

# Translation and Language Change with Reference to Popular Science Articles

Malamatidou, Sofia

*License:*

None: All rights reserved

*Document Version*

Peer reviewed version

*Citation for published version (Harvard):*

Malamatidou, S 2011, Translation and Language Change with Reference to Popular Science Articles: The interplay of Diachronic and Synchronic Corpus-Based Studies. in H Hedeland, T Schmidt & K Worner (eds), *Multilingual Resources and Multilingual Applications: Proceedings of the Conference of the German Society for Computational Linguistics and Language Technology (GSCL) 2011*. vol. 96, University of Hamburg, Hamburg, pp. 135-140, Conference of the German Society for Computational Linguistics and Language technology, Hamburg, Germany, 26/09/11.

[Link to publication on Research at Birmingham portal](#)

**General rights**

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

**Take down policy**

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

# **Translation and Language Change with Reference to Popular Science Articles: The Interplay of Diachronic and Synchronic Corpus-Based Studies**

**Sofia Malamatidou**

University of Manchester

Oxford Road, M13 9PL

E-mail: sofia.malamatidou@manchester.ac.uk

## **Abstract**

Although a number of scholars have adopted a corpus-based approach to the investigation of translation as a form of language contact and its impact on the target language (Steiner, 2008; House, 2004; 2008; Baumgarten et al. 2004), no sustained corpus-based study of translation involving Modern Greek has so far been attempted and very few diachronic corpus-based studies (Amouzadeh & House, 2010) have been undertaken in the field of translation. This study aims to combine synchronic and diachronic corpus-based approaches, as well as parallel and comparable corpora for the analysis of the linguistic features of translated texts and their impact on non-translated ones. The corpus created captures a twenty-year period (1990-2010) and is divided into three sections, including non-translated and translated Modern Greek popular science articles published in different years, as well as the source texts of the translations. Unlike most studies employing comparable corpora, which focus on revealing recurrent features of translated language independently of the source and target language, this study approaches texts with the intention of revealing features that are dependent on the specific language pair involved in the translation process.

Keywords: corpus-based translation studies, language change, diachronic corpora, Modern Greek, passive voice

## **1. Introduction**

Translation as a language contact phenomenon is a phenomenon that neither linguistics nor translation studies has addressed in depth. However, in the era of the information society, the translation of popular science texts tends to be very much a unidirectional process from the dominant lingua franca, which is English, into less widely spoken languages such as Modern Greek. This process is likely to encourage changes in the communicative conventions of the target language. Given the fact that the genre of popular science was developed in Greece mainly through translations from Anglophone sources in the last two decades, it is interesting to examine whether and how the translations from English encouraged the dissemination of particular linguistic features in the target language in the discourse of this particular genre. A number of scholars, mostly within the English-German context, have taken interest in investigating translation as a form of language contact and its effects on the target language. Steiner (2008) has

investigated grammatical and syntactic features of explicitness as a result of the contact between English and German, which however did not involve diachronic analyses of corpora. Most importantly, House and a group of scholars have investigated how translation from English affects German, but also Spanish and French (House, 2004; 2008; Baumgarten et al. 2004; Becher et al. 2009). However, these studies mainly involved manual analyses of texts, that is, they were not corpus-based studies as they are understood by Baker (1995), i.e. they did not involve the automatic or semi-automatic analysis of machine-readable texts. Diachronic corpus-based approaches to translation are limited (Amouzadeh & House, 2010) and in terms of Modern Greek, no similar study has ever been conducted.

This study aims to examine whether and how translation can encourage linguistic changes in the target language by investigating a diachronic corpus of non-translated and translated Modern Greek popular science articles, along with their source texts, in order to examine how

translation can be understood as a language contact phenomenon. The linguistic change that is examined is the frequency of the passive voice, since it has been claimed to be found more frequently in translated Modern Greek texts (Apostolou-Panara, 1991), especially those translated from English.

This paper first presents the theoretical model that informs the study, namely the Code-Copying Framework (Johanson, 1993; 1999; 2002). Then the research methodology is presented in detail and data analysis techniques are analysed. Finally, some preliminary findings are discussed. It must be mentioned, that this is still an ongoing project and for that reason the results are limited to a number of small sample studies.

## 2. The Code-Copying Framework

The Code-Copying Framework is a widely applicable linguistic model that is suitable for the description of phenomena that have consistently been neglected, such as translation as a form of language contact and a propagator of change. Some of its concepts have recently been used by translation scholars to describe similar phenomena (Steiner, 2008), suggesting that it is a conceptual model suitable for analysing diverse cases of language contact, in particular cases where translation plays a central role in the dissemination of linguistic features.

The Code-Copying Framework was developed by Johanson (1993; 1999; 2002) who is critical of the terminology, especially that of borrowing, used in the field of language change studies and it is this critique that serves as a point of departure towards developing a new explanatory framework of language contact, where ‘copying’ replaces traditional terms and provides a different vintage point from which to analyse the phenomenon. Johanson (1999:39) argues that in any situation of code-interaction, that is, in a situation where two or more codes interact, two linguistic systems, i.e. two codes are employed. The Model Code is the source code, whereas the Basic Code is the recipient code which also provides the necessary morphosyntactic and other information for inserting and adapting the copied material (Johanson, 2008:62). Although, there are different directions of copying, this study focuses on the

case of ‘adoption’ which involves elements being inserted from the Model Code into the Basic Code and views translation as a language contact situation where translators are likely to copy elements from the source language, i.e. the Model Code, when translating into the target language, which is the Basic Code.

Two types of copying are possible within this model: global and selective copying. The linguistic properties that can be copied are material (i.e. phonic), semantic, combinational (i.e. collocations and syntax) and frequential properties, namely the frequency of particular linguistic units. In the case of global copying, a linguistic item is copied along with all its aforementioned properties. In the case of selective copying, one or more properties are copied resulting in distinct types of copying. Thus, there is material (M), semantic (S), combinational (C) and frequential (F) copying.

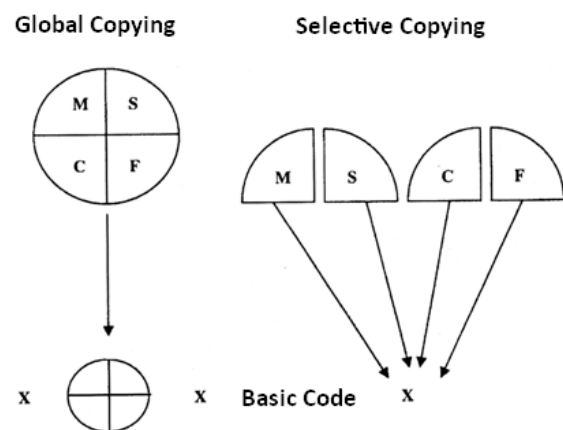


Figure 1: The Code-Copying Framework (Johanson, 2006:5)

During the process of translation, selective copying is more probable than global copying (Verschik, 2008:133). For that reason, the type of copying that is dealt with in this study is selective copying and in particular frequential copying, which results in a change in the frequency patterns of an existing linguistic unit. Apostolou-Panara (1991) notes that the passive constructions are used more frequently in Modern Greek than they once were. Traditionally, it has been argued that the passive voice structures are used in Modern Greek though not as often as in English (Warburton, 1975:576), where the passive voice is quite frequent especially in

terms of informative texts such as popular science articles. As far as translation is concerned, different frequencies and proportionalities of native patterns often result in texts having a ‘non-native’ feeling (Steiner, 2008:322). The frequent translation of source text patterns with grammatical, yet marginal, target language linguistic patterns may ultimately override prevailing patterns and result in new communicative preferences in the target language (Baumgarten & Özçetin 2008:294).

Copies usually begin as momentary code-copies, that is, occasional instances of copying. When copies start being frequently and regularly used by a group of individuals or by a particular speech community, they become habitualised code-copies. Copies may also become conventionalised and become integrated and widely accepted by a speech community. The final stage is for copies to become monolingual, i.e. when copies are used by monolinguals and do not presuppose any bilingual ability (Johanson, 1999:48). Since momentary copies are difficult to trace (Csató, 2002:326), emphasis in this study is placed on habitualised code-copies. Translators are considered as part of a particular speech community and copies are regarded as habitualised when they are frequently and regularly used by translators. Conventionalised copies are not examined in this study, since they presuppose measuring social evaluation that is outside the scope of this research. However, it is safe to assume that if a copy is monolingualised, that is, it is used in non-translated texts; it is also in general terms socially approved.

Translation in this study is understood as a social circumstance facilitating copying. It is not considered as a cause of change, but rather as an instance of contact during which copying may occur and change may proliferate through language, since translated texts, especially newspaper and magazine articles, are widely circulating texts that are likely to exert a powerful linguistic impact on a large audience. The main factors of copying are considered to be extra-linguistic, especially the cultural dominance of English in relation to Modern Greek, as far as the production of scientific texts is concerned, and the prestige that English enjoys as a prominent language and culture, both in the general sense of a lingua franca and in terms of scientific research.

### **3. Data and Methodology**

#### **3.1. Corpus design**

Based on the availability of data and the research aims of this thesis, an approximately 500,000 corpus of Modern Greek non-translated and translated popular science articles, along with their source texts was created. The corpus is named TROY (TRANslation Over the Years) and covers a 20-year period (1990-2010), which is considered to be an adequate time span for language change to occur and is amenable to being systematically observed.

Newspapers and magazines dedicated to scientific issues provide are the two main sources of popular science articles. The corpus is specialised in terms of both genre and domain, i.e. it involves popular science articles from the domain of technology and life sciences. These domains were chosen due to the fact that the majority of articles, especially translations, seem to belong to either one of the two domains. This in turn indicates that interest is expressed for these domains from the general public, which consequently suggests that a high number of people will read articles belonging to the domains of technology and life sciences, a fact that is likely to result in a powerful linguistic impact on a large audience.

The TROY corpus is divided into three subcorpora. The first subcorpus consists of non-translated Modern Greek popular science articles published in 1990-1991. The second subcorpus consists of non-translated and translated Modern Greek popular science articles published in 2003-2004, as well as the source texts of the translations. The years 2003-2004 were selected because translations of popular science texts started circulating more widely in Greece during that period than in previous years. The third subcorpus includes non-translated as well as translated texts and their source texts, all published in 2009-2010. The subcorpora are evenly balanced, both in terms of their overall size and between the two domains.

#### **3.2. Corpus Methodology**

The corpus methodology employed in this study has three aims. Firstly, it aims to investigate whether certain features have changed over time in Modern Greek. Secondly, it aims to examine whether this change is

related or mirrored in the process of translation. Finally, it aims to investigate whether influence can be traced back to the English source texts. Ultimately, this methodology aims at combining most corpus-based methodologies under one research aim. Thus, synchronic and diachronic corpus-based approaches, as well as parallel and comparable corpora are employed in order to illustrate the way in which combined methodologies can assist in the analysis of the linguistic features of translated texts and their impact on non-translated ones.

Firstly, the corpus methodology aims at examining language change in Modern Greek and in particular to investigate whether the frequency of the passive voice has changed over time. This involves a longitudinal corpus-based study, during which a comparable corpus is analysed diachronically. For the purposes of this study, the non-translated articles published in 1990-1991 will be compared to the non-translated articles published in 2009-2010.

The second aim of this corpus-based methodology is to examine the role of translation in this language change phenomenon. This involves a comparable corpus-based analysis where translated and non-translated Modern Greek popular science articles are analysed synchronically. First, the non-translated articles published in 2003-2004 will be compared to the translated articles published during the same years. Then, the same type of analysis will be conducted for articles published in 2009-2010. Two separate phases of analysis are included in order to investigate the extent to which the linguistic features in the translated texts differ from those of the non-translated ones at different time periods. More particularly, the first phase of analysis focuses on a period of time when the influence from English translations of popular science articles was at its initial stage. The second phase of analysis focuses on a later stage of the contact between English and Modern Greek through translation, as far as the particular genre or popular science is concerned.

Finally, this corpus-based methodology aims to investigate the role of the source texts in this language contact situation. This involves the synchronic analysis of a parallel corpus of translated articles and their

originals, which consists of two phases of analysis, i.e. the translated popular science articles that were published in 2003-2004 will be compared to their source texts and the same analysis will be conducted for the articles published in 2009-2010.

The analyses will be conducted with the help of the Concordance tool of WordSmith Tools 5.0 and will be based on semi-automatic methods, since at points where a closer examination of the texts is required, they will be analysed manually. The verb form is considered to be the unit of analysis and auxiliary verbs are excluded from the counts, since they do not provide any lexical information. For the sample studies discussed below, a part-of-speech (POS) tagger is not being used due to the fact that available Modern Greek POS taggers score relatively low on accuracy and Modern Greek verbs can be quite accurately identified from their suffixes with the use of wildcards.

## 4. Preliminary Results

Although this is still an ongoing project, a number of sample studies indicate that a corpus-based methodology that combines synchronic and diachronic corpus-based approaches, as well as parallel and comparable corpora can considerably assist in the analysis of the linguistic features of translated texts and their impact on non-translated ones. Articles for the sample studies are taken from the newspaper *Βήμα* (The Tribune), which includes a section dedicated to scientific issues.

### 4.1. Language Change in Modern Greek

In terms of the first aim of this corpus-based methodology, that is, the examination of language change in Modern Greek, a sample study of popular science articles published in 1991 and 2010 involving 4,000 words was conducted in order to examine changes in the frequency of the passive voice. Although this is a very small sample study, it was found that the passive voice has become more frequent in Modern Greek in the last 20 years, at least in terms of the specific genre of popular science articles. In particular, in the articles published in 1991, 273 verb forms were found, 42 of which involved passive verb forms. In the articles published in 2010, 217 instances of verb forms were identified, 42 of which were passive. This means that there is an approximately 5%

increase in the frequency of the distribution of passive voice constructions in Modern Greek. However, this 5% increase may be attributed to a number of factors that are irrespective of contact-induced language change, i.e. it may be a result of internal language changes. An analysis of translated texts is necessary in order to establish the extent to which contact through translation has encouraged a frequential copying of passive voice structures from English.

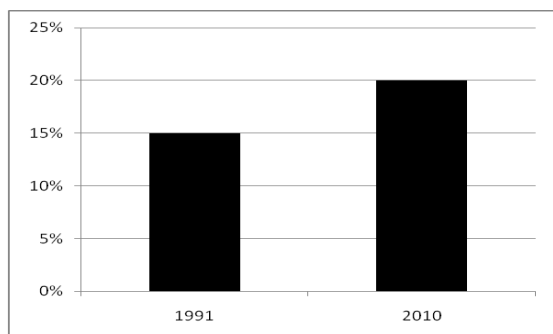


Figure 2: Change in the frequency of the passive voice in Modern Greek (1991-2010)

#### 4.2. The Role of the Translations

A second sample study was conducted in order to examine the role of the translation in this language change situation. In particular, a small corpus of 20,000 words taken from translated and non-translated Modern Greek popular science articles published in 2010 was analysed. The analysis revealed that the frequency of the passive voice in the translated and non-translated articles is very similar, i.e. approximately 20%. In the non-translated articles, 1,081 verb instances were identified, 215 of which were passive, whereas the translated articles included 1,234 verb forms, out of which 243 involved passive voice occurrences.

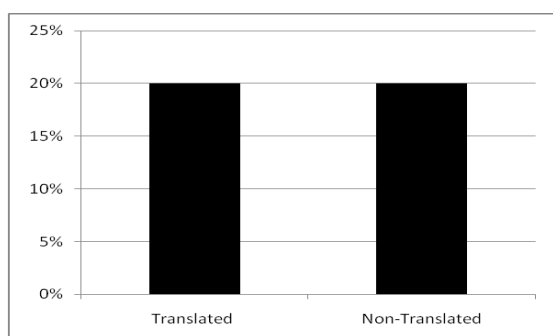


Figure 3: Frequency of the passive voice in translated and non-translated articles published in 2010

This similarity in terms of the proportions of the passive voice suggests that the translated texts at least mirror the changes in the frequency of the passive voice that is attested in Modern Greek. This sample study focuses on a later stage of contact between English and Modern Greek in terms of popular science publications and it is assumed that this later stage indicates more established instances of copying, if we accept that some kind of copying has taken place. Although a comparable analysis of articles published in 2003-2004, when the influence from Anglophone source texts was at its initial stage, has not yet been attempted, such an analysis is likely to reveal a different patterning than the one discussed above, i.e. that the frequency of the passive voice is higher in translated texts than in non-translated ones. This will indicate that the frequential copying of passive voice gradually habitualised in the context of translation.

#### 4.3. The Role of the Source Texts

Finally, in terms of the last aim of this corpus-based methodology, namely the investigation of the role of the English source texts in this language change phenomenon, it should be mentioned that although a sample study is not available at the moment for this type of analysis, it can be predicted based on the previous sample study that translated texts are likely to follow the patterns of the source texts. Corpus studies (Biber et al. 1999:476) suggest that the English passives account for approximately 25% of all finite verbs in academic prose and for 15% in news. Popular science articles are considered to be somewhere in between these two genres, since they present scientific issues using a journalistic language. Thus, the frequency of the passive voice in English popular science articles can be expected to be somewhere between these two percentages, i.e. 20%. The distribution of the frequency of the passive voice in the previous sample study represents exactly this proportion. If this prediction is confirmed, it will suggest that the translation of popular science articles from Anglophone sources tends to encourage the frequential copying of the passive voice in Modern Greek. In that case, Modern Greek being the Basic Code copied the frequency of the passive voice patterns from the Model Code, which is English. The copies first habitualised in the discourse of the translation and then spread into the general linguistic community and became monolingual copies.

## 5. Conclusion

Although the results are only preliminary, the importance of this corpus-based study lies in a number of factors. Firstly, it is one of the first diachronic corpus-based studies ever to be attempted within the field of translation studies and it raises collective awareness of how translation can encourage the dissemination of particular source language linguistic features. If this scholarly strand is to be consolidated, more research across a wider range of language pairs and linguistic features has to be conducted. Secondly, it is one of the first sustained corpus-based studies ever to be conducted in the Modern Greek context within the field of translation studies, which aims at analysing systematically and in depth the Modern Greek linguistic features of translated texts. Finally, this study combines all corpus-based methodologies, i.e. diachronic, synchronic, comparable and parallel, under one research aim: the investigation of translation as a language contact phenomenon. This is probably the most important aspect of this study since it stresses the numerous advantages of collaborative techniques and engages them in a mutually profitable dialogue.

## 6. References

- Amouzadeh, M., House, J. (2010): Translation and Language Contact: The case of English and Persian. *Languages in Contrast*, 10(1), pp. 54-75.
- Apostolou-Panara, A. (1991): English Loanwords in Modern Greek: An overview. *Terminologie et Traduction*, 1(1), pp. 45-60.
- Baker, M. (1995): Corpora in Translation Studies: An overview and some suggestions for future research. *Target*, 7(2), pp. 223-243.
- Baumgarten, N., House, J., Probst, J. (2004): English as a Lingua Franca in Covert Translation Processes. *The Translator*, 10(1), pp. 83-108.
- Baumgarten, N., Özçetin, D. (2008): Linguistic Variation through Language Contact in Translation. In E. Siemund & N. Kintana (Eds.), *Language Contact and Contact Languages*. Amsterdam: John Benjamins, pp. 293-316.
- Becher, V., House, J., Kranich, S. (2001): Convergence and Divergence of Communicative Norms through Language Contact in Translation. In K. Braunmüller & J. House (Eds.), *Convergence and Divergence in Language Contact Situations*. Amsterdam: John Benjamins, pp. 125-152.
- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E. (1999): *Longman Grammar of Spoken and Written English*. Harlow: Longman.
- Csató, É.Á. (2002): Karaim: A high-copying language. In M.C. Jones & E. Esch (Eds.), *Language Change: The interplay of internal, external and extra-linguistic factors*. Berlin: Mouton de Gruyter, pp. 315-327.
- Johanson, L. (1993): Code-Copying in Immigrant Turkish. In G. Extra & L. Verhoeven (Eds.), *Immigrant Languages in Europe*. Clevedon, Philadelphia and Adelaide: Multilingual Matters, pp. 197-221.
- Johanson, L. (1999): The Dynamics of Code-Copying in Language Encounters. In B. Brendemoen, E. Lanza & E. Ryen (Eds.), *Language Encounters across Time and Space*. Oslo: Novus Press, pp. 37-62.
- Johanson, L. (2002): Structural Factors in Turkic language Contacts. London: Curzon.
- Johanson, L. (2008): Remodelling Grammar: Copying, conventionalisation, grammaticalisation. In E. Siemund & N. Kintana (Eds.), *Language Contact and Contact Languages*. Amsterdam: John Benjamins, pp. 61-79.
- House, J. (2004): English as Lingua Franca and its Influence on Other European Languages. In J.M. Bravo (Ed.), *A New Spectrum of Translation Studies*. Valladolid: Universidad de Valladolid, pp. 49-62.
- House, J. (2008): Global English and the Destruction of Identity?. In P. Nikolaou & M.V. Kyritsi (Eds.), *Translating Selves: Experience and identity between languages and literatures*. London and New York: Continuum, pp. 87-107.
- Steiner, E. (2008): Empirical Studies of Translations as a Mode of Language Contact: 'Explicitness' of lexicogrammatical encoding as a relevant dimension. In E. Siemund & N. Kintana (Eds.), *Language Contact and Contact Languages*. Amsterdam: John Benjamins, pp. 317-346.
- Verschik, A. (2008): *Emerging Bilingual Speech: From Monolingualism to Code-Copying*. London: Continuum.
- Warburton, I. (1975): The Passive in English and Greek. *Foundations of Language*, 13(4), pp. 563-578.