abbia ricevuto cure prenatali almeno due mesi prima del parto, il 14% non è stato sottoposto a screening per la sifilide in modo tempestivo.
And of the mothers who were, 50 percent were not retested in the third trimester — a crucial time for preventing transmission — and 68 percent were not retested at all.	E delle madri che lo sono state, il 50% non è stato sottoposto a nuovi test nel terzo trimestre - un momento cruciale per prevenire la trasmissione - e il 68% non è stato sottoposto a nuovi test.	E delle madri che lo erano, il 50 percento non fu ripetuto nel terzo trimestre - un momento cruciale per prevenire la trasmissione - e il 68 percento non fu ripetuto affatto.
John Vanchiere, vice chairman for pediatric research at Louisiana State University Health Shreveport, said his team evaluates two to four babies per month for congenital syphilis.	John Vanchiere, vice presidente per la ricerca pediatrica presso la Louisiana State University Health Shreveport, ha detto che il suo team valuta da due a quattro bambini al mese per la sifilide congenita.	John Vanchiere, vicepresidente della ricerca pediatrica presso la Shreveport della Louisiana State University, ha dichiarato che il suo team valuta da due a quattro bambini al mese per la sifilide congenita.
In the vast majority of cases, Vanchiere said, pregnant women are screened early in their pregnancies and either test negative or positive and are successfully treated for the infection, but then acquire syphilis later in the pregnancy.	Nella stragrande maggioranza dei casi, ha detto Vanchiere, le donne incinte vengono sottoposte a screening precocemente nelle loro gravidanze e sia il test negativo o positivo e sono trattati con successo per l'infezione, ma poi acquisire la sifilide più tardi nella gravidanza.	Nella stragrande maggioranza dei casi, ha affermato Vanchiere, le donne in gravidanza vengono sottoposte a screening nelle prime fasi della gravidanza e risultano negative o positive e vengono trattate con successo per l'infezione, ma successivamente acquisiscono la sifilide in gravidanza.
In one recent case in Shreveport, Vanchiere said, a woman who had initially tested negative for syphilis became ill about halfway through her pregnancy after acquiring the STD.	In un caso recente a Shreveport, Vanchiere ha detto, una donna che inizialmente era risultata negativa per la sifilide si è ammalata circa a metà della	In un recente caso a Shreveport, ha detto Vanchiere, una donna che inizialmente era risultata negativa alla sifilide si
She went into labor, and her son was delivered at about		

25 weeks of gestation. Within 48 hours, the baby died, Vanchiere said.	gravidanza dopo aver acquisito la STD. E 'andata in travaglio, e suo figlio è stato consegnato a circa 25 settimane di gestazione. Entro 48 ore, il bambino è morto, disse Vanchiere.	ammalò a metà della sua gravidanza dopo aver acquisito la MST. Andò in travaglio e suo figlio fu consegnato a circa 25 settimane di gestazione. Entro 48 ore, il bambino è morto, ha detto Vanchiere.

Here, in the second sentence, DeepL does not translate “at all”, rendering the output quite weird from a fluency point of view. GT instead translates it correctly as “affatto”, without translating again the verb “were not retested” as “non è stato sottoposto a nuovi test”, choosing just the verb “ripetere”. In the following sentence, both outcomes have the mistranslation of the verb “evaluate” as “valutare”, but in this context translating it as “esaminare” would be more befitting. In the next sentence we notice some issues in DeepL’s outcomes: the term “test negative” and “test positive” is rendered correct just by Google Translate as “risultano negative” and “risultano positive”. The final clause contains an error in DeepL’s outcome, where the verb “to acquire” is translated as an infinitive. In the last sentence of this section there is an error in both outcomes, the translation of the verb “to deliver” as “consegnare”, when in this case the proper equivalent is “far nascere”.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
“Mothers and fathers, especially when there is infectious disease transmission to the baby, often carry a lot of guilt about that because these are preventable diseases.”	"Madri e padri, soprattutto quando c'è la trasmissione di malattie infettive al bambino, spesso portano con sé molte colpe per questo, perché queste sono malattie prevenibili".	"Le madri e i padri, specialmente quando c'è una trasmissione di malattie infettive al bambino, spesso portano molto senso di colpa perché sono malattie prevenibili".

In this sentence we can notice a literal translation of the verb “to carry” as “portare”, but in Italian the right collocation with the noun “guilt” is “sentirsi in colpa”.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE

But in Louisiana, preliminary numbers for 2018 show a 22 percent drop in the number of cases of congenital syphilis — a drop that public health authorities say comes after the implementation of policies and programs to prevent infection. For example, in 2014, state health officials worked with lawmakers to pass state legislation requiring women to be screened upon their first visit and during their third trimester, as well as at the time of delivery, to help determine whether the newborn should be treated.	ma in Louisiana, numeri preliminari per il 2018 mostrano un calo del 22 per cento nel numero di casi di sifilide congenita - un calo che le autorità sanitarie pubbliche dicono che viene dopo l'attuazione di politiche e programmi per prevenire le infezioni. Ad esempio, nel 2014, i funzionari sanitari statali hanno lavorato con i legislatori per approvare la legislazione statale che richiede alle donne di essere controllate alla loro prima visita e durante il loro terzo trimestre, così come al momento del parto, per aiutare a determinare se il neonato deve essere trattato.	Ma in Louisiana, i numeri preliminari per il 2018 mostrano un calo del 22% nel numero di casi di sifilide congenita - un calo che secondo le autorità della sanità pubblica arriva dopo l'implementazione di politiche e programmi per prevenire l'infezione. Ad esempio, nel 2014, i funzionari sanitari statali hanno lavorato con i legislatori per approvare la legislazione statale che impone alle donne di essere sottoposte a screening alla loro prima visita e durante il loro terzo trimestre, nonché al momento del parto, per determinare se il neonato debba essere curato.

In one of the first sentences, we can notice a mistranslation of the noun “implementation” as “implementazione”, due to the phonological and graphical similarity. The verb “comes after” is translated as “arriva dopo”, which is not properly an error, however a better translation can be “deriva”. We can say the same for DeepL’s outcome, where the verb “comes after” is translated simply as “viene”. In the following sentence, DeepL translates overly literal the verb “treated” as “trattato”: the correct equivalent in this case is the one used by GT, “curato”.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
Louisiana has assigned case managers in two high-risk regions of the state to act as liaisons between doctors and patients, assisting	La Louisiana ha assegnato case manager in due regioni ad alto rischio dello stato per fungere da collegamento tra medici e	La Louisiana ha incaricato i case manager in due regioni ad alto rischio di agire come collegamenti tra medici e pazienti,

women throughout their pregnancies and helping them and their partners access the proper STD care. Public health nurses in several regions make contact with pregnant women who have been diagnosed with syphilis and offer to make house calls to administer antibiotics. DeAnn Gruber, director of the Bureau of Infectious Diseases for the Louisiana Department of Health, said the efforts send the message that “this is a very urgent health-care condition that needs to be taken care of." In 2017, Nevada saw 21 cases of congenital syphilis, at a rate of 57.9 per 100,000 live births.	pazienti, assistendo le donne durante la gravidanza e aiutando loro e i loro partner ad accedere alle cure adeguate per le malattie sessualmente trasmissibili. Le infermiere sanitarie pubbliche di diverse regioni entrano in contatto con le donne incinte a cui è stata diagnosticata la sifilide e si offrono di effettuare visite domiciliari per la somministrazione di antibiotici. DeAnn Gruber, direttore del Bureau of Infectious Diseases per il Dipartimento della Louisiana, ha detto che gli sforzi inviano il messaggio che "questa è una condizione sanitaria molto urgente che deve essere curata". Nel 2017, il Nevada ha visto 21 casi di sifilide congenita, ad un tasso di 57,9 per 100.000 nati vivi.	assistendo le donne durante le loro gravidanze e aiutando loro e i loro partner ad accedere alle cure STD adeguate. Gli infermieri di sanità pubblica in diverse regioni entrano in contatto con donne incinte a cui è stata diagnosticata la sifilide e si offrono di effettuare chiamate a domicilio per somministrare antibiotici. DeAnn Gruber, direttore del Bureau of Infectious Diseases per il Dipartimento della Salute della Louisiana, ha affermato che gli sforzi inviano il messaggio che "questa è una condizione sanitaria molto urgente che deve essere curata". Nel 2017, il Nevada ha visto 21 casi di sifilide congenita, ad un tasso di 54,9 per 100.000 nascite vive.

In the first sentence, the term “case managers” is not translated in both outcomes. And it is a proper choice, since in Italian the term is a borrowing from English. Case managers are figures who assist people in difficult situations, helping them find the service they need and create plans for treatments or recovery. They are a sort of a hybrid figure between a nurse and social worker. Maybe, to respect Italian fluency, it would be better to place an indefinite article before the noun, as “dei case manager”. Google Translate’s outcome has some fluency and accuracy errors due to the mistranslation of the verb “to

assign” as “incaricare” and the verb “to act” as “agire”. DeepL instead manages to give a correct outcome both from the fluency and the accuracy point of view. In the next sentence we can observe how DeepL translates the plural noun “nurses” as the feminine plural noun “le infermiere”: however in Italian, given that a neutral gender does not exist, the masculine one is usually preferred if the gender subject is not specified. It is possible that DeepL translate it as a feminine noun because nursing is a profession which is considered inherently natural to the female gender, then the majority of texts on which DeepL is trained could contain the translation of “nurses” as “infermiere”. Therefore, Google Translate’s choice of translating it as plural masculine noun, “gli infermieri”, is more correct. The adjective “public health” is mistranslated by DeepL as “sanitarie pubbliche”, that sounds really weird in Italian. Google Translate’s outcome, that change the adjective into the adjunct “di sanità pubblica”, is more suitable. The term “house calls” is literally translated by Google Translate as “chiamate a domicilio”, while DeepL translates it correctly as “visite domiciliari”. In the following sentence, both machine translation translate incorrectly the cluase “the efforts send the message […]” as “gli sforzi inviano il messaggio che […]”; it would be more useful to the fluency of the texts to add a demonstrative adjective as “tali” before the noun “sforzi” and translate the verb “send” as “mandare”. Or if we avoid a literal translation of the verb phrase but want to maintain its meaning, “send the message” could be translated as “dimostrano”. In the last sentence of this section, the translation of the verb “see” as “vedere” is not proper in this context; however the machine model should have changed too much the structure of the sentence to reach a correct outcome. In fact, the subject “Nevada” should be translated as a locative adjunct “in Nevada” and the verb “saw” should be translated as “ci sono stati”. However, given the difficulty of the text typology, it is a modification too complicated to perform by neural machine translation.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
“Given how preventable these cases are with appropriate testing and treatment and the really effective antibiotics we	"Dato come prevenibili questi casi sono con test e trattamenti appropriati e gli antibiotici veramente efficaci che abbiamo oggi,	"Dato che questi casi sono prevenibili con test e trattamenti adeguati e gli antibiotici veramente efficaci che abbiamo oggi,

have today, we shouldn't be here," she said. The United States was able to get syphilis under control after an epidemic decades ago, she noted, which "gives me hope that we can do this again."	non dovremmo essere qui", ha detto. Gli Stati Uniti sono stati in grado di tenere la sifilide sotto controllo dopo un'epidemia di decenni fa, ha osservato, che "mi dà la speranza che possiamo fare di nuovo questo".	non dovremmo essere qui", ha detto. Gli Stati Uniti sono stati in grado di avere il controllo della sifilide dopo un'epidemia di decenni fa, ha osservato, che "mi fa sperare che possiamo farlo di nuovo".

In the last part of the article, DeepL mistranslates completely the first sentence, and the outcome is quite incomprehensible. Instead GT's outcome is practically perfect: it does not translates it literally as DeepL and the conjunction "given how" is translated as "dato che". It is really impressive how Google Translate has made such a perfect outcome changing the word order and giving to the conjunction at the beginning of the sentence a slight different and correct meaning. In the last sentence DeepL's outcome sounds a little bit weird because of the translation of "this" as "questo", rendering the result overly literal. Google Translate instead renders it better, translating "this" as the Italian direct pronoun "lo" joined to infinitive verb "fare".

The last article is about human papilloma virus and the valuable benefit of immunization. It is from The Guardian, and it takes up a medical research published in the British Medical Journal which is way here we find more terms compare with the other two newspaper articles. It is as long as the other articles just analyzed, and more than the specialized abstracts. Let us begin to analyze the first part of the article.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
HPV rates tumble after routine vaccination Routine HPV vaccination in Scotland has led to a dramatic drop in the cervical disease which can lead to cancer, new data reveals. The vaccine protects	I tassi di HPV cadono dopo le vaccinazioni di routine La vaccinazione ordinaria HPV in Scozia ha portato ad un drastico calo della malattia cervicale che può portare al cancro, rivelano nuovi dati. Il vaccino protegge dal	I tassi di HPV crollano dopo la vaccinazione di routine La vaccinazione di routine per l'HPV in Scozia ha portato a un drammatico calo della malattia cervicale che può portare al cancro, rivelano nuovi dati.

against human papillomavirus (HPV), which causes most strains of cervical cancer. In 2008, routine HPV immunisation of 12 and 13-year-old girls was introduced in schools across the UK. Vaccination has led to far fewer HPV infections, but there has not yet been much published evidence of the impact this has had on cervical disease and cancer. A new study in the British Medical Journal has found a huge drop in the proportions of vaccinated women aged 20 detected at screening with cervical disease – the abnormal cells and lesions that are the precursor of cancer. The researchers found an 89% drop in the numbers of those with the most severe levels of disease, called cervical intraepithelial neoplasia or CIN. There are three grades of CIN, with the most severe, giving the highest risk of cancer, being CIN3+.	papillomavirus umano (HPV), che causa la maggior parte dei ceppi di cancro cervicale. Nel 2008, l'immunizzazione HPV di routine delle bambine di 12 e 13 anni è stata introdotta nelle scuole di tutto il Regno Unito. La vaccinazione ha portato a un numero molto inferiore di infezioni da HPV, ma non sono state ancora pubblicate molte prove dell'impatto che questo ha avuto sulla malattia cervicale e sul cancro. Un nuovo studio del British Medical Journal ha rilevato un enorme calo nelle percentuali di donne vaccinate di 20 anni di età rilevate durante lo screening con la malattia cervicale - le cellule anormali e le lesioni che sono il precursore del cancro. I ricercatori hanno trovato un calo dell'89% nel numero di quelle con i livelli più gravi di malattia, chiamata neoplasia intraepiteliale cervicale o CIN. Ci sono tre gradi di CIN, con il più grave, che dà il più alto rischio di cancro, essendo CIN3+.	Il vaccino protegge dal papillomavirus umano (HPV), che causa la maggior parte dei ceppi di cancro cervicale. Nel 2008, l'immunizzazione di HPV di routine di ragazze di 12 e 13 anni è stata introdotta nelle scuole di tutto il Regno Unito. La vaccinazione ha portato a molte meno infezioni da HPV, ma non sono ancora state pubblicate molte prove dell'impatto che ciò ha avuto sulle malattie cervicali e sul cancro. Un nuovo studio del British Medical Journal ha riscontrato un enorme calo delle proporzioni di donne vaccinate di 20 anni rilevate allo screening con malattia cervicale - le cellule anomale e le lesioni che sono precursori del cancro. I ricercatori hanno riscontrato un calo dell'89% nel numero di persone con i livelli più gravi di malattia, chiamati neoplasia cervicale intraepiteliale o CIN. Esistono tre gradi di CIN, con il più grave, che dà il più alto rischio di cancro, essendo CIN3 +.

In the title, the translation of the verb “to tumble” as “cadono” in DeepL’s outcome does not convey the drastic nosedive of HPV rates as the verb “crollare” does. In the following sentence, there is an error seen before in the previous articles, namely a word order error in the Italian outcome: the clause “new data reveals” should be collocated at the beginning of the sentence, placing the subject before the verb and adding a relative conjunction “che”; the sentence should sound as “nuovi dati rivelano che […]”. In the following sentence, the term “strain” is incorrectly translated as “ceppi” in both machine translations’ outcomes. Actually the correct equivalent is “sottotipi”, that are indeed the varieties through which a cancer may occur. The next sentence is translated incorrectly by both neural models, since English is a language that tends to place several adjectives and modifiers before the noun, making translation difficult for machines. Sure enough, the modifier “routine” is linked just to the noun “immunizzazione”, while the noun “HPV” that here acts as a modifiers, is linked to the noun phrase adjective phrase “routine immunisation”. So, the correct translation should be “l’immunizzazione di routine per l’HPV”. In the following sentence there is again a word order error: DeepL translates overly literal the sentence, rendering in the outcome even the noun “aged” which does not need a translation. Plus, the prepositional phrase “at screening” should be placed after the prepositional phrase “with cervical disease”. Google Translate renders it a little better, however its output has the same word order errors of DeepL. The term “cervical disease” is translated as “malattie cervicali”, that is misleading since the term “cervicale” in Italian identifies also a part of the neck. It would be better translating it as “malattie legate alla cervice uterina”, turning the term “disease” into a plural in Italian. In the following sentence, the acronym CIN is not translated: this is the correct equivalent in Italian, since, even if the term of the disease has a corresponding, the Italian acronym stand for the English term.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
There was a bigger decline in disease among those vaccinated at the age of 12 or 13 (89%) than in those vaccinated at 17 (51%).	C'è stato un declino maggiore della malattia tra coloro che sono stati vaccinati all'età di 12 o 13 anni (89%) rispetto a coloro	C'è stato un più grande declino della malattia tra quelli vaccinati all'età di 12 o 13 anni (89%) rispetto a quelli vaccinati a 17 anni

The vaccine is known to be more effective in those who have not yet encountered HPV, which is a sexually transmitted infection. But there was also a drop in the levels of disease among women who have not been vaccinated, which suggests a "herd immunity" effect, say the authors – because there has been less HPV virus in the community. Routine cervical cancer screening is still highly recommended for women, even if they have been vaccinated, but this may change, say the authors. "Low levels of cervical cancer after routine vaccination clearly have ramifications for screening vaccinated women," the study said. "Although major disease is reduced in Scotland, it has not been eradicated, and continued screening is therefore necessary, particularly as for some years most women within the screening programme will not have been vaccinated."	che sono stati vaccinati a 17 anni (51%). Il vaccino è noto per essere più efficace in coloro che non hanno ancora incontrato l'HPV, che è un'infezione sessualmente trasmissibile. Ma c'è stato anche un calo dei livelli di malattia tra le donne che non sono state vaccinate, il che suggerisce un effetto "immunità mandria", dicono gli autori - perché c'è stato meno virus HPV nella comunità. routine screening del cancro cervicale è ancora altamente raccomandato per le donne, anche se sono state vaccinate, ma questo può cambiare, dicono gli autori. "bassi livelli di cancro cervicale dopo la vaccinazione di routine hanno chiaramente ramificazioni per lo screening delle donne vaccinate", ha detto lo studio. "Anche se la malattia maggiore è ridotta in Scozia, non è stata eradicata, ed è quindi necessario continuare lo screening, soprattutto perché per alcuni anni la maggior parte delle donne che partecipano al programma di screening	(51%). Il vaccino è noto per essere più efficace in coloro che non hanno ancora riscontrato HPV, che è un'infezione a trasmissione sessuale. Ma c'è stato anche un calo dei livelli di malattia tra le donne che non sono state vaccinate, il che suggerisce un effetto di "immunità alla mandria", affermano gli autori - perché c'è stato meno virus HPV nella comunità. Lo screening routinario di carcinoma cervicale è ancora altamente raccomandato alle donne, anche se sono state vaccinate, ma questo può cambiare, affermano gli autori. "Bassi livelli di cancro cervicale dopo la vaccinazione di routine hanno chiaramente conseguenze per lo screening delle donne vaccinate", ha detto lo studio. "Sebbene la malattia maggiore sia ridotta in Scozia, non è stata eradicata e pertanto è necessario uno screening continuo, in particolare poiché per alcuni anni la maggior parte delle donne nell'ambito del

	non saranno state vaccinate".	programma di screening non sarà stata vaccinata".

In the first sentence of this section, both machine translations commit a terminological error: the term “decline” is literally rendered as “declino” which does not fit in this context and it is not even one of the main corresponding meanings. The correct equivalent term is “calo”. In the following part there is another terminological error with the translation of the verb “to encounter”, which again is literally translated by DeepL with “incontrare”. Google Translate instead renders it appropriately as “riscontrato”, however makes some syntactical errors which prejudice the sentence’s fluency: indeed, the verb should be translated as an Italian passive form, that is “essere riscontrato”, in that the focus of the sentence is on the term “vaccine”; therefore, the relative pronoun “who” cannot be literally translated as “che”, but it has to be rendered with a prepositional phrase such as “in cui”. So the correct grammatical construction of this sentence is “Il vaccino è noto per essere più efficace in coloro in cui non è stato ancora riscontrato HPV, che è un'infezione a trasmissione sessuale”. The next sentence have two terminological errors due to an overly literal translation. The term “herd immunity” is translated by both DeepL and GT as “immunità mandria” while the equivalent term is “immunità di gregge”. Furthermore, the finale clause of this sentence includes a terminological error in the translation of the expression “HPV virus”, which is rendered as “virus HPV”, that in this context sounds quite nonsense. The correct equivalent is “infezione da HPV”. In this case, also the translation of the adjective “less” needs to be revised, since “meno” does not fit with the noun “infezione”. In our opinion a better rendering could be “minore”. In the next sentence DeepL leaves untranslated the term “routine screening” and again commits, as GT, a word order error, not placing the clause “say the authors” at the beginning of the sentence. DeepL seems not to get the figurative meaning of the noun “ramifications”, translating it as “ramificazioni” and not with the correct equivalent “conseguenze” - something that instead Google understands. In the last sentence of this section we have a mistranslation of the noun “malattia maggiore” that does not have sense in Italian; a correct translation could be “malattia peggiore”, rendering so the adjective “major” with the Italian relative superlative.

SOURCE TEXT	DEEPL	GOOGLE TRANSLATE
But in the long term, it may be that “two or three screens in a lifetime using HPV testing might be sufficient,” the paper said. Julia Brotherton, the medical director of Australia’s National HPV Vaccination Program Register at VCS Foundation, said in a linked editorial in the journal that the findings are “dramatic and document a considerable reduction in high grade cervical disease over time.” She called for scaling up of HPV vaccination to countries where it is not yet available or accepted. “We must work towards a world in which all girls and their families are offered, and the majority accept, HPV vaccination, wherever they live. We must also actively develop, resource, and scale up more effective, feasible and culturally acceptable strategies for cervical screening, such as self-collection of specimens, if we are ever to effectively reduce the global burden of cervical cancer.”	Ma a lungo termine, può essere che "due o tre schermi in una vita utilizzando il test HPV potrebbe essere sufficiente", ha detto il documento. Julia Brotherton, direttore medico del National HPV Vaccination Program Register presso la VCS Foundation, ha detto in un editoriale collegato nella rivista che i risultati sono "drammatici e documentano una notevole riduzione della malattia cervicale di alto grado nel tempo". Ha chiesto di aumentare la vaccinazione HPV nei paesi in cui non è ancora disponibile o accettata. "Dobbiamo lavorare verso un mondo in cui tutte le ragazze e le loro famiglie sono offerti, e la maggioranza accetta, la vaccinazione HPV, ovunque essi vivano. Dobbiamo anche sviluppare attivamente, risorse e sviluppare strategie più efficaci, fattibili e culturalmente accettabili per lo screening cervicale, come l'auto-collezione di campioni, se vogliamo ridurre efficacemente il	Ma a lungo termine, potrebbe essere che "due o tre schermi in una vita usando i test HPV potrebbero essere sufficienti", ha detto il documento. Julia Brotherton, direttrice medica dell'Australian National HPV Vaccination Program Register presso VCS Foundation, ha dichiarato in un editoriale collegato sulla rivista che i risultati sono "drammatici e documentano una notevole riduzione della malattia cervicale di alto grado nel tempo". Ha chiesto di aumentare della vaccinazione contro l'HPV in paesi in cui non è ancora disponibile o accettato. “Dobbiamo lavorare verso un mondo in cui tutte le ragazze e le loro famiglie sono offerte e la maggior parte accetta la vaccinazione contro l'HPV, ovunque vivano. Dobbiamo anche sviluppare, risorse e scalare attivamente strategie più efficaci, fattibili e culturalmente accettabili per lo screening cervicale,

Robert Music, the chief executive of Jo's Cervical Cancer Trust, said: "The findings of this research are highly exciting and clearly demonstrate the impact of the HPV vaccine in protecting the cervical health of future generations. We are lucky to have such an effective prevention programme which means the elimination of cervical cancer is firmly on the horizon. Focusing on communities and areas where take up is below the national average should be a priority."	peso globale del cancro cervicale". Robert Music, l'amministratore delegato di Jo's Cervical Cancer Trust, ha detto: "I risultati di questa ricerca sono molto interessanti e dimostrano chiaramente l'impatto del vaccino HPV nella protezione della salute cervicale delle generazioni future. Siamo fortunati di avere un programma di prevenzione così efficace, il che significa che l'eliminazione del cancro al collo dell'utero è decisamente all'orizzonte. Concentrarsi sulle comunità e sulle aree in cui l'assorbimento è inferiore alla media nazionale dovrebbe essere una priorità".	come l'auto-raccolta di campioni, se vogliamo mai ridurre efficacemente l'onere globale del cancro cervicale. " Robert Music, amministratore delegato del Cervical Cancer Trust di Jo, ha dichiarato: "I risultati di questa ricerca sono estremamente entusiasmanti e dimostrano chiaramente l'impatto del vaccino HPV sulla protezione della salute cervicale delle generazioni future. Siamo fortunati ad avere un programma di prevenzione così efficace che significa che l'eliminazione del cancro cervicale è saldamente all'orizzonte. Concentrarsi su comunità e aree in cui la diffusione è inferiore alla media nazionale dovrebbe essere una priorità. "

In the first sentence of this section, there is a terminological error with the translation of the term "screen": both models translate it as "schermi", that is the general meaning of this noun. But in the medical terminology, especially that related to the public health, "screen" means "any systematic activity which attempts to identify a particular disease in persons in a particular population" (https://medical-dictionary.thefreedictionary.com/screen). Therefore, the befitting equivalent is "controlli". The adjective "linked" is translated as "collegato", which sounds pretty awkward. Here, maybe a better solution would be translating as the prepositional phrase "a riguardo". In the next sentence, DeepL translates the noun

“scaling up” with the verb “aumentare”, however, here the verb that fits better with the noun “vaccination” is “incrementare”. GT does the same, but translates the preposition “of” anyway, which is an error, since in Italian the verb “aumentare” is transitive and does not require any preposition. If we wanted to translated as a noun, the correct equivalent would be “incremento”. In this case, the verb “to call for” should be translated as “esige”, in that fits better in the context and conveys the urgency existing with HPV infections. The following sentence is translated overly literally by both machine translation. They keep the passive voice of the verb “to offer”, but do not turn the source text’s subject in an indirect object headed by the preposition “a” and the source text’s direct object into the subject. Indeed, the correct translation is “dobbiamo lavorare verso un mondo in cui il vaccino per l’HPV viene offerto a tutte le ragazze […] e accettato dalla maggior parte di loro”. The adverb “actively” could have been omitted easily, as Google Translate does, however translating it does not represent an error. In the last sentence of the final section in DeepL’s outcome there is a mistranslation of the verb “take up” as “assorbimento”, while the correct translation is Google’s one, “diffusione”. However, it is impressive that both machine translation understand that this verb had to be translated as a noun in order to avoid spoiling the target text’s fluency.

Analyzing these articles, we can observe how the language is completely changed compared to the specialized one: there are less terms, a more complex structure characterized by the use of subordinate dependent clauses, and of course a style more comprehensible by a lay audience. It seems that Google Translate can give a better output than DeepL, maybe thanks to the huge database it can access – let us think for a moment to Google News, Google Books, and probably Gmail too, that allows it to have a giant training dataset. However, the differences between the two translation models are actually subtle.

4. ERROR-BASED ANALYSIS' DISCUSSION

4.1 DeepL vs Google Translate: who is better?

We hold that two machine translation are very similar. DeepL outperforms GT as for the technical papers, probably because it is trained on the Linguee dictionary, which can be compared to a sort of technical translation memory. At the other hand, GT delivers better outcomes when translating newspaper's articles: this tendency could be attributed to the fact that GT has a huge material on which is trained, namely Google Search Engine, including Google Books, Google News and who knows maybe Gmail. The content of these engines are pretty colloquial and narrative, so that GT has the opportunity to glean from a huge data set. The two machine models make similar errors, mainly relating to word order, terminology, overly literal translation, and morphological aspect. With regard to the first issue, one of the solution seems to be the aforementioned Attention mechanism and back-translation. It seems that sometimes both neural models face some difficulties reordering the sentence in the target output, but maybe this problem is to be attributed to the fact they lose sight of the context, something on which we have see some improvement in a matter of months. Furthermore, English and Italian present several differences with regard to word order: Italian allows greater variation in word order than English, something that generate issues in the translation of large sequences. As for the terminology issue, we cannot say for sure whether DeepL and GT have a terminological database, since they are able to translate some term, while other translated incorrectly or left untranslated. We assume in some respects both have a terminological database, just think of the translation of the term "screening" or "HPV" with their corresponding Italian equivalents. The third issue is related to the first one: we believe that overly literal translations due to missing features are allowing the model to be able to translate long sequences according to context. Besides, it seems that often both models choose the first equivalent meaning appearing in the dictionary, a fact due to probably to statistical calculations, which weights the first result more probable than others. As for the morphological aspect, we have noticed how often there are some errors related to inflection and agreement. This kind of error is caused probably to an

overly literal translation as well, but also to the fact that the English and Italian are very different languages. As we have seen in the previous chapters, they belong to different families and morphological typology: indeed, English is an analytical language, while Italian is inflected language. For instance, this difference lead often to errors related to nouns' grammatical gender which is distinct in Italian and not in English. However, except for this issue, we can admit that both DeepL and GT perform greatly, allowing the reader to understand the content of the texts despite sometimes they are a bit rough. Neural machine models have incredibly improved machine translation, however the role of the translator is needed both in the post-editing phase and in the research field, because it is not possible to keep advancing without specialist figure of the linguistic sector.

In the comparative analysis we have seen how DeepL has made major advancements in a very short period of time. As we have seen above, we discovered it translating again a piece of source text already translated. The improvement were impressive, and we wonder how DeepL could do such as big developments in so few time. The answer may lay in the method of back-translation of the target-side monolingual data (Ranzato et al. 2018; Abdulmimin et al. 2019; Graça et. al. 2019). We know that DeepL architecture is kept secret; we only know that it avails itself of a super computer based in Iceland, with a huge capacity, by which it trains neural networks on several set of multilingual texts. It is thought that probably utilizes attention based convolutional neural networks (Yin et al. 2015), a model that it considers the relationship/influence existing between the different parts or words of one input sentence with the other, and furthermore provides an interdipendent sentence pair representation that can be used in subsequent tasks. Besides this, we hold that one of the reasons for which DeepL is able to improve itself in a short period of time is back-translation. Even if this model is utilized mostly for languages for which the pool of available translation training resources is so small that it cannot be used with existing systems, we believe that it improves languages pair with already existing parallel data as well. Indeed it allows to make enhancement as for large parallel corpora. The back-translation is an unsupervised approach providing more accurate translation. As we can see in the work of some reasearchers (Senrich et al. 2016), NMT has obtained

outstanding performance using only parallel data for training. However, target-side monolingual data plays a fundamental role in augmenting fluency for phrase-based SMT. So this approach was investigated also in the NMT. In back-translation synthetic source sentences are created by an automatic translation of the monolingual target text into the source language. Then during the training, synthetic parallel text are mixed into the original parallel text. Basically the result of this approach leads to a boosting of the fluency of the outcome, and back-translation reveals itself to be really effective in improving NMT's results.

Another important role is played by the attention mechanism (Vaswani et al 2017). The main idea behind attention mechanism is to improve the traditional Seq2Seq model, which discards all the intermediate state of the encoder and use only the final state to initialize the decoder. This method works for smaller sequences, "but as the length of the sequence increases, a single vector becomes a bottleneck and it gets very to summarize long sequences into a single vector" (Lamba 2019). So central idea of attention mechanism is to keep the intermediate encoder states and utilize them to construct the context vectors required by the decoder to generate the output sentence. In other words, thanks to the attention mechanism the model has an overview of the context of the source text allowing it to create better outcome. Notwithstanding we do not know how exactly DeepL implements these mechanism in its model, we believe that they are some of the main features that allows it to reach incredible results. These are methods used by GT as well, but evidently the neural architecture possess different way of utilize and implement them, so that the results are not always the same.

4.2 Some considerations on speech analysis

Register analysis is one of the most discussed practices of Translation Studies, about which many scholars have debated. One of the most important figure in this sense is Karl Bülher, who categorized language into three functions through the organon model, according to which linguistic communication can be identified as an expressive function, representation function, or conative function. Because of Bülher's work, many attempts have been made in order to set up a typology of texts. Texts generally have

been classified according to criteria such as field of discourse, that according to Halliday is "the total event, in which the text is functioning, together with the purposive activity of the speaker or writer; it thus includes subject-matter as one element in it" (Halliday 1994). Halliday is one of the most influential scholar who investigated the textual meaning and the concept of register, defining it as "a configuration of meanings that are typically associated with a particular situation of field, mode and tenor" (Halliday 1990). Over the years, the concept of field has been declined by other theorist in different notions, such as domain. Domain has been taken as "the basis for developing a different classification of texts into types such as literary, poetic and didactic" (Beaugrande and Dressler 1981). Although it was criticized in that it is considered a vague concept standing between discourse and function, we hold that is useful to analyze technical texts paying attention to conveyance of meaning in the target text and its reception by the addressee. Texts own a semantic and pragmatic function which suit a clear communicative intent. When we talk about a pragmatic function, we mean that texts are the result of an intentional communication activity, we could say a purpose, between and addresser and an addressee.

Another notion in the register and discourse analysis theory is the one of genre. Genre draws on Aristotelian principles expressed in one of his most important work, Rhetoric (Mapelli 2014). According to Aristotle there exist three main genres of rhetoric – deliberative, forensic, and epideictic. This categorization has been taken up by other linguists, as Bajtín, who applied it to the linguistic analysis of speech. In his speech theory, he maintains that in order to write a text which fulfills its communicative function, the author needs to choice the suitable genre for the addressee. Therefore, the purpose of the text establishes its structural and textual features. Genre notion is pivotal in the analysis of text discourse, keeping in mind that in writing a text pertaining to a specific domain entails a shared language and terminology, a specific pattern to follow, and a strict content set-up. For our analysis we chose two different type of discourse – the medical article's abstract and the newspaper article relating to medical matters. Both have their specific structure and language. The former has a strict textual architecture and a precise terminology, the latter is more apt to author's modification, its structure

can be less rigorous, however its purpose is conveying to the reader news in an effective manner.

Therefore, media discourse as well is part of a specialized genre, performing a mediation between the audience and the reality events. Journalists need to convey the message bearing in mind who their readers are and their knowledge of the argument faced in the article, that is if they are experts of the subject or lay readers. This is especially true in articles with medical topic as their subject, whose aim is making complex issues belonging to a highly technical field accessible to a wide public, and not only to the scientific one. The main objective of journalism is disseminating information effectively with a neutral tone. Journalism discourse features are conciseness, clarity, and consistency. News macro-structure is represented by the well-known metaphor of the inverted pyramid – the part at the top represents the most important information that the writer means to convey, while the lower portion illustrates background details (Mapelli 2014). One of the fundamental part of the news structure is the title, whose function is attracting readers' attention and encouraging them to read the article. Being news relating to medical topics, there are terms and technical expression explained by the journalist, who often is as lay as the reader. However, as we said above, the purpose is making available complex information, so syntactic structure and lexicon are quite simple, with a prevalence of verb active voice and pretty low subordination. There is a huge use of scientists' quotations, which help the writer to explain in a complete manner, avoiding misunderstanding complex facts relating to the vaccination and immunization. We have to notice how this particular subject is pretty thorny, since over the last years vaccines have been at the center of fierce debate, dividing people all over the world between those for and against immunization programs.

Given that MTs are not yet provided with that kind of sensitivity needed to understand context and discourse, they translate based on solely the source text. So, linguistic choices are taken without taking in consideration readers knowledge or education. MTs are not trained to understand which the environment of texts is, but they have just the mere sentence sequences to translate. However, the source text represents a solid foundation from which to begin, since what we said above about the discourse analysis is valid in any language. What changes are the purely linguistic features of a

linguistic system, that need to be rendered according to the target language fluency. According to these considerations, we can observe how in the three popular newspaper articles analysis GT manages to respect the media discourse structure in the target texts, utilizing an appropriate syntactic pattern and recognizing idiomatic expression when present. Also the terminology used in this context is well translated by GT, while DeepL often overly translates the article, rendering the target language too far from its real appearance.

With regard to specialized medical texts, we have decided to translate three medical article's abstract. They are less long than newspaper article, but their structure is fixed and there are much more technical terms. Abstracts purpose is informing the audience about the content of the article in a complete and concise fashion. Specialized texts are in between argumentative and descriptive writing, in that try to explain something through a neuter and objective manner and at the same time they want to demonstrate and persuade to have reached a given objective. Technical language is constantly changing and terms need to craft a language which cannot be misunderstood. Being the texts chosen research abstracts, the addressee is a peer audience who share the same knowledge of the writer. The macro-structure follows the format known as IMRaD (Mapelli 2014) that stands for introduction, methods, results and discussion. In the introduction the topic is shown. In the method part, usually the steps of the trial are described. In two of our articles – the one about HPV and the one about syphilis – are not described trials, but a report is provided on the current situation in facing these diseases. So the main part is not about methods, but it is about which are the current and future challenges implicated by HPV and syphilis and appropriate methods to fight them are proposed – for instance vaccination programs to restrict HPV spread. In our case, as for these articles, textual pattern does not follow the IMRaD organizational structure. However, being a genre subordinate to a shared standard, scientific abstracts are characterized by a syntactic structure that MTs are able to translate much better that newspaper articles. Indeed, both GT and DeepL provide more correct outcomes compared to popular articles translations. Anyway some terminological errors are there, but the whole syntactic structure is respected and lexical choices, with some exceptions, prove to be correct. It is precisely because of source texts genre's differences that we

obtain translations so different from each other. In fact both MTs translate better specialized than popular texts, because they are trained to dataset which probably contains more sequence of this genre. So, it would be wrong maintaining that MT does not recognize – and thus translate texts incorrectly – different genres. Indeed, depending on the source text they give different outcome and we can say without doubt that specialized texts are better rendered. Of course, our opinion that it is sufficient having just the text in order to translate texts in efficient manner and to convey the source meaning would call into question years of theory about discourse and genre analysis. Of course, this is not the what we mean: we agree with the fact that to translate a text an extra-linguistic awareness is needed, however we believe that texts have pragmatical and socio-linguistic features that allows the translator to understand the nuances that distinguish different writings from each others. It is the language that make the discourse.

Conclusions

The introduction of the neural network system revolutionized machine translation applications. As demonstrated by the comparative analysis of Google and DeepL, both guarantee high quality result outcomes in translating specialist and popular language. Recognising their undoubted merits, it is impossible not to note how far we are from the utopia (or, for some, dystopia) of a technology that could work independently, making the translator's work superfluous. Errors and misunderstandings continue to be multiple and although, as demonstrated by the evolution of DeepL in so less amount of time, one can hope for the ability of machine learning to continuously improve itself, it is reasonable to hypothesize an intrinsic impossibility for technological devices to keep up with the continuous reinvention of human language: mutations and evolution can be pursued but not anticipated. It is perhaps the paradox of "Achilles and the turtle" by Zenon that can represent this unequal relationship in the most effective way possible. As told by the Argentine writer Jorge Luis Borges:

> Achilles, symbol of speed, must reach the turtle, symbol of slowness. Achilles runs ten times faster than the turtle and gives him ten meters of advantage. Achilles runs those ten meters and the turtle runs a meter; Achilles runs that meter, the turtle runs a decimeter; Achilles runs that decimeter, the turtle runs a centimeter; Achilles runs that centimeter, the turtle runs a millimeter; Achilles runs that millimeter, the turtle runs a tenth of a millimeter, and so on to infinity; so that Achilles can run forever without reaching it. (Borges 1974) (My translation).

However, this image should not cause any discomfort. On the contrary, the awareness of the "intrinsic" limits of automated translation can be the starting point for a less distrustful relationship between this and machine translation. Once considered as a support, useful to speed up some aspects and to provide a basic draft on which to intervene later, it is possible to use these features to improve the translations themselves and to go in a direction of wider dissemination of knowledge. The analysis focused on the field of medical language (in particular sexually transmitted diseases) in order to emphasize, more or less implicitly, how much machine translation can help to democratize a knowledge that, most of the time, continues to be inaccessible to most

people, not only for the “mumbo jumbo” typicl of specialized language, but for that preliminary barrier that is the almost exclusive use of English within the scientific community: this is in fact one of the mumbo jumbo discourse that needs to be addressed. On the other hand, another of the objectives of the thesis was to demonstrate how, to date, the different translation softwares works not only in a similar way, but above all provides qualitatively similar results. The error patterns are the same and, as it was to be expected, they work better with a more standardized language, such as the specialized one, than with popular article whose rules are less strict. In any case, the revolution of the neural network can only be welcomed and prelude to important improvements. The fact that the IT community has welcomed and tried to translate into a mathematical language many of the reflections of structuralist philosophy and sociological constructivism can only be a positive fact from the point of view of interdisciplinarity and a greater openness to what are the reflections of a knowledge, such as philosophical-anthropological knowledge, which too often the mathematical and computer sciences tend to ignore. The contamination and hybridization of knowledge can only be productive. As can be seen from the thesis, the challenge therefore remains that of rethinking the role of translators, recognizing and valuing their indispensability as connoisseura of the human and cultural contexts within which each writing is located but, at the same time, training them in the use and knowledge of increasingly performing machine applications. In this way they can contribute, starting from their own needs, to the continuous improvement of software. In the serene awareness that, however fast he may go, Achilles will never be able to reach the turtle of human language.

References

Abdulmumin, Idris, Galadanci, Bashir Shehu, Isa, Abubakar. 2019. Iterative Batch Back-Translation for Neural Machine Translation: A Conceptual Model.

Brenes, Patricia. 2017. *Terminology for beginners*.

Borges, Jorge Luis. 1974. *Obras completas*. Buenos Aires: Emecé Editores.

Cabré, M. Teresa. 1992. *La Terminolgia*. *La teoria, els mètodes, les aplicacions*. [Terminology. Theory, methods and applications]. Barcelona: Emúries.

Castilho, Sheila, Doherty, Stephen, Gaspari, Federico, Moorkens, Joss. 2018. Approaches to Human and Machine Translation Quality Assessment. In Joss Moorkens, Sheila Castilho, Federico Gaspari, Stephen Doherty, *Translation Quality Assessment*, 39-68. Cham: Springer International Publishing AG.

Cheng, Yong. 2019. *Joint Training for Neural Machine Translation*. Singapore: Springer Theses.

Costa, Ângela, Ling, Wang, Luís, Tiago, Correia, Rui, Coheur, Luísa. 2015. A linguistically motivated taxonomy for Machine Translation error analysis. *Machine Translation 29 (2)*. 127-161.

De Beaugrande, R., & Dressler. 1981. *Introduction to text linguistics*. London/New York: Longman.

Firth, John Rupert. 1957. *Papers in Linguistics 1934-1951*. London: Oxford University Press.

Džuganová, Božena. 2002. *Bratislavske lekarske listy 103(6)*. 223-7

Gabutti, G. 2014. Meningococco B: controllo di due focolai epidemici mediante vaccinazione. *Journal of Preventive Medicine and Hygiene*. 55 (2). 35-41.

Graça, Miguel, Kim, Yunsu, Schamper, Julian, Khadivi, Shahram, Ney, Hermann. 2019. Generalizing Back-Translation in Neural Machine Translation. Paper presented at the 4th Conference on Machine Translation (WMT 2019)

Halliday, Michael A.K. & Ruqaiya Hasan. 1990. *Language, context, and text: aspects of language in a social-semiotic perspective.* Oxford: Oxford University Press.

Halliday, Michael A.K. & Ruqaiya Hasan. 1994. *Cohesion in English*. London: Longman.

House, Juliane. 2015. *Translation Quality Assessment*. Abingdon/New York: Routledge.

Kenny, Dorothy. Machine Translation. 2020. In Mona Baker, Gabriela Saldanha, Routledge *Encyclopedia of Translation Studies*, 305-309 [1998]. Abingdon/New York: Routledge.

Koehn, Philip. 2010 [2009]. *Statistical Machine Translation*. New York: Cambridge University Press.

Koponen, Maarit. 2010. Assessing Machine Translation Quality with Error Analysis. Paper presented at Kääntämisen ja tulkkauksen tutkimuksen symposiumin verkkojulkaisu Electronic proceedings of the MikaEL KäTu symposium on translation and interpreting studies 4.

Lamba, Harshall. 2019. Intuitive Understanding of Attention Mechanism in Deep Learning. https://towardsdatascience.com/intuitive-understanding-of-attention-mechanism-in-deep-learning-6c9482aecf4f (accessed 27 February 2020).

Lanciotti, Eudes. 2012. *Biologia e microbiologia sanitaria con tecnologie di controllo*. Bologna: Zanichelli.

Lommel, Arle. 2018. Metrics for Translation Quality Assessment: A Case for Standardising Error Typologies. In Joss Moorkens, Sheila Castilho, Federico Gaspari, Stephen Doherty, *Translation Quality Assessment*, 39-68. Cham: Springer International Publishing AG.

Mapelli, Giovanna. 2014. Texto y Género. In Maria Vittoria Calvi, Cristina Bordonaba Zabalza, Giovanna Mapelli, Javier Santos López, *Las Lenguas De Especialidad En Español*. Roma: Carocci.

Mapelli, Giovanna. 2014. El lenguaje de los medios de comunicación. In Maria Vittoria Calvi, Cristina Bordonaba Zabalza, Giovanna Mapelli, Javier Santos López, *Las Lenguas De Especialidad En Español*. Roma: Carocci.

Mapelli, Giovanna. 2014. El lenguaje técnico-científico. In Maria Vittoria Calvi, Cristina Bordonaba Zabalza, Giovanna Mapelli, Javier Santos López, *Las Lenguas De Especialidad En Español*. Roma: Carocci.

McMorrow, Leon. 1998. Breaking the Greco-Roman Mold in Medical Writing: The Many Languages of 20th Century Medicine. In Henry Fischbach, *Translation and Medicine,* 13-28. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Mikolov, Tomas, Sutskever, Ilya, Chen, Kai, Corrado, Greg, Dean, Jeffrey. 2013. Distributed representations of Words and Phrases and Their Compositionality.

Pascale Raffaele. Google vs. Facebook Machine Translation – What is the difference? https://www.vengaglobal.com/blog/google-vs-facebook-nmt-difference/ (accessed 27 February 2020).

Popović, Maya. 2018. Error Classification and Analysis for Machine Translation Quality Assessment. In Joss Moorkens, Sheila Castilho, Federico Gaspari, Stephen Doherty, *Translation Quality Assessment*, 39-68. Cham: Springer International Publishing AG.

Ranzato, Marc'Aurelio, Lample, Guillaume, Ott, Myle, Conneau, Alexis, Denoyer, Ludovic. 2018. Phrase-Based & Neural Unsupervised Machine Translation.

Rubenstein, Herbert & Goodenough, John. 1965. Contextual correlates of synonyms. *Communications of the ACM 8(10)*. 627-633.

Salhgren, Magnus. 2008. The distributional hypothesis. *Rivista di Linguistica 20(1)*. 33-53.

Saussure, Ferdinand de. 1983. Course in General Linguistics. Eds. Charles Bally & Albert Sechehaye. Trans. Roy Harris. La Salle, Illinois: Open Court.

Senrich, Rico, Haddow, Barry, Birch, Alexandra. 2016. Improving Neural Machine Translation Models with Monolingual Data. Paper presented at the Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Berlin, Germany.

Shannon, Claude. 1948. A Mathematical Theory of Communication. *Bell System Technical Journal.*

Skelac, Ines, Jandrić Andrej. 2020. Meaning as Use: From Wittgenstein to Google's Word2vec. In Sandro Skansi, *Guide to Deep Learning Basics. Logical, Historical and Philosophical Perspectives*, 41-54. Cham: Springer International Publishing AG.

Spivak, Gayatri. 1993. *Outside in the teaching machine*. London/New York: Routledge.

Vaswani, Ashish, Shazeer, Noam, Parmar, Niki, Uszkoreit, Jakob, Jones, Llion, Gomez, Aidan N., Kaiser, Lukasz. 2017. Attention is all you need. Paper presented at the 31st Conference on Neural Information Processing Systems, Long Beach, CA, USA.

Weaver, Warren. 1949. *Translation*. Rockefeller Foundation.

Wittgenstein, Ludwig. 1953. *Philosophical Investigation.* Oxford: Basil Blackwell.

Wu, Yonghui, Schuster, Mike, Chen, Zhifeng, Le, Quoc V., Norouzi, Mohammad, Macherey, Wolfgang, Krikun, Maxim, Cao, Yuan, Gao, Qin, Klaus, Macherey, Klinger, Jeff, Shah, Apurva, Johnson, Melvin, Liu, Xiaobing, Kaiser, Lukasz, Gouws, Stephan, Kato, Yoshikiyo, Kudo, Taku, Kazawa, Hideto, Stevens, Keith, George, Kurian, Patil, Nishant, Wang, Wei, Young, Cliff, Smith, Jason, Riesa, Jason, Rudnick, Alex, Vinyals, Oriol, Corrado, Greg, Macduff, Hughes, Dean, Jeffrey. 2016. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation.

Wulff, Henrik R. 2004. The language of medicine. *Journal of the Royal Society of Medicine 97(4)*. 187-188.

Yin, Wenpeng, Schütze, Hinrich, Xiang, Bing, Zhou Bowen. 2018. ABCNN: Attention-Based Convolutional Neural Network for Modeling Sentence Pairs.