Check for updates

RESEARCH ARTICLE

REVISED Investigating syntactic priming cumulative effects in MT-human interaction

[version 2; peer review: 2 approved]

Natália Resende

Haddad Assistant Professor in Literary Translation, Trinity Centre for Literary and Cultural Translation, Trinity College Dublin, Dublin, Ireland

v2 **First published:** 12 Aug 2021, **1**:93
https://doi.org/10.12688/openreseurope.13902.1
Latest published: 13 Dec 2023, **1**:93
https://doi.org/10.12688/openreseurope.13902.2

Abstract

Background

A question that deserves to be explored is whether the interaction between English language learners and the popular Google neural machine translation (GNMT) system could result in learning and increased production of a challenging syntactic structure in English that differs in word order between speakers first language and second language.

Methods

In this paper, we shed light on this issue by testing 30 Brazilian Portuguese L2 English speakers in order to investigate whether they tend to describe an image in English with a relation of possession between nouns using a prepositional noun phrase (e.g. the cover of the book is red) or re-use the alternative syntactic structure seen in the output of the GNMT (e.g. the book cover is red), thus manifesting syntactic priming effects. In addition, we tested whether, after continuous exposure to the challenging L2 structure through Google Translate output, speakers would adapt to that structure in the course of the experiment, thus manifesting syntactic priming cumulative effects.

Results

Our results show a robust syntactic priming effect as well as a robust cumulative effect.

Conclusions

Open Peer Review

Approval Status ✓ ✓

	1	2
version 2 (revision) 13 Dec 2023		✓ view
version 1 12 Aug 2021	✓ view	? view

1. **Cândido Oliveira**, Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, Brazil
2. **Michael Carl**, Kent State University, Kent, USA
   Longhui Zou, Kent State University, Kent, USA

Any reports and responses or comments on the article can be found at the end of the article.

These results suggest that GNMT can influence L2 English learners linguistic behaviour and that L2 English learners unconsciously learn from the GNMT with continuous exposure to its output.

Plain language summary
In this research, we aim to investigate whether web-based popular translation tools such as google neural machine translation (GNMT) are capable of influencing the language behaviour of the English students that use them for learning purposes. We also investigate whether English students learn from the output of the systems and how this learning takes place, i.e., whether consciously or unconsciously.

Keywords
Google Translate, Portuguese, English, Language Learning, SyntacticPriming, Machine Translation



This article is included in the Horizon 2020 gateway.



This article is included in the Marie-Sklodowska-Curie Actions (MSCA) gateway.



This article is included in the Machine Learning collection.

Corresponding author: Natália Resende (natcarol2000@gmail.com)

Author roles: Resende N: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 843455. ADAPT Centre is funded by the Science Foundation Ireland Research Centres Programme (Grant 13/RC/2106) and is co-funded by the European Regional Development Fund.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2023 Resende N. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Resende N. **Investigating syntactic priming cumulative effects in MT-human interaction [version 2; peer review: 2 approved]** Open Research Europe 2023, **1**:93 https://doi.org/10.12688/openreseurope.13902.2

First published: 12 Aug 2021, **1**:93 https://doi.org/10.12688/openreseurope.13902.1

REVISED **Amendments from Version 1**

Major differences from the previous version includes:

- I have expanded a section in the revised paper, elaborating on how NP (noun phrase with a relation of possession) production in English poses challenges for Brazilian Portuguese (BP) speakers. This expansion is informed by studies in other languages which indicate that when an L2 syntactic structure lacks a counterpart in the L1, speakers are inclined to opt for an L2 structure that closely resembles their L1. This preference stems from the ease of forming associations through L1 transfer processes.

- I provide a clearer explanation for administering the proficiency test after the experimental sessions were completed.

- I have eliminated an erroneously included sentence that implied a priming effect from the baseline phase, despite the absence of priming stimuli presented to participants during this phase

- The first research question has been rephrased for clarity: "Does the use of Google Translate for language production aid in processing more complex structures in a second language?"

- Consistency in terminology has been improved by exclusively using the acronym GNMT, rather than alternating between GNMT and GT

Any further responses from the reviewers can be found at the end of the article

Introduction

Over the last few years, the output quality of machine translation (MT) has improved considerably thanks to the emergence of the neural machine translation (NMT) systems. The NMT systems represent a new type of statistical engine that uses neural networks algorithms originally inspired by the functioning of the biological brain[1]. With this new paradigm, the NMT systems began to "learn" from the data through the association of patterns, confirming hypotheses in the field of psycholinguistics, according to which the brain organizes language in a series of neural networks that operate based on the association of frequent patterns[2].

In addition to improvements in translation quality, MT technology has also improved users' access to those systems. Nowadays, users can choose to access an MT through a mobile application or a web browser. Moreover, MT users can choose between translation modalities, such as speech, text or image translation. These advances in translation quality coupled with easier access have led NMT systems to become the heroes of intercultural environments. The last world cup in Russia, for example, was considered the *Google Translate world cup*. Google reported a 30% increase in user access to the Google Translate (GNMT) system application during the tournament in Russia, especially translating from Spanish into Russian using the voice translation mode[3]. This event shows that the way people are engaging with MT technology is changing. Instead of relying traditionally on tourist booklets for basic phrases, questions and expressions to communicate in the country's native language, the ease of access to the tools has led tourists to use MT as a tool for quick translation of simple sentences for basic communication in Russian.

Although it can be thought that advances in the development of even more accurate translation systems may compromise interest in language learning, research in the field of computer assisted language learning (CALL) shows that, contrary to these concerns, MT systems are being used as tools supporting second language learning[4]. The study by Garcia and Pena[5] illustrates this trend. The researchers have shown that free online MT systems help students with little command of the second language (L2) to improve their writing skills. With MT mediation, the study has shown that L2 learners communicated better, i.e., with higher quality, in the L2. Results have also revealed that the lower the L2 level, the greater the difference between the number of words composed using an MT and the number of those written directly into L2. Another study[6] with undergraduate students from Duke University has shown that, even against the advice of their teachers, students admitted to using the GNMT in their foreign language tasks.

Considering the current scenario in which, on one hand, we see changes in the way people are engaging with MT technology and, on the other hand, its use for language learning purposes, a question that deserves to be explored is related to the role the popular GNMT system plays in the cognitive processing of a second language, in particular, English as a second language (L2 English) due the global spread of interest in English learning[7] and its emergence as *lingua franca*[8].

In this paper, we hypothesize that GNMT output is facilitating L2 English learners to move from more semantic to syntactic processing, especially those in the early stages of language learning. It has already been observed that less proficient learners rely more on the lexical memory of isolated words, so they encounter difficulties and less automaticity in language tasks that involve the proper combination of words in sentences[9]. Therefore, it is at this stage and for that purpose that translation tools represent an important source of support. Thus, it may be that, when using GNMT as a tool to support sentence construction in English, the output of GNMT may not only facilitate learners' awareness of the differences in word order between their first language and the target language but also influence students to unconsciously produce sentences in the target language using the structure seen in the MT output. While in the field of second language acquisition (SLA) the interaction hypothesis[10] claims that human-human interaction facilitates second language development as it raises learners' awareness of language forms, it may be that human-MT interaction has the potential to raise the same awareness in the target language and enable learners to learn challenging structures and modify grammatically incorrect syntactic structures for better communication.

Rationale for the present study

The goal of the present study is to investigate whether the interaction between L2 English learners and GNMT can result in the learning of structures that are challenging for L2 English learners to process. It is also our aim to investigate whether

any learning trends that can possibly be observed are of implicit (unconscious) or explicit (conscious) nature.

In order to accomplish this paper's goal, we employed a syntactic priming study which is an experimental methodology widely used in the field of psycholinguistics to address issues related to language syntactic processing[11–14]. This experimental paradigm will allow us to identify learning trends in our data by examining whether L2 English speakers manifest so-called *cumulative effects*.

Cumulative effects are associated with the implicit learning account of syntactic priming[11,13]. According to this view, syntactic learning is of unconscious nature and it emerges as a result of continuous exposure to a certain structure. Although in the syntactic priming literature there is a controversy concerning the nature of the repetition phenomenon,i.e., whether it is of implicit (i.e. unconscious) or explicit (i.e. conscious) nature, a number of studies (e.g. 15–17) have demonstrated that continuous exposure to a certain syntactic structure across trials leads to long-lasting adaptation (or learning) within the language production system. In other words, increasing the subjects' experience with a certain structure affects the magnitude of the syntactic priming effect.

Following Shin and Christianson (2012), we adopted a pre-test-priming design as it enabled us to study the influence of the GNMT on participants' performance in a picture description priming task as compared to the pre-test baseline. We have chosen this methodology because, as we will see in the next section, its ecological validity to study both L2 learning and human language behaviour when interacting with artificial partners has been already tested by previous studies[16,18–20].

In our previous paper[19], we tested 20 Brazilian Portuguese L2 English students to check whether, after Google translating Portuguese sentences expressing a relationship of possession between nouns (e.g. *A janela do quarto está fechada* - "the bedroom window is closed") into English, they would describe images in English using the same syntactic structure previously seen in the GNMT output (which differs in word order from learners' first language), that is, whether they would be primed by the MT output or whether they would choose a syntactic alternative that resembles the most common syntactic alternative in Portuguese to convey the same idea. We observed a robust priming effect suggesting that participants were influenced by GNMT syntactic alternative when describing images in English. In a subsequent study[21], we introduced a post-test phase the day after the priming test phase, in which participants were asked to describe the same images they described in the baseline pre-test phase (without using GNMT). With this design, we were able to observe whether the priming effect lasted from the priming session until the next day. Results showed that participants tended to describe the images using the syntactic alternative they had seen in the GNMT output the day before, rather than the syntactic alternative that resembled the word order in their native language. These results suggest that participants learned a challenging structure in English through the GNMT output. However, it was unclear to us how this syntactic alignment with the output of GNMT emerged. Specifically, it was unclear whether the priming effect observed in our two previous studies was the result of an implicit learning process triggered by the system output during the experimental session or the result of a learning effect that emerged from the first priming trials presented to the participants. Thus, in this paper, we tested whether the probability of producing the GNMT syntactic alternative when describing images in English increased over time as the experiment proceeded, thus manifesting cumulative effects with the participants' continued exposure to this structure, or the priming effect was triggered by the GNMT output from the beginning of the experimental session onwards. In addition, we investigate whether any cumulative effect observed will vary as a function of participants' English proficiency.

In the present follow-up study, we address the following research questions:

- **RQ1** *When interacting with Google translate for language production purposes, can the output of Google translate facilitate the processing of more challenging structures in the second language?*
- **RQ2** *If RQ1 is true, can learning emerge from the interaction between users and the GNMT system through continuous exposure to that structure via MT output?*
- **RQ3** *If RQ2 is true, does this learning vary as a function of participants' English proficiency?*

As in the two previous studies, we will examine participants' language behaviour when producing an English noun phrase expressing a relationship of possession between nouns (e.g. *the cutlery handles are colourful* or *the handles of the cutlery are colourful*). We focused on this structure as this type of noun phrase varies across the participants' native and non-native languages.

In Portuguese, only one syntactic alternative exists to represent a relationship of possession between nouns. The relationship is always encoded in the preposition *do* (de + o) or *da* (de + a) (e.g. *a mesa do escritório está cheia* or *a porta da casa está fechada*). However, in English, this relationship can be represented using either a prepositional noun phrase (PNP), which follows the same word order as in Portuguese (e.g. *the table of the office is full*) or a non-prepositional noun phrase (NP) (e.g. *the office table is full*), which differs from Portuguese in word order. Given that NP structures do not exist in Portuguese, we posit that they pose more difficulty for native Brazilian Portuguese speakers. As a result, these speakers might produce more PNP sentences due to their similarity with Portuguese syntax that may be transferred to their second language. While no research indicates that producing NP structures in English is more difficult for Brazilian Portuguese speakers than PNP, we assumed it is a more challenging structure because some studies (e.g. 22,23) with Brazilian

Portuguese speakers of English as L2 have shown that when a syntactic structure in the L2 does not exist in the L1 but there is in the L2 a syntactic alternative that matches the L1 structure speakers tend to choose the one that is closer to the L1 structure because it is an easier association due to L1 transfer processes. Specifically for the syntactic structure to express a relation of possession, the study from 24 shows that, when rating preferred English sentences with possessive structures, German speakers of English as L2 often rate higher the sentences using the PNP (of-genitive) with an animate subject instead of the s-genitive case because, differently from English, the s-genitive in German is only used with proper nouns. Therefore, this study is another evidence for transfer trends interfering in the L2 process. Hence, studying how Brazilian Portuguese speakers process these structures in English can help determine if syntactic priming by the GNMT output prompts them to use the NP format more frequently in English.

Before describing our methodology in detail, we review experimental syntactic priming evidence to demonstrate that the syntactic priming rationale represents an ecologically valid approach to address issues involving L2 learning as well as issues involving the interaction between humans and artificial systems.

Background

Cumulative effects in syntactic priming experiments

Syntactic priming, also known in the literature as *structural priming* or *syntactic alignment* (for a complete review on the terminology see 25), can be defined as the tendency of speakers to repeat a syntactic structure previously processed[13]. It also refers to the facilitation of syntactic processing when a syntactic structure is repeated across consecutive sentences[26,27].

The first laboratory study to investigate the syntactic repetition effect dates back to the 80's[13]. In this first study, participants were exposed to transitive sentences in either active (e.g. *One of the fans punched the referee*) or passive (e.g. *The referee was punched by one of the fans*) forms and to dative sentences in either prepositional-object (e.g. *A rock climber sold some cocaine to an undercover agent*) or double-object (e.g. *A rock climber sold an undercover agent some cocaine*) forms. After requesting participants repeat the sentences out loud, researchers asked participants to describe images depicting transitive scenes unrelated to the repeated sentences. The results of this study revealed that participants tended to describe the images using the same structure they had previously produced, i.e., they described the images using passive sentences if they had previously produced passive sentences, double-object sentences if they had previously produced double-object sentences, and so on. This seminal study has paved the way for the emergence of a number of other studies using the same methodological paradigm to investigate the repetition effect. Such interest is based on the assumption that investigating the repetition effect of syntactic structures could help researchers to understand aspects of the nature of human syntactic knowledge, the mechanisms underlying this knowledge as well as aspects related to its learning and development.

Over the few last decades, the effort devoted to research on the syntactic priming effect has brought some insights concerning the nature of the relationship between lexical and syntactic processing. Research has revealed, for instance, that the syntactic priming effect is more frequently driven by the least preferable syntactic structure. A number of studies (e.g. 16,17,26,28) have shown that passive constructions, which are less frequent in English and Dutch, prime more than the highly frequent active constructions. Similarly, the less common dative constructions in English and Dutch prime more than the frequent prepositional double-object constructions. Such an effect, known in the literature as the *inverse preference effect*, is explained by researchers as a consequence of the cognitive mechanism reorganization driven by the surprise of an unexpected structure[17,26,29] as well as by the so-called *cumulative effect* or *accommodation effect*.

The *cumulative effect* occurs when the priming probability (i.e. repetition of syntactic structures previously processed) increases with one's continuous exposure to that syntactic structure over time[17]. Studies reporting cumulative effects show, therefore, that participants adapt their language behaviour to the least preferable structure throughout the entire experimental session.

Some studies (e.g. 11,30) revealed that, when participants were more frequently exposed to one of two structures, the production of the other alternative structure was reduced. For instance, if prepositional datives were more frequently used by participants after repeated exposure to prepositional datives, the double-object datives become less frequently used.

In a study analysing priming effects for word order of auxiliary verbs and the past participle in Dutch subordinate clauses[30], researchers found that participants tended to keep the same word order as the prime sentence in the target sentence. In addition, this experiment has shown a cumulative effect as the use of the least preferred structure (the auxiliary-final structure) became the more preferred in later trials of the experiment compared to the earlier trials, replicating results reported in their previous studies[31].

Another source of evidence comes from the corpus of spontaneous speech[17]. The researchers counted the number of primes that were either comprehended or produced by the same speaker up to the point of the target. Results showed that the more passives the speaker had previously produced in the conversation, the more likely they were to produce another passive in the subsequent utterance. They also found that the more actives speakers produced previously, the less likely they were to produce a passive. Again, these observations suggest that continuous exposure of an infrequent structure increases the probability of that structure being produced later.

In a study testing patients with Korsakoff's syndrome, cumulative effects were also found[32]. Although the aim of the study was to investigate the memory system that supports syntactic priming effects, the researchers analysed learning trends in the data by calculating the proportion of passives out of the total number of transitive responses produced in the target

trials before the current target trial. Results showed that the more passives produced, the stronger the effect, revealing a learning effect of priming in patients with Korsakoff's syndrome. Similarly, in the L2 literature, studies investigating L2 learning through syntactic priming have demonstrated that continuous exposure to a challenging structure to process in the L2 leads to the learning of that structure[33–35].

Syntactic priming is also manifested as facilitation in syntactic processing. A number of studies (e.g. 26,36–38) have shown that participants exposed to a certain syntactic structure tend to process the subsequent sentence with the same structure faster, suggesting that the prime sentence plays a key role in the amount of language resources recruited to produce a sentence. The facilitatory characteristic of the syntactic priming effect has been observed not only in speech production contexts but also in contexts of sentence comprehension.

As syntactic priming is a well-studied phenomenon and can inform researchers about various aspects of syntactic processes, other fields of research have adopted the syntactic priming methodological paradigm to address a variety of research questions related to language processing. In the field of second language acquisition (SLA), syntactic priming methods have been used to investigate the differences and similarities in syntactic processing between the first (L1) and second languages (L2) as well as the impact of the priming effect in the acquisition and processing of a second language[20]. Studies on human-computer interaction (HCI) have also benefited from the syntactic priming methodological paradigm when investigating linguistic behaviour in speech-based interactions between humans and intelligent systems[15,16,18,39]. Below we review studies using the rationale of the syntactic priming methodological paradigm to investigate the influence of intelligent systems on language behaviour and to address issues related to syntactic processing in a second language.

Syntactic priming in HCI

Although syntactic priming effects have been largely observed in monolingual studies in different languages, using different methodological paradigms and different syntactic structures, linguistic alignment has also been demonstrated to occur between humans and computers and artificial systems[16,40]. Research[39] has found, for instance, that children talking to computer partners spontaneously adapt several basic acoustic and prosodic features of their speech by 10–50%, with the largest adaptations involving utterance pause structure and amplitude. Prosodic alignment has also been demonstrated to occur between humans and computers[41]. Moreover, work has found that humans align their lexical choice and gesture handedness in similar ways when interacting with human partners and virtual partners[42].

Regarding syntactic alignment, in human-computer speech-based interactions syntactic priming has been found to occur for both dative structures (e.g. *give the waitress an apple* vs. *give the apple to the waitress*) and noun phrase structures (e.g. *a purple circle* vs. *a circle that is purple*) evidencing that a computer system can also influence a speakers' grammatical choices in speech-based interactions[18]. More recently, virtual reality studies have shown syntactic alignment between humans and computer avatars[16]. The effect has been observed for passives and actives, although the priming effect was stronger for passives than for actives. In a naturalistic experiment with the dialogue system *Let's Go!*, as in laboratory experiments, users adapt to the system's lexical and syntactic choices[43].

All these pieces of evidence in the literature suggest that, even though aware of the artificial nature of the partners, in a communication context, humans tend to consider artificial systems as social actors and, when interacting with them, apply human communication strategies.

Syntactic priming in the field of SLA

Syntactic priming experiments have also been proven to be useful to study second language learning because the syntactic repetition effect is interpreted as a learning process that occurs during the mapping between message form and its meaning, which leads to the subsequent use of the sentence form[44].

In the SLA literature, the syntactic priming experiments follow the same rationale as the syntactic priming experiments within the L1 literature; that is, researchers in the field of SLA test whether a speaker's production is influenced by the structure that was present in the preceding discourse despite the availability of another accepted structure to convey the same meaning.

A seminal study was carried out by Hartsuiker, Pickering and Veltkamp[45]. They investigated whether Spanish L2 English bilinguals would describe pictures in English using passive structures after they listened to passives structures in Spanish or whether they would choose the active structure. In this experiment, researchers found cross-linguistic syntactic priming as participants produced more passive picture descriptions in English after they had just heard a Spanish passive sentence. This work has shown that structural priming plays a beneficial role in L2 development in the production of active and passive structures in English[33].

Different structures have also been tested such as *Wh-* question production in L2 English[35,46,47]. In these studies, researchers tested whether interacting with more advanced English learners would improve learners' performance in producing *Wh-* questions with the supplied auxiliary verb (e.g. *why do people buy products?*) instead of the interlanguage form in which the obligatory auxiliary verb is missing (e.g. *why people buy products?*). The researchers assumed that hearing or producing the advanced *Wh-* question would function as a template for the production of the subsequent use of that form as opposed to the less advanced form. The results of both experiments showed that syntactic priming played a role in the development of *Wh-* question formation.

Syntactic priming for prepositional-object datives has also been tested in the L2[34,47]. Results have shown that participants produced more prepositional-object datives when they had previously heard or produced the prepositional-object structure themselves than when they had not.

A study of Korean L2 English learners[20] investigated whether structural priming improves performance in producing complex, double-object dative (e.g. *The boy is handing the singer a guitar*) and simple, separated phrasal-verb structures (e.g. *The man is putting the fire out*), which are structures that Korean L2 English learners have difficulties in producing. Results showed that syntactic priming improved complex dative production and this improvement was observed to persist over time.

More recently, syntactic priming has been investigated in adverb–verb–subject structures (e.g. *In the winter, jack wears a jacket*) vs. subject–verb–adverb order structures (e.g. *Jack wears a jacket in the winter*) among intermediate English–German second language learners. Participants exhibited comparable short-term priming for adverb-first word order.

From the literature presented above, it can be noted that syntactic priming represents a fruitful methodological approach to study both syntactic processing in L1 and L2, as well as to investigate the impact of machines on human linguistic behaviour. However, unlike the present study, the studies reported above did not address the issue of English students' internal cognitive processes related to implicit learning via structural priming emerging in HCI. Nor has this earlier work examined the effects of structural priming from language comprehension (reading) to production (speaking) through a cross-linguistic task (translation). Therefore, in this paper we intend to contribute to the HCI, MT and SLA fields by carrying out an experiment that involves all these issues. We believe that the present study will bring important insights to the question of whether peoples' grammatical choices can be influenced by the MT systems' grammatical choices and whether this experience could result in (implicit) second language learning.

Methods

In this paper, using a syntactic priming paradigm we investigate cumulative effects with the objective of detecting learning trends elicited by GNMT in the processing of NP in L2 English learners. The experiment was constructed to be carried out in two phases: pre-test phase and priming phase. We will analyse whether participants will change their language behaviour from the pre-test phase to the priming phase. Thus, we considered the pre-test our baseline as, in this phase, participants were not influenced by the MT output.

The rationale behind our experiment is:

- If, in the priming phase, participants describe the images in English using the same structure previously seen in the GNMT output more frequently than the structures used in the pre-test phase (which does not involve any interaction with GNMT), then our results suggest that GNMT is capable of influencing participants' grammatical choices.
- If we observe that the use of the challenge structure increases in the later trials with continuous exposure to the same structure leading participants to adapt to that structure in the course of the experiment, our data suggest that a implicit learning process took place in the course of the experiment as every instance of syntactic structure updates the speaker's knowledge of that structure.

According to some researchers[48], learning occurs because speakers adapt to the context with the aim of reducing errors and uses all the information available to them for this purpose. Therefore, we hypothesise that will we see a priming effect emerging in the course of the experiment, suggesting that GNMT is playing a role in the learning of English syntactic structures as well as in shaping speakers' syntactic processing.

Ethics

The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of Dublin City University (protocol code: DCUREC/2019/110, date of approval: 21 June 2019.) Written informed consent in which participants consented to voluntarily take part in the study and have their demographic data published in journals and elsewhere was obtained from all subjects involved in the study.

Participants

We analysed data from 30 volunteer Brazilian Portuguese L2 English speakers (10 men, mean age=35.7 - sd=5.3). Participants were recruited through posts on Facebook groups of Brazilians living in Dublin and word-of-mouth recommendations. All participants were requested to read a plain language statement and sign the informed consent form to take part in the experiment. The inclusion criteria to take part in the study was to be a native speaker of Brazilian Portuguese, to live in Dublin at the time of the data collection, to have used GNMT as a tool supporting spoken English and be currently enrolled or have previously attended an English school in Dublin, with English proficiency at either intermediate or advanced levels. If not actively enrolled, participants had to share their most recent proficiency level upon exiting the English school. In order to confirm participants' English proficiency levels we asked them to complete the General English test online with 25 questions immediately after completing the pre-test and priming test phases. Participants were classified at the basic level (level A2) with scores ranging from 2 to 13; pre-intermediate level (B1) scores from 14 to 17; intermediate level (B2) scores from 18 to 19; advanced level scores from 20 to 22 and proficient level scores from 23 to 25. Participants reporting an intermediate level from their English school but scoring at the basic level on our test were excluded. The mean score of the participants

included in the experiment was 14.1 (SD=4.7) or level B1). In return for taking part in the experiment, all participants received a €10 voucher.

Materials

Baseline pre-test. For the baseline pre-test, we created a total of 26 sentences in Portuguese and we selected 26 images depicting those sentences from an online image repository[49].

The experimental trials consisted of 20 sentences and images depicting those sentences. The sentences expressed a relationship of possession between nouns and they were composed of a noun phrase + to be verb + complement (e.g. *a janela do escritório está quebrada*, which could be translated by participants using either a prepositional noun phrase structure (PNP) such as *The window of the office is broken* or a noun phrase structure (NP) such as *The office window is broken*. The remaining six trials (30% of the experimental trials) were filler trials which were constructed with sentences composed of subject + to be verb + complement (e.g. *A porta está trancada - "the door is locked"*) and images depicting those sentences. Figure 1 illustrates the trials presented to participants in the baseline pre-test phase.

Priming test. To construct the priming phase trials, we created a total of 26 prime-target triplets trials[49]. As in the pre-test phase, 20 prime-target triplets out of 26 prime-target triplets were experimental trials while the remaining six prime-target triplets were filler trials (30% of the experimental trials). Figure 2 illustrates the 20 experimental trials presented to participants in the priming phase and Figure 3 illustrates the filler trials of the priming phase.

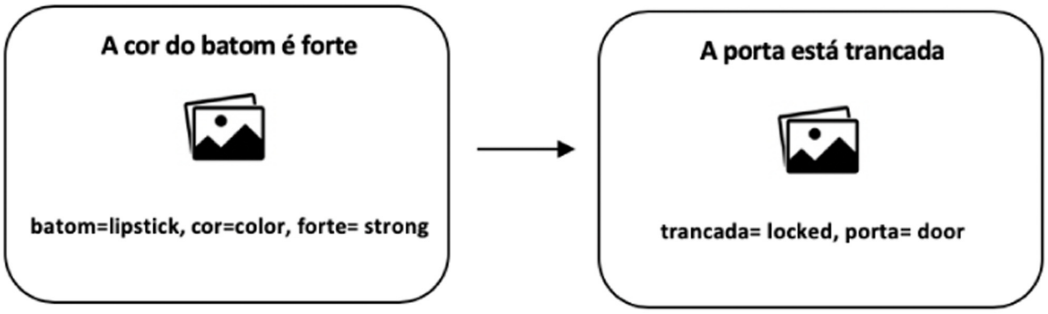


Figure 1. Example of trials presented in the baseline pre-test. The sentences were presented in the written form. Participants were asked to read and then translate the sentences corresponding to the images. They were instructed to verbalize their translations using the words given below each image.

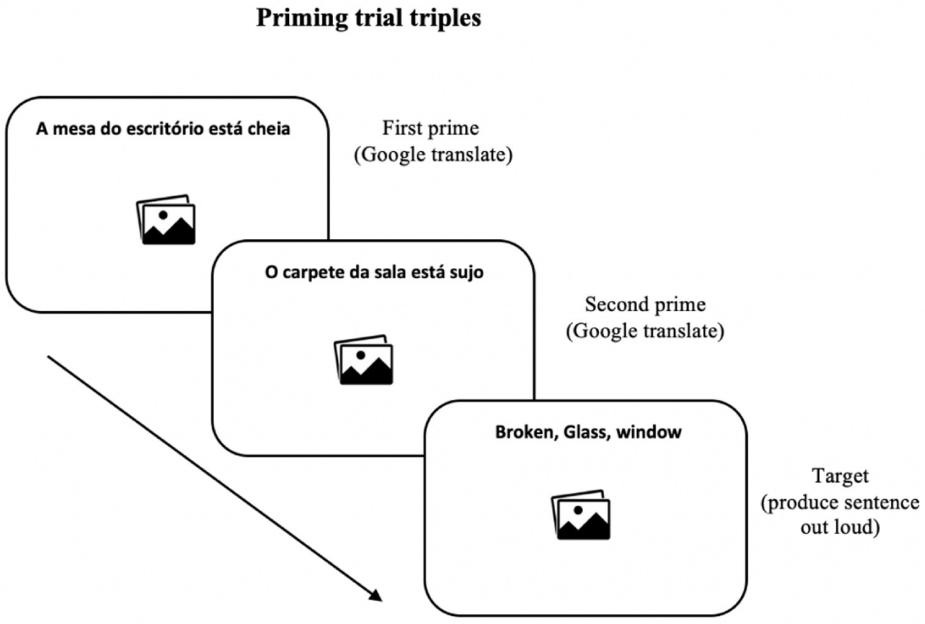


Figure 2. Example of trial presented in the priming test. Participants used Google Translate for the sentences from the first and second primes. For the target items, they were guided to verbally describe the images using the words listed above each image.

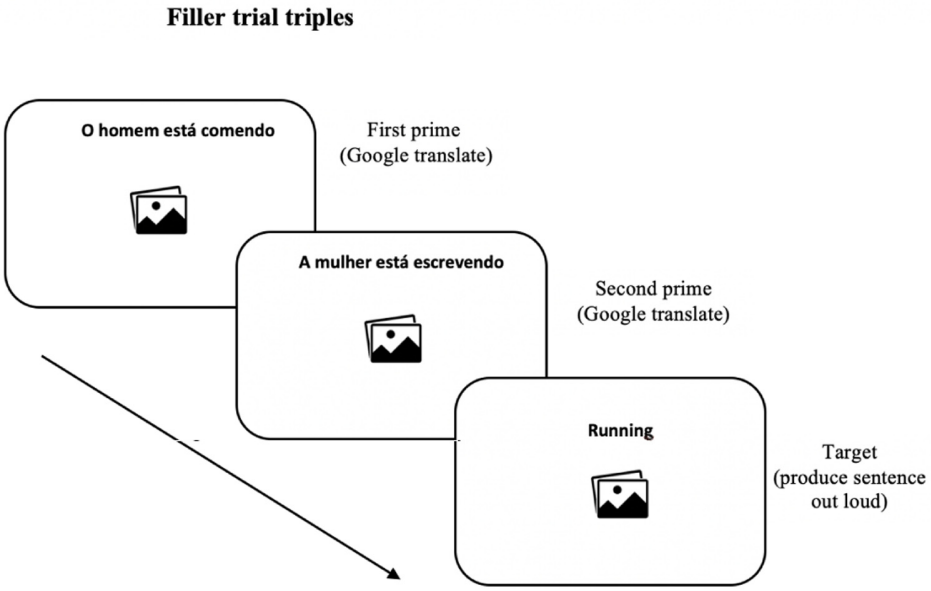


Figure 3. Example of trial presented as fillers.

The 26 prime-target triplets. The 20 prime-target triplets were composed of two prime items preceding a target item, thus we used 40 sentences and 40 images depicting those sentences to create the prime items and 20 sentences and 20 images to depicting those sentences to create the targets, totalling 60 sentences and images depicting those sentences.

To construct the six filler prime-target triplets we used 12 sentences and 12 images depicting those sentences to create the prime items and six sentences and six images depicting those sentences to create the targets, totalling 18 sentences and 18 images depicting those sentences[49]. Thus, in total, the priming test phase used a total of 78 sentences and 78 images depicting those sentences (60 sentences and 60 images for the 20 prime-target triplets and 18 images and 18 sentences for the filler prime-target triplets).

The two prime items preceding the target of the 20 experimental trials (the 20 prime-target triplets) consisted of a sentence in Portuguese composed of a noun phrase expressing a relationship of possession between nouns + to be verb + complement (e.g. *A capa do livro é vermelha* - "The book cover is red"), which were presented to participants to translate from Portuguese into English using a GNMT application on their own mobile device. The target consisted of an image and three English words (randomised across target trials) appearing above the images.

The two prime items preceding the target of the remaining six filler trials (the six prime-target triplets) consisted of sentences in Portuguese composed of a subject + an intransitive verb (e.g. *The man is writing*) presented to participants to translate orally from Portuguese into English using a GNMT application on their own mobile device. The target consisted of an image for description in English and an intransitive verb (e.g. *knitting*) appearing above the image.

It is important to note that when we created the prime sentences, they had all been machine-translated using GNMT on various days and different times spanning two months. We observed that the sentence structures remained consistent over time as Google consistently translated all sentences using an NP structure. We were aware, however, that despite this consistency over time, GNMT output could still vary across experiment trials during sessions. However, because we noticed that GNMT's preferred structure is an NP structure, we thought it would be worth testing whether GNMT output would be capable of eliciting priming effects even if in a few cases the system outputs an alternative structure.

Procedures

Following Cowan *et al.*[18], the experimenter noted the syntactic structure (NP, PNP, 'S or Other) used by the 30 participants to describe the target images in English. The experimenter noted as "PNP" if the image was described by the participant using the English preposition "of" (e.g. *the table of the office*); and as "NP" if participant used a noun phrase a structure in English without the preposition in the correct word order (e.g. *The office table*) mirroring the GNMT output, and "other" if they used a noun phase structure in the incorrect order (e.g. *the table office*), or the genitive case (e.g. *table's office*) to express a relation of possession between nouns or for any other structure used (e.g. *the table in the office*) that differed from GNMT preferred syntactic alternative. These data were then considered the categorical dependent variable used for data analyses. All experimental sessions were audio recorded using *Quick Player* voice recorder and stimuli were presented in a computer screen (Macbook Pro with 13-inch

Retina display screen) using *Psychopy* software version 3.0 which allowed us to randomise the stimuli set of both baseline and priming test phases across participants. The recordings of the sessions allowed a sanity check carried out by an independent blind rater to make sure that all data was correctly coded by the experimenter. The set of stimuli was presented visually and each participant performed the tasks individually in a silent room in the school of computing of the Dublin City University. Prior to the experimental sessions, the experimenter presented the instructions of the tasks to each participant while they appeared on the screen.

Baseline pre-test phase. The baseline pre-test was presented to participants before the priming test. The order of presentation of the 20 trials and the six filler trials were randomised across participants. In this phase, participants were instructed to speak their translation of the sentences from Portuguese into English out loud using the English words presented at the bottom of the screen in order to avoid lexical retrieval issues in the L2 during the task (see Figure 1). All sentences could be translated from Portuguese into English using either a PNP structure or a NP structure.

Priming test phase. The order of the presentation of the 26 prime-target triplets was randomised across participants. In the two prime items preceding the target item, participants were asked to translate the prime sentences above the images using the GNMT application on their own mobile device and repeat the translation out loud in order to trigger the syntactic priming effect[11,13,20].

Immediately after machine translating the prime sentences, participants were presented with the target item and were instructed to orally describe the image on the screen with a simple sentence using the three words appearing above the image (experimental trials) or the intransitive verb (filler trials). Participants were also instructed to avoid including words that were not on the computer screen and avoid describing the images using prepositions of location (such as *in*, *on*, *at*, etc). All images could be described using either a NP structure or a PNP structure.

Two prime-target triplets in the two conditions (experimental and filler condition), not included in the main experiment, were used for participants' training prior to the start of the experimental session.

Coding and analysis

We analysed 40 data points per participant, totalling 1200 data points (40 × 30 = 1200). Data points from filler trials were not included in the analysis.

Participants' responses for both baseline and priming tests were manually coded by the experimenter. In the baseline phase, responses with an NP structure were coded as "1", while those with a PNP structure received a "0". For the priming phase, if participants used the same structure in the target item as seen in the GNMT output (specifically, NP structures) to describe the images, it was coded as "yes" (1). If they used a PNP structure or another variant like "*the office's window is broken*" or "*the window in the office is broken*", it was coded as "no" (0).

Analysis was carried out in R Studio version 1.1.423[50] package lme4[51]. We used a logistic mixed effects maximal model including participants and items as random effects. The factors included as fixed effects were: *cumulative NP proportion* (continuous) to investigate learning trends in data, *English proficiency test score* (continuous) and *test type* (factorial). Following[15,16], *cumulative NP proportion* was calculated as the proportion of NP structures out of all structures produced in the target items before the current target item. We included the *English proficiency test score* (continuous) and *test type* (factorial) with two levels: baseline or priming test. Factorial predictors were dummy coded (all means compared to a reference group) and all numeric predictors were centered.

Predictions based on GNMT output

Before creating our experimental trials, we tested how GNMT would translate Portuguese sentences containing a relation of possession between nouns. We noticed that GNMT translated sentences from Portuguese into English using a NP structure far more frequently than a PNP structure[49]. We also tested all the 40 sentences used to construct the prime items of the priming phase of the experiment and, again, we observed that GNMT translated them using a NP structure which is, as already mentioned, a more challenging syntactic alternative to Portuguese speakers due to differences in word order between learners' first and second languages.

Because the English PNP structures with a relationship of possession between nouns resemble the noun phrase structure word order in Portuguese and there is tendency of speakers to use the structure that is closer to the structure in the L1 (e.g., 22,23) we predict that, in the baseline pre-test, participants will produce more English PNP structures than NP structures, while in the priming phase, we will see a decrease in the use of this structure and, at the same time, an increase in the production of NP structures in the course of the experiment triggered by GNMT output.

In addition, based on a number of syntactic priming monolingual and bilingual studies showing learning trends in the data, we predict that the increase in production of NP structure will occur at later trials regardless of participants' English proficiency levels.

Results

Table 1 displays the percentage of structures produced by participants in both the pre-test and priming phases. As predicted, in the pre-test phase, participants produced 59.7% of constructions in PNP form while in the priming phase this percentage drops to 38.7% (a difference of 21%). In the priming phase, the average percentage production of NP structures

increased by 26.7% and other structures (such as *the house's door is yellow* or with the incorrect order of the noun phrase such as **the door house is yellow*) decreased by 5.6%.

Table 2 displays the model that best explains our data set. The intercept estimate is negative, meaning that NP structures in the priming phase were more frequent than in the pre-test baseline. The model shows a significant priming effect ($p < 0.001$) as well as an effect of cumulative NP proportion ($p = 0.01$) to express the relationship of possession in the second language. This effect suggests that the amount of NP structures previously produced by the participants influenced the probability of producing NP structures in the subsequent utterances. This effect indicates a learning process taking place as NP production increases over time. This increase in the proportion of NP structures produced over time in the priming phase is demonstrated in Figure 5.

To investigate if the learning process varied as a function of participants' English proficiency levels, we tested the interaction between factors *cumulative NP proportion* and *English proficiency test score*. As we can see from Table 2, we did not find a significant interaction between these two factors ($p=0.5$), suggesting that, at all language proficiency levels, the amount of NP structures previously produced by the participants can influence the probability of producing NP structures in subsequent utterances.

It is possible to observe in Figure 4 that this increase in the proportion of NP structures occurs at all language proficiency levels, except for a few highly proficient participants with an English test grade above 20 who produced NP structures from the beginning to the end of the experiment in both the baseline and priming phases. This therefore explains the lack of interaction between factor *English proficiency test score* and *cumulative NP proportion*, suggesting that learning and accommodation to a challenging structure can occur at all English proficiency levels.

Table 1. Count and percentages of structures produced by 30 participants in the pre-test and priming phases with 20 trials each.
PNP = prepositional noun phrase,
NP = non-prepositional noun phrase.

Structures produced	Pretest	Priming
PNP	358 (59.7%)	232 (38.7%)
NP	168 (28%)	328 (54.7%)
Other	74 (12.3%)	40 (6.7%)
Total	600 (100%)	600 (100%)

Discussion and conclusion

Fully in line with our predictions, the effects observed in the present study clearly show that the GNMT system influences the processing of English syntax and that this influence is a consequence of cumulative exposure to the syntactic structure. Thus, our results answer in the affirmative RQ1 (*When interacting with Google translate for language production purposes, can the output of Google translate facilitate the processing of more challenging structures in the second language?*) and RQ2 (*Can learning emerge from the interaction between users and the GNMT system through continuous exposure to that structure via MT output?*). Such an influence is demonstrated in our results when a structure in English that is challenging for Brazilian Portuguese native speakers to process becomes more frequent in their subsequent speech after continuous exposure to the GNMT output with that structure. Thus, the syntactic priming effect observed in the present experiment is in line with previous results of studies investigating syntactic priming

Table 2. Summary of the best mixed effects logit model for participants' structure choices. Number of obs: 1200; groups: subjects = 30, items = 20. NP = non-prepositional noun phrase.

Random effects				
Group name	Variance	Std.Dev.		
Subject (Intercept)	5.767e-01	0.7593928		
Items (Intercept)	4.326e-10	0.0000208		
Fixed effects				
	Estimate	Std. Error	z value	Pr(>—z—)
Intercept (PNP structures or other)	-0.99042	0.18402	-5.382	7.35e-08 ***
English test grade	0.67375	0.17795	3.786	0.000153 ***
Cumulative NP structures	0.48263	0.10650	4.532	5.85e-06 ***
Priming test phase	113.027	0.16352	6.912	4.77e-12 ***
English test grade: cumulative	-0.06264	0.09733	-0.644	0.519863

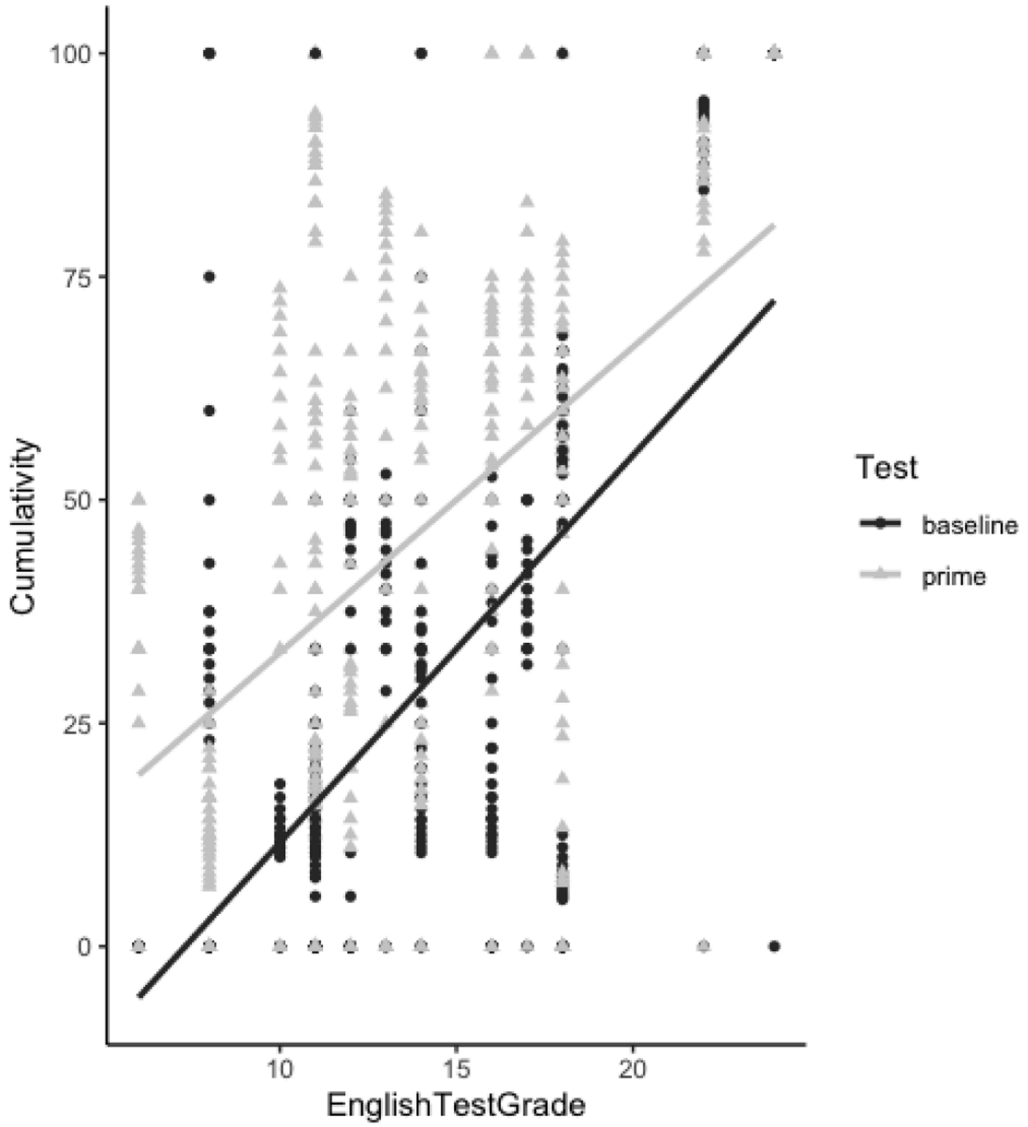


Figure 4. Cumulativity of non-prepositional noun phrase (NP) responses throughout experimental trials in both the baseline and priming phases. The proportion of NP responses produced increases over the course of the priming phase for all English proficiency levels (except for one outlier highly proficient participant who used the NP structure from the beginning until the end of the experimental session).

between humans and artificial systems[15] as well as in line with theoretical assumptions of surprisal and cumulativity[17], which predict that long-term priming results in a change in participants' syntactic preference. The less frequent or the most challenging structure (in this case NP structures) becomes more frequent in the course of the experiment.

As we can observe from Table 1 and Figure 5, there is a difference in participants' syntactic choice between the pre-test (without any influence of the MT output) and priming test. In the priming test phase, while the amount of PNP structures decreases, the amount of NP structures increases. Therefore, just like the surprisal and cumulativity theories claim, the learning mechanism was triggered by a less frequent structure and through continuous exposure to this structure the repetition effect increased. Therefore, our results suggest that an implicit learning mechanism can be activated in the second language by the syntactic repetition of an unusual structure in line with previous studies using a similar methodological paradigm[20,30,31]. Regarding RQ3 (*Does this learning vary as a function of participants' English proficiency?*), our results suggest that this phenomenon can occur both at lower levels of English proficiency and at more advanced levels.

Our results can bring insights to the MT and HCI fields as they show that interaction with an MT system could help students

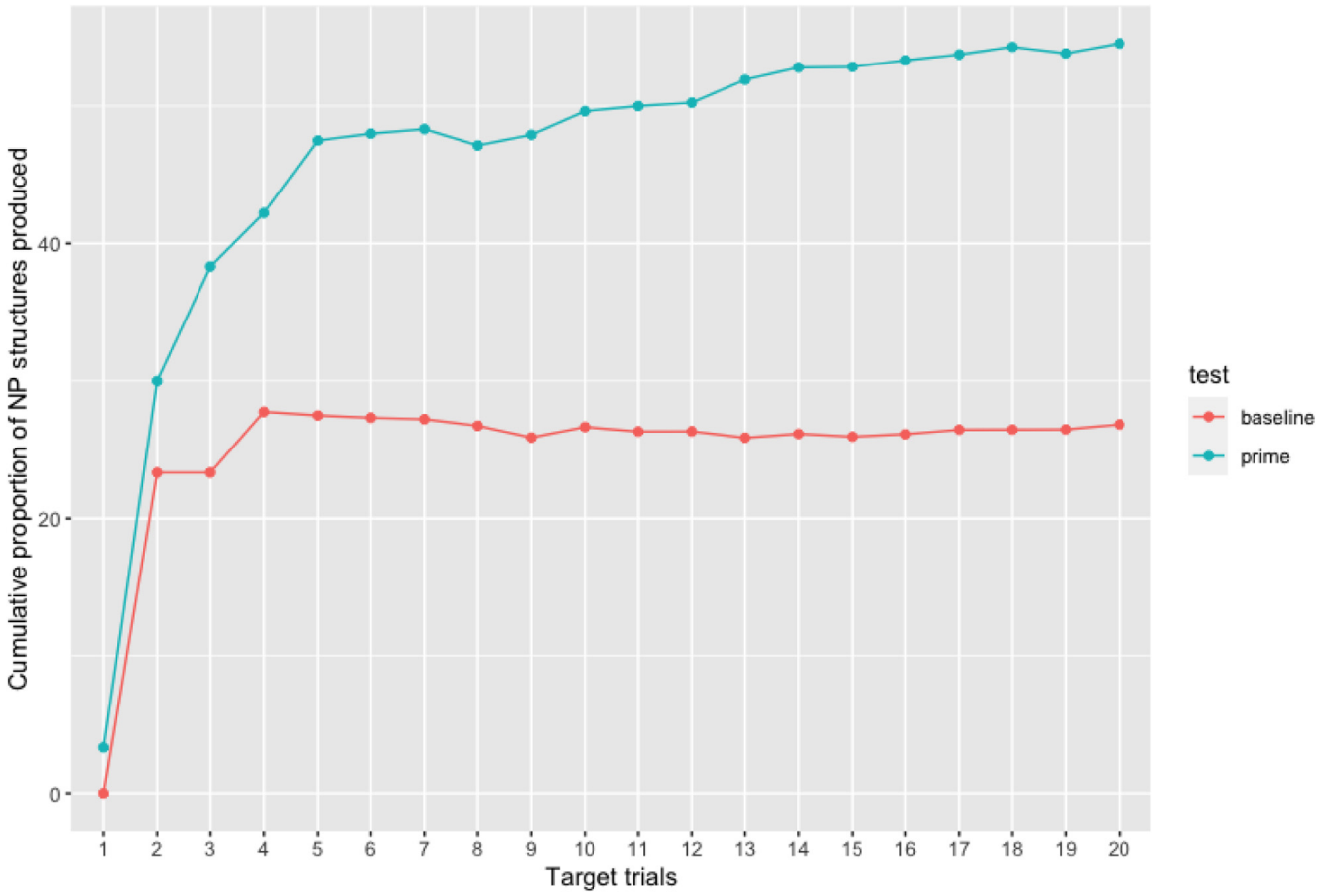


Figure 5. Cumulativity of non-prepositional noun phrase (NP) responses throughout experimental trials in both the baseline and priming phases. As compared to the baseline test, the proportion of NP responses produced increases over the course of the priming phase (from target 1 to target 20).

internalise and unconsciously learn a difficult structure in the second language regardless of English proficiency levels. In addition, it is also curious to see that, regardless of the easily observable gaps in the translation quality of raw MT output, speakers trust Google translate enough to use the same syntactic structure seen in the MT output in their speech and learn from it. These results add to the growing body of evidence that speakers tend to align their linguistic behaviour with their conversational partners not only in in human-human interactions, but also in human-computer interactions in a cross-linguistic task.

We question, nevertheless, if the same effect could be observed in the interaction between a less popular MT system and humans. Within the syntactic priming literature, research has demonstrated that the social opinion one has of the interlocutor shapes the priming effect. Thus, it might be that less popular MT systems could fail to elicit syntactic priming effects or elicit less robust priming effects compared to GNMT. In future research we aim to address this question.

In future studies we also plan to investigate in depth the implicit nature of the learning effect observed here by testing if the cumulative effect emerges from different structures using less popular MT systems. If so, this would suggest that MT systems play a relevant role in the cognitive mechanism of sentence processing in the second language.

In summation, our results allow us to conclude that the syntactic priming paradigm represents an ecologically valid method to study MT-human interaction as well as the impact of MT in second language learning as our results replicate findings of a number of previous studies in the fields of HCI, SLA and Psycholinguistics. Thus, our study opens the possibility of addressing other research questions in MT-human interaction using the same methodological paradigm.

Data availability

Underlying data

Zenodo: Dataset MTrill project. https://doi.org/10.5281/zenodo.5138028[49].

This project contains the following underlying data:

- data_30_subj cumulativity.csv (anonymised data for 30 participants)

Extended data

Zenodo: Dataset MTrill project. https://doi.org/10.5281/zenodo.5138028[49].

This project contains the following extended data:

- The 26 images and sentences used in the baseline phase (pre-test phase): 20 items used in the trials of interest and six items used as filler trials (in PDF format)
- The 78 images and sentences used in the Priming phase: 60 items used in the trials of interest and 18 items used as filler trials (in PDF format)
- translated_sentences_experiment_MTrill.xlsx (All sentences used to test Google Translation before running the experiments)
- stimuli_noun_phrase.docx and Noun_phrases_2.docx (Files with sentences used to test Google Translate prior choosing the sentences used in the experiment)
- MTrill.psyexp_version2.psyexp copy and images_prime_target_version2 copy.xlsx (Files used to run the experiment on Psychopy software)

Data are available under the terms of the Creative Commons Attribution International license (CC-BY 4.0).

References

1. McCulloch WS, Pitts W: **A logical calculus of the ideas immanent in nervous activity.** *Bull Math Biophys.* 1943; **5**(4): 115–133.
Publisher Full Text
2. Collins AM, Loftus EF: **A spreading-activation theory of semantic processing.** *Psychol Rev.* 1975; **82**(6): 407–428.
Publisher Full Text
3. Smith R: **Russia 2018: the Google Translate World Cup**. (visited on 10/07/2020).
Reference Source
4. Groves M, Mundt K: **Friend or foe? Google Translate in language for academic purposes.** *English for Specific Purposes.* 2015; **37**: 112–121.
Publisher Full Text
5. Garcia I, Pena MI: **Machine translation-assisted language learning: writing for beginners.** *Comput Assist Lang Learn.* 2011; **24**(5): 471–487.
Publisher Full Text
6. Clifford J, Merschel L, Munné J: **Surveying the Landscape: What is the Role of Machine Translation in Language Learning?** *Revista d'innovacio educativa.* 2013; **10**: 108–121.
Publisher Full Text
7. Tin TB: **Exploring the Development of 'Interest' in Learning English as a Foreign/Second Language.** *RELC Journal.* 2013; **44**(2): 129–146.
Publisher Full Text
8. Seidlhofer B: **English as a lingua franca.** *ELT Journal.* 2005; **59**(4): 339–341.
Publisher Full Text
9. Weber K, Luther L, Indefrey P, *et al.*: **Overlap and Differences in Brain Networks Underlying the Processing of Complex Sentence Structures in Second Language Users Compared with Native Speakers.** *Brain Connect.* 2016; **6**(4): 345–355.
PubMed Abstract | Publisher Full Text
10. Gass S, Mackey A, Ross-Feldman L: **Task-Based Interactions in Classroom and Laboratory Settings.** *Lang Learn.* 2005; **55**(4): 575–611.
Publisher Full Text
11. Bock JK, Kroch AS: **The Isolability of Syntactic Processing.** In: *Linguistic Structure in Language Processing.* Ed. by Greg N. Carlson and Michael K. Tanenhaus. Dordrecht: Springer Netherlands, 1989; 157–196.
Publisher Full Text
12. Bock JK: **Meaning, sound, and syntax: Lexical priming in sentence production.** *J Exp Psychol Learn Mem Cogn.* 1986; **12**(4): 575–586.
Publisher Full Text
13. Bock JK: **Syntactic persistence in language production.** *Cogn Psychol.* 1986; **18**(3): 355–387.
Publisher Full Text
14. Bock K, Loebell H, Morey R: **From conceptual roles to structural relations: Bridging the syntactic cleft.** *Psychol Rev.* 1992; **99**(1): 150–171.
PubMed Abstract | Publisher Full Text
15. Heyselaar E, Hagoort P, Segaert K: **How social opinion influences syntactic processing-An investigation using virtual reality.** *PLoS One.* 2017; **12**(4): e0174405.
PubMed Abstract | Publisher Full Text | Free Full Text
16. Heyselaar E, Hagoort P, Segaert K: **In dialogue with an avatar, language behavior is identical to dialogue with a human partner.** *Behav Res Methods.* 2017; **49**: 46–60.
Publisher Full Text
17. Florian Jaeger T, Snider N: **Implicit Learning and Syntactic Persistence: Surprisal and Cumulativity.** In: *University of Rochester Working Papers in the Language Sciences.* 2007; **3**(1): 26–44.
Reference Source
18. Cowan BR, Branigan HP, Obregón M, *et al.*: **Voice anthropomorphism, interlocutor modelling and alignment effects on syntactic choices in human–computer dialogue.** *Int J Hum Comput Stud.* 2015; **83**: 27–42.
Publisher Full Text
19. Resende N, Cowan B, Way A: **MT syntactic priming effects on L2 English speakers**. In: *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation.* 2020; 245–253.
Reference Source
20. Shin JA, Christianson K: **Structural Priming and Second Language Learning.** *Lang Learn.* 2012; **62**(3): 931–964.
Publisher Full Text
21. Resende N, Way A: **Can Google Translate Rewire Your L2 English Processing?** *Digital.* 2021; **1**(1): 66–85.
Publisher Full Text
22. Vilela AC, Oliveira FLP: **"I cut my hair" e "I did my nails": evidência de transferência linguística na interlíngua de falantes brasileiros aprendizes de inglês como segunda língua?** *Trabalhos de Linguística Aplicada.* Campinas, 2010; **49**(1): 223–239.
Publisher Full Text
23. Zara JV, Oliveira FLPD, Souza RAD: **Selective transfer in the acquisition of English double object constructions by Brazilian learners.** *Alfa: Revista de Linguística (São José do Rio Preto).* 2013; **57**: 519–544.
Publisher Full Text
24. Ghilzai SA: **The role of L1 transfer in the processing of L2 English possessive constructions: evidence from German.** *Eur Academ Res.* 2014; **2**: 7454–7481.
25. Kopp S: **Social resonance and embodied coordination in face-to-face conversation with artificial interlocutors.** In: *Speech Communication.* Speech and Face-to-Face Communication, 2010; **52**(6): 587–597.
Publisher Full Text
26. Segaert K, Menenti L, Weber K, *et al.*: **A Paradox of Syntactic Priming: Why Response Tendencies Show Priming for Passives, and Response Latencies Show Priming for Actives.** *PLoS One.* 2011; **6**(10): e24209.
PubMed Abstract | Publisher Full Text | Free Full Text
27. Traxler MJ: **Structural priming among prepositional phrases: Evidence from eye movements.** *Mem Cognit.* 2008; **36**(3): 659–674.
PubMed Abstract | Publisher Full Text
28. Schoot L, Hagoort P, Segaert K: **Stronger Syntactic Alignment in the Presence of an Interlocutor.** *Front Psychol.* 2019; **10**: 685.
Publisher Full Text
29. Pickering MJ, Ferreira VS: **Structural priming: A critical review.** *Psychol Bull.* 2008; **134**(3): 427–459.
PubMed Abstract | Publisher Full Text | Free Full Text

30. Hartsuiker RJ, Westenberg C: **Word order priming in written and spoken sentence production.** *Cognition.* 2000; **75**(2): B27–B39.
PubMed Abstract | **Publisher Full Text**

31. Hartsuiker RJ, Kolk HH: **Syntactic Persistence in Dutch.** *Lang Speech.* 1998; **41**(Pt 2): 143–184.
PubMed Abstract | **Publisher Full Text**

32. Heyselaar E, Segaert K, Walvoort SJ, *et al.*: **The role of procedural memory in the skill for language: Evidence from syntactic priming in patients with amnesia.** In: 2016.
Reference Source

33. Kim Y, McDonough K: **The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners**. *Lang Teach Res.* 2008; **12**(2): 211–234.
Publisher Full Text

34. McDonough K: **Interaction and syntactic priming: English L2 speakers' production of dative constructions.** *Stud Second Lang Acquis.* 2006; **28**(2): 179–207.
Publisher Full Text

35. McDonough K, Mackey A: **Syntactic priming and ESL question development.** *Stud Second Lang Acquis.* 2008; **30**(1): 31–47.
Publisher Full Text

36. Corley M, Scheepers C: **Syntactic priming in English sentence production: Categorical and latency evidence from an Internet-based study.** *Psychon Bull Rev.* 2002; **9**(1): 126–131.
PubMed Abstract | **Publisher Full Text**

37. Smith S, Wheeldon L: **Syntactic priming in spoken sentence production - an online study.** *Cognition.* 2001; **78**(2): 123–164.
PubMed Abstract | **Publisher Full Text**

38. Wheeldon L, Smith M: **Phrase structure priming: A short-lived effect.** *Lang Cogn Process.* 2003; **18**(4): 431–442.
Publisher Full Text

39. Oviatt S, Darves C, Coulston R: **Toward Adaptive Conversational Interfaces: Modeling Speech Convergence with Animated Personas.** *ACM Trans Comput-Hum Interact.* 2004; **11**(3): 300–328.
Publisher Full Text

40. Branigan HP, Pickering MJ, Pearson J, *et al.*: **Syntactic alignment between computers and people: the role of belief about mental states.** English. In: *Proceedings of the 25th Annual Conference of the Cognitive Science Society, July 31 - August 2 2003, Boston, Massachusetts.* Ed. by Richard Alterman and David Kirsh. Lawrence Erlbaum Associates, 2003; 186–191.
Reference Source

41. Suzuki N, Katagiri Y: **Prosodic alignment in human-computer interaction.** *Connection Science.* 2007; **19**(2): 131–141.
Publisher Full Text

42. Bergmann K, Branigan HP, Kopp S: **Exploring the alignment space – lexical and gestural alignment with real and virtual humans.** *Frontiers in ICT.* 2015.
Publisher Full Text

43. Stoyanchev S, Stent A: **Lexical and syntactic priming and their impact in deployed spoken dialog systems**. en. In: *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume: Short Papers on - NAACL '09.* Boulder, Colorado: Association for Computational Linguistics, 2009; 189.
Publisher Full Text

44. van Gompel RPG, Arai M: **Structural priming in bilinguals.** *Bilingualism: Language and Cognition.* 2018; **21**(3): 448–455.
Publisher Full Text

45. Hartsuiker RJ, Pickering MJ, Veltkamp E: **Is Syntax Separate or Shared between Languages? Cross-Linguistic Syntactic Priming in Spanish-English Bilinguals.** *Psychol Sci.* 2004; **15**(6): 409–414.
PubMed Abstract | **Publisher Full Text**

46. McDonough K, Kim Y: **Syntactic Priming, Type Frequency, and EFL Learners' Production of *Wh*-Questions.** *The Modern Language Journal.* 2009; **93**(3): 386–398.
Publisher Full Text

47. McDonough K, Mackey A: **Responses to Recasts: Repetitions, Primed Production, and Linguistic Development.** *Lang Learn.* 2006; **56**(4): 693–720.
Publisher Full Text

48. Florian Jaeger T, Snider NE: **Alignment as a consequence of expectation adaptation: syntactic priming is affected by the prime's prediction error given both prior and recent experience.** *Cognition.* 2013; **127**(1): 57–83.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

49. Resende N: **Dataset MTrill project.** *Zenodo.* 2021.
http://www.doi.org/10.5281/Zenodo.5138028

50. R R Core Development Team: **R: A language and environment for statistical computing**. Viena, Austria, 2011.

51. Bates D, Mächler M, Dai B: **lme4: Linear Mixed-Effects Models Using S4 Classes.** 2011; 0.999375-33.

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 2

Reviewer Report 15 February 2024

https://doi.org/10.21956/openreseurope.18312.r36793

© **2024 Carl M et al.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Michael Carl**
Kent State University, Kent, OH, USA
Longhui Zou
Kent State University, Kent, OH, USA

I read the paper and found it to be nicely written and structured, complete review of literature, appropriate description of the experiment and impactful conclusion.

I think it can be indexed as is.

Competing Interests: No competing interests were disclosed.

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 04 February 2022

https://doi.org/10.21956/openreseurope.14986.r28381

© **2022 Carl M et al.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

? **Michael Carl**
[1] Kent State University, Kent, OH, USA
[2] Kent State University, Kent, OH, USA

Longhui Zou
[1] Kent State University, Kent, OH, USA
[2] Kent State University, Kent, OH, USA

This paper investigates effects of syntactic priming for L2 learner using Google Neuro Machine Translation (GNMT). It addresses several topics including implicit language learning, process of syntactic priming, human-computer interaction (HCI), computer assisted second language learning (SLA). The introduIdction and background give an informative overview of current research on syntactic priming in the fields of HCI and SLA. Different from previous studies, this paper investigates the cumulative effect of syntactic priming on SLA via GNMT output.

The paper presents an empirical study with a total of 30 participants (native Brazilian Portuguese, L2 English speakers) translate a Portuguese prepositional noun phrase (PNP) into English. Portuguese only allows for the PNP structure, while in English it can be expressed by either PNP: "The *window of the office* is broken", or a noun phrase structure, NP: "The *office window* is broken". It is assumed that the less proficient English L2 speakers will more often (re)produce the PNP structure in English in an unprimed pre-test phase of the study', but when primed with the GNMT output (which produces more often NP than PNP) they would learn and prefer the more common English NP rendering. Using a "logistic mixed effects maximal model", the results show that, indeed, NP structures were more frequently used by the participants in the priming phase than the pre-test phase. The authors conclude that GNMT can facilitate the processing of more challenging structures in the second language and can trigger implicit learning of a second language by syntactic priming effect. The study also finds that implicit learning occurs at both lower levels and more advanced levels of English proficiency.

Overall, the methods, materials and procedures of the experiment are clearly illustrated, and the statistical analysis are in sufficient details, which can largely benefit replication and future studies.

However, there are still certain space for improvement:

- The study investigates the learning process of challenging syntactic structures for L2 English learners with GNMT, but there is inadequate literature review or empirical evidence to support that NP production is more challenging or complex to the participants than PNP production in their L2.

- On page 7, the author mentions that one of the inclusion criteria of the participants was "to be at intermediate or advanced English levels", but the participants' English proficiency levels were tested after completing both the pre-test and priming test phases. This sequence of experimental design seems unclear to me.

- The authors state that in their manual coding, "we coded the baseline as "1" if participants produced a NP structure, suggesting they were primed by the GT output and "0" if participants produced a PNP structure". According to my understanding of the research procedures, in the baseline phase, the participants did not use GNMT, so this paragraph is somewhat confusing.

- I have some concern about the priming test phase, during which the participants were asked to translate the prime sentences using GT application on their own mobile device. There might be certain possibilities that the MT output of the same priming sentence would

be different across the 30 participants during the experiment, which is undesirable and should be controlled.

It is unclear what exactly the study wants to test. The RQ1: “Can Google translate facilitate the processing of more challenging structures in the second language?” is somewhat not really a ‘research’ question, given that we know google translates every day billons (or so) words (and structures) for millions of users, who would otherwise not be able to process the second language at all. When reading, the paper appears to investigate: 1) the syntactic priming effect when translating Portuguese PNP or 2) how much is MT helpful for L2 learners. The paper provides plenty of evidence for syntactic priming effects (i.e. 1) from previous research, but, unfortunately does not address 2) very much (where it is not identical to 1). Only in the conclusion the authors argue that “regardless of the easily observable gaps in the translation quality of raw MT output, speakers trust Google translate enough to use the same syntactic structure seen in the MT output in their speech and learn from it”. Ah!

- it should be made clearer, from the beginning, that the texts were spoken. Would they speak the texts into their mobile phones? Was the prime also an auditory signal?

- It might be interesting to quantify MT quality and include the variable in the data analysis part, to investigate whether the difference in MT quality of the priming sentences would affect the NP production or SLA.

- A quite different experimental set-up should be used to assess ‘trust’ in the MT output, with different MT systems, different baselines, and maybe also translationese.

While the study shows “that the GNMT system influences the processing of English syntax”, it is unclear to what extent this is due to the particular brand, how MT errors impact the learning effect, and what effect might have the mode (spoken vs. written translation).

Minor issues:

- Use only one of the two acronyms GT or GNMT.

- Figures do not show the pictures.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Author Response 04 Dec 2023

NATALIA RESENDE

First, I want to express my sincere gratitude for the thorough review of this paper, and I truly value all the insightful comments that have been provided. Your feedback has proven to be immensely valuable in enhancing the overall quality of the paper. In my response to your comments, I have diligently addressed each one individually. Please find below my responses to each of the comments :

- The study investigates the learning process of challenging syntactic structures for L2 English learners with GNMT, but there is inadequate literature review or empirical evidence to support that NP production is more challenging or complex to the participants than PNP production in their L2. Thank you for bringing up this point. I agree that the paper did not provide specific evidence supporting the claim that NP (noun phrase) production is more challenging or complex for the participants in their L2 (second language) compared to PNP (post-nominal possessive) production. As far as I'm aware, there is no existing evidence to support this assertion.

However, in the revised version of the paper, I have addressed this issue by expanding a paragraph where I explain that NP production is considered more challenging for BP (Brazilian Portuguese) speakers. This assertion is based on previous studies conducted for other languages, which have shown that when a syntactic structure in the L2 does not exist in the L1 but there is a syntactic alternative in the L2 that closely matches the L1 structure, speakers tend to choose the one that is closer to the L1 structure. This preference arises because it represents an easier association due to L1 transfer processes.

The expanded paragraph is provided below: In Portuguese, only one syntactic alternative exists to represent a relationship of possession between nouns. The relationship is always encoded in the preposition do (de + o) or da (de + a) (e.g. a mesa do escritório está cheia or a porta da casa está fechada). However, in English, this relationship can be represented using either a prepositional noun phrase (PNP), which follows the same word order as in Portuguese (e.g. the table of the office is full) or a non-prepositional noun phrase (NP) (e.g. the office table is full), which differs from Portuguese in word order. Given that NP structures do not exist in Portuguese, we posit that they pose more difficulty for native Brazilian Portuguese speakers. As a result, these speakers might produce more PNP sentences due to their similarity with Portuguese syntax that may be transferred to their second language. While no research indicates that producing NP structures in English is more difficult for Brazilian Portuguese speakers than PNP, we assumed it is a more challenging structure because some studies (e.g. Vilela et al, 2010, Zara et al., 2013) with Brazilian Portuguese speakers of English as L2 have shown that when a syntactic structure in the L2 does not exist in the L1 but there is in the L2 a syntactic alternative that matches the L1 structure speakers tend to choose the one that is closer to the L1 structure because it

is an easier association due to L1 transfer processes. Specifically for the syntactic structure to express a relation of possession, the study from Ghilzai (2014) shows that, when rating preferred English sentences with possessive structures, German speakers of English as L2 often rate higher the sentences using the PNP (of-genitive) with an animate subject instead of the s-genitive case because, differently from English, the s-genitive in German is only used with proper nouns. Therefore this study is another evidence for transfer trends interfering in the L2 process. Hence, studying how Brazilian Portuguese speakers process these structures in English can help determine if syntactic priming by the GNMT output prompts them to use the NP format more frequently in English. • On page 7, the author mentions that one of the inclusion criteria of the participants was “to be at intermediate or advanced English levels”, but the participants’ English proficiency levels were tested after completing both the pre-test and priming test phases. This sequence of experimental design seems unclear to me. I clarify this in the paper as can be read below: We analysed data from 30 volunteer Brazilian Portuguese L2 English speakers (10 men, mean age=35.7 - sd=5.3). Participants were recruited through posts on Facebook groups of Brazilians living in Dublin and word-of-mouth recommendations. All participants were requested to read a plain language statement and sign the informed consent form to take part in the experiment. The inclusion criteria to take part in the study was to be a native speaker of Brazilian Portuguese, to live in Dublin at the time of the data collection, to have used GNMT as a tool supporting spoken English and be currently enrolled or have previously attended an English school in Dublin, with English proficiency at either intermediate or advanced levels. If not actively enrolled, participants had to share their most recent proficiency level upon exiting the English school. In order to confirm participants’ English proficiency levels we asked them to complete the General English test online with 25 questions immediately after completing the pre-test and priming test phases. Participants were classified at the basic level (level A2) with scores ranging from 2 to 13; pre-intermediate level (B1) scores from 14 to 17; intermediate level (B2) scores from 18 to 19; advanced level scores from 20 to 22 and proficient level scores from 23 to 25. Participants reporting an intermediate level from their English school but scoring at the basic level on our test were excluded. The mean score of the participants included in the experiment was 14.1 (SD=4.7) or level B1). In return for taking part in the experiment, all participants received a €10 voucher.

• The authors state that in their manual coding, “we coded the baseline as ”1” if participants produced a NP structure, suggesting they were primed by the GT output and ”0” if participants produced a PNP structure”. According to my understanding of the research procedures, in the baseline phase, the participants did not use GNMT, so this paragraph is somewhat confusing. Thanks for pointing out this oversight of my part. I have reviewed this paragraph by removing the sentence “suggesting they were primed by the GT output” Reviewed version: Participants’ responses for both baseline and priming tests were manually coded by the experimenter. In the baseline phase, responses with an NP structure were coded as "1", while those with a PNP structure received a "0". For the priming phase, if participants used the same structure in the target item as seen in the GNMT output (specifically, NP structures) to describe the images, it was coded as "yes" (1). If they used a PNP structure or another variant like "the office’s window is broken" or "the window in the office is broken", it was coded as "no" (0).

• I have some concern about the priming test phase, during which the participants were asked to translate the prime sentences using GT application on their own mobile device. There might be certain possibilities that the MT output of the same priming sentence would

be different across the 30 participants during the experiment, which is undesirable and should be controlled.
Yes, I was aware that some variation in the output of GT could occur. In the first version, I mistakenly did not add a paragraph in this article to explain what was the procedure to address this issue. However, this issue was addressed in this new version as below: It is important to note that when we created the prime sentences, they had all been machine-translated using GNMT on various days and different times spanning two months. We observed that the sentence structures remained consistent over time as Google consistently translated all sentences using an NP structure. We were aware, however, that despite this consistency over time, GNMT output could still vary across experiment trials during sessions. However, because we noticed that GNMT's preferred structure is an NP structure, we thought it would be worth testing whether GNMT output would be capable of eliciting priming effects even if in a few cases the system outputs an alternative structure. It is unclear what exactly the study wants to test. The RQ1: "Can Google translate facilitate the processing of more challenging structures in the second language?" is somewhat not really a 'research' question, given that we know google translates every day billons (or so) words (and structures) for millions of users, who would otherwise not be able to process the second language at all. When reading, the paper appears to investigate: 1) the syntactic priming effect when translating Portuguese PNP or 2) how much is MT helpful for L2 learners. I reworded the RQ1 (see below). Hope it is clearer now: "When interacting with Google translate for language production purposes, can the output of Google translate facilitate the processing of more challenging structures in the second language?"
The paper provides plenty of evidence for syntactic priming effects (i.e. 1) from previous research, but, unfortunately does not address 2) very much (where it is not identical to 1). Only in the conclusion the authors argue that "regardless of the easily observable gaps in the translation quality of raw MT output, speakers trust Google translate enough to use the same syntactic structure seen in the MT output in their speech and learn from it".
• it should be made clearer, from the beginning, that the texts were spoken. Would they speak the texts into their mobile phones? Was the prime also an auditory signal?

I improved the wording when describing the procedures to clarify that the prime sentences were presented in the written form but the translation was carried out verbally • It might be interesting to quantify MT quality and include the variable in the data analysis part, to investigate whether the difference in MT quality of the priming sentences would affect the NP production or SLA. it's worth noting that throughout our experimental sessions, we consistently observed Google Translate using the NP (noun phrase) structure. Consequently, this particular variable didn't appear to have any noticeable impact on the results. In fact, it was precisely this consistency that gave rise to the observed priming effect. • A quite different experimental set-up should be used to assess 'trust' in the MT output, with different MT systems, different baselines, and maybe also translationese. While the study shows "that the GNMT system influences the processing of English syntax", it is unclear to what extent this is due to the particular brand, how MT errors impact the learning effect, and what effect might have the mode (spoken vs. written translation). This is a good suggestion for future research. This is one of the questions that I want to address in future experiments.
Minor issues: • Use only one of the two acronyms GT or GNMT. Acronym issue fixed

- Figures do not show the pictures

Competing Interests: No competing interests were disclosed.

Reviewer Report 19 October 2021

https://doi.org/10.21956/openreseurope.14986.r27633

© 2021 Oliveira C. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Cândido Oliveira
[1] Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, Brazil
[2] Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, Brazil

Thank you for the opportunity to review **Investigating syntactic priming cumulative effects in Mt-human interaction**. The study and its findings are very interesting. I hope the following feedback will provide helpful guidance in shaping the manuscript.

The article presents a study that aims to investigate the interaction between Google neural machine translation system (GNMT) and English language learners whose first language is Brazilian Portuguese (BP). The author conducted a syntactic priming experiment with a pre-test-priming design to analyze whether the interaction with GNMT could lead these participants to use the structure that is the most preferred one among English native speakers to indicate a relation of possession (the book cover is red) more often than the structure whose equivalent in Portuguese is the most preferred one among BP native speakers (the cover of the book is red / *a capa do livro é vermelha*). The results indicate that participants (i) tended to use the former structure to form a sentence after they had just been exposed to it in a translation process (priming effect) and (ii) this tendency increased as they were continuously exposed to this structure (cumulative effect). The author suggest that these results indicate that GNMT can influence learners' behavior in the L2 as well as their learning process.

The text reads well and has a good organization. The references are relevant and up to date. The theoretical and applied motivations are clearly established. I found the justifications to investigate MT-human interaction through the use of google translate by language learners very interesting. Moreover, the methodology, especially the method, and well justified and described in great details. I have only three suggestions to the author, which I present below:

On page four, the author presents two studies that seem to be part of the same project this article is inserted in. In order for the readers to have a better picture of these previous articles, I recommend the author includes information about the participants proficiency level. Also, it is not clear if there was any empirical evidence that participants indeed had a tendency to use the PNP instead of the NP structure (the present article does that) before the experiments were carried out.

I was under the impression that the low number of distractors in the present study could have had an effect on the results. I suggest the author include a justification for the number of distractor items utilized in the experiment and/or a discussion about it. Other studies using priming, including some that were cited by the author (ex: Hartsuiker & Westenberg, 2000) use a much higher number of distractors.

The use of psycholinguistic methods to bridge the gap between second language acquisition theories and language learning itself has been growing exponentially in recent years and this study adds to this process. Even though, it is not the focus of the study, I believe the author could include some discussions about the implications of the results to the field of Second Language Acquisition and to language learning in general.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Bilingualism, Psycholinguistics and Second Language Acquisition

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.