Testing Google translate

This short report will present the results of a test on Google Translate to evaluate the quality of the translations using texts in Italian, Greek, and Dutch.

The following is part of the Italian constitution. It is a text written in clear and standard Italian.

> "Tutti i cittadini hanno pari dignità sociale e sono eguali davanti alla legge, senza distinzione di sesso, di razza, di lingua, di religione, di opinioni politiche, di condizioni personali e sociali. È compito della Repubblica rimuovere gli ostacoli di ordine economico e sociale, che, limitando di fatto la libertà e l'eguaglianza dei cittadini, impediscono il pieno sviluppo della persona umana e l'effettiva partecipazione di tutti i lavoratori all'organizzazione politica, economica e sociale del Paese". (Assemblea Costituente, 1947)

It is translated in English as:

> "All citizens have equal social dignity and are equal before the law, without distinction of sex, race, language, religion, political opinion, personal and social conditions. It is the duty of the Republic to remove the obstacles of an economic and social nature which, by limiting the freedom and equality of citizens, prevent the full development of the human person and the effective participation of all workers in the political, economic and of the country"

The translation is perfectly clear and faithful to the original.

The following is the beginning of a novel written by Andrea Camilleri in Italian mixed with Sicilian words. Sicilian belongs to Italo-Dalmatian languages like Italian and it is generally understandable by a native Italian speaker. The words in italic are not pure Italian.

> "Lume d'alba non filtrava nel *cortiglio* della Splendor, la società che aveva in appalto la nettezza urbana di Vigata, una nuvolaglia bassa e densa *cummigliava* completamente il cielo come se fosse stato tirato un telone grigio da cornicione a cornicione, foglia non si *cataminava*, il vento scirocco tardava ad *arrisbigliarsi* dal suo sonno *piombigno*, già si faticava a *scangiare* parole." (Camilleri, 2007).

It is translated in English as:

> "Light of dawn did not filter into the *courtyard* of the Splendor, the company that had contracted out the garbage collection of Vigata, a low and dense cloud *covered* the sky completely as if a gray tarpaulin had been pulled from cornice to cornice, leaf did not *cataminate*, the sirocco wind was late **in** *risking* from its *leaden* sleep, it was already hard to *swallow* words."

Surprisingly the overall quality of the translation is good and simply lacks the dialectical tone of the original. Half of the Sicilian words are not correctly translated, but "*cortiglio*", "*cummigliava*", and "*piombigno*" have an accurate translation.

The following is the introduction of the page dedicated to Turin in Wikipedia in Piedmontese. Piedmontese is Gallo-Iberian language, closer to French than Italian and it cannot be understood by native Italian speakers.

> "Turin a l'é na metròpoj ant l'ëstat italian con 902.137 abitant, capleugh dla sità metropolitan-a omònima e capital ëd Piemont. A l'é la quarta comun-a

italian-a scond j'abitant dòp Roma, Milan e Nàpoli ma a l'é la tersa sità da la mira conòmica. Con l'àrea metropolitan-a a së spantia sna surfassa ëd 1.100 km2 con 2.200.000 abitant (conforma a l'OCSE). A l'é un fra ij pì amportant sènter universitari, colturaj, torìstich e sientìfich dël Pais. A l'é na sità industrial mersì a soe meison viturìstiche. An dzorpì, a l'é sènter dl'edission, dla telecomunicassion, dël cine, dj'areclam, dla cusin-a, dl'architetura, dla progetassion e dlë spòrt." (PMS Wikipedia, N.D.)

It is translated in English as:

"Turin is a metropolis between the Italian state with 902,137 inhabitants, capital of the metropolitan city of the same name and capital of Piedmont. It is the fourth most common Italian city inhabited by Rome, Milan and Naples, but it is the third largest city in the world. The largest metropolitan area covers an area of 1,100 km2 with 2,200,000 inhabitants (according to the OECD). It is one of the most important university, cultural, tourist and scientific centers in the country. It is an industrial site thanks to its viticultural harvest. An dzorpì, a l'é sènter dl'edission, dla telecomunicassion, dël cine, dj'areclam, dla cusin-a, dl'architetura, dla progetassion e dlë spòrt."

The translation is clearly full of errors. It is noted that Google Translates detect the Piedmontese as Corsican.

The following is a translation in Greek of the beginning of a novel originally written in Italian by Italo Calvino:

"Ο άνεμος, καθώς έρχεται στην πόλη από μακριά, της φέρνει απρόσμενα δώρα που μόνο λίγες ευαίσθητες ψυχές τ' αντιλαμβάνονται, όπως αυτοί που

πάσχουν από τον πυρετό του άχυρου και φταρνίζονται εξαιτίας της γύρης άλλων τόπων." (Calvino, 1963; translated by Kallifatidis, 1989)

It is translated in Italian as:

> "Il vento, arrivando da lontano in città, le porta doni inaspettati che solo poche anime sensibili percepiscono, come quelle che soffrono di febbre da **paglia** e starnutiscono a causa del polline di altri luoghi."

The translation contains one error: "*paglia*" means "straw" while the correct translation is "*fieno*" ("f*ebbre da fieno*" = "hay fever"). Overall the translation is very similar to the original text of Calvino:

> "Il vento, venendo in città da lontano, le porta doni inconsueti, di cui s'accorgono solo poche anime sensibili, come i raffreddati del fieno, che starnutano per pollini di fiori d'altre terre" (Calvino, 1963).

Although correct, the style of Google's translation is in line with standard Italian and it sounds less poetic than the original to a native speaker.

The following is the text of Dutch song by Marco Borsato:

> "Open m'n ogen
>
> Kijk om me heen
>
> Alles lijkt veranderd
>
> M'n buik doet raar en ik voel me vreemd
>
> En ik vraag me af wat dit gevoel veroorzaakt heeft
>
> M'n gedachten lijken zelfs niet meer van mij
>
> Kan mezelf niet vinden
>
> En de reden dat ben jij" (Borsato & Ewbank, 1999)

It is translated in Italian as:

"apri i miei occhi

guardami intorno

Tutto sembra essere cambiato

Il mio stomaco si comporta in modo strano e mi sento strano

E mi chiedo cosa abbia causato questa sensazione

I miei pensieri non sembrano nemmeno più miei

Non riesco a trovare me stesso

E il motivo sei tu"

The translation of the second sentence is technically correct, but it sounds unnatural. "Guarda intorno a me" would have been better. The same song gets a better translation in English:

"open my eyes

Look around me

Everything seems to have changed

My stomach is acting weird and I feel strange

And I wonder what caused this feeling

My thoughts don't even seem like mine anymore

Can't find myself

And the reason is you"

It is useful to notice that Dutch and English are both West-Germanic languages and they share characteristics, possibly making easier the translation of such a fragmented text.

Overall it is clear that Google Translate doesn't rely on a database to improve its translations. It works best with standard languages. It has a limited ability to translate

unexpected words in dialect. Being based on machine learning, the best way to improve the technology is to scale up the application in terms of ability to learn and knowledge. The function to get alternative translations may provide good suggestions for the platform. The application should learn a new dimension to correctly interpret the tone of the input to match it in the output.

References

Borsato M. & Ewbank J. (1999) *Binnen*.

Camilleri A. (2007) *La forma dell'acqua.* 42nd ed. Palermo: Sellerio.

Calvino I. (1963) *Marcovaldo*. Turin: Einaudi.

Costituzione della Repubblica Italiana. (1947) Available from: https://www.cortecostituzionale.it/documenti/download/pdf/Costituzione_della_Repubblica_italiana.pdf [Accessed 24/08/2021]

Kallifatidis E. (1989) translation of Calvino I. (1963) *Marcovaldo*. Athens: Kastaniotis.

PMS Wikipedia (N.D.) Turin. Available from: https://pms.wikipedia.org/wiki/Turin [Accessed 24/08/2021]