

Μηχανική μετάφραση δομών με υποστηρικτικό ρήμα από τα ιταλικά στα ελληνικά: Η περίπτωση του ιταλικού *fare*

Κωνσταντίνος Χατζηθεοδώρου

Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

Abstract

One of the most challenging tasks for a translator (whether human or automatic) is the proper treatment of constructions such as *fare una presentazione* or *fare una telefonata* which contain verbs ‘void’ of meaning, here called ‘support verbs’. In the last few years, many different approaches of recognition and paraphrases have been proposed in order to improve the machine translation output. The aim of this paper is to present the way a support verb construction is automatically translated by widely-used free web translation software (such as Google Translate or Bing Translator) and if pre-editing helps. In addition, a discussion is attempted about whether the generation of the equivalent paraphrases in the target language and especially the nominalizations ensure better output. The evaluation results show that currently the outcome is far from perfect, as the provided translation is very often literal and unnatural. However, when the support verb construction is recognized by the machine, the results are significantly better.

Keywords: nominalization, support verb construction, paraphrase, fare, machine translation, pre-editing

1. Εισαγωγή

Η Αυτόματη ή Μηχανική Μετάφραση (στο εξής MM) (Machine Translation - MT) είναι ένα πεδίο της Υπολογιστικής Γλωσσολογίας (Computational Linguistics - CL) που ερευνά τη μεταφορά κειμένου από μία φυσική γλώσσα σε μία άλλη με τη χρήση προηγμένης τεχνολογίας, χωρίς την παρέμβαση ανθρώπινου παράγοντα (Hutchins 1992). Το εμπορικό αλλά και ακαδημαϊκό ενδιαφέρον για ανάπτυξη συστημάτων MM είναι έντονο και απασχολεί ερευνητές από διάφορους τομείς (στατιστική, ορολογία, σημασιολογία, συντακτική - μορφολογική ανάλυση κ.ά.). Επιπροσθέτως, η έρευνα στη MM δημιουργεί σχέσεις μεταξύ γλωσσών και επιλύει καίρια μεταφραστικά προβλήματα.

Από το 1954, χρονολογία που παρουσιάστηκε ένα από τα πρώτα συστήματα MM μέχρι και σήμερα, πολλές είναι οι προσεγγίσεις που έχουν προταθεί από τους ερευνητές (Bennett και Slocum 1985, Brown κ.ά.1993, Nagao 1984, Mitamura κ.ά.1991). Από τη χρονοβόρα και δαπανηρή προσέγγιση μετάφρασης βάσει κανόνων (Bennett και Slocum 1985) έως την γρήγορη αλλά όχι πάντα ακριβή (accurate) Στατιστική MM (ΣΜΜ) (Koehn 2010) παρουσιάζονται πολλές διακυμάνσεις. Για το λόγο αυτό, τα τελευταία χρόνια οι έρευνες στρέφονται σε υβριδικές λύσεις, δηλαδή σε συστήματα που λειτουργούν συνδυάζοντας τεχνικές από τις προτεινόμενες προσεγγίσεις. Η επικρατέστερη λύση είναι, ένα σύστημα να εκπαιδεύεται με ένα σώμα κειμένων και στη συνέχεια χρησιμοποιούν έναν πεπερασμένο αριθμό κανόνων ώστε να βελτιώσουν τις παραγόμενες μεταφράσεις. Από την άλλη μεριά, δεν είναι πάντα εφικτή αυτή η λύση, καθώς για τις «μικρές» γλώσσες, αφενός δεν υπάρχουν διαθέσιμα σώματα κειμένων αλλά και αφετέρου γλωσσολόγοι με σχετικές γνώσεις.

Οι δομές με υποστηρικτικό ρήμα δεν αναγνωρίζονται σε όλες τις περιπτώσεις από τα συστήματα MM, καθώς η δομή τους δεν είναι πάντα σταθερή. Επιπλέον, πολλές φορές το ρήμα κενής σημασίας που εμπεριέχουν αγνοείται από το σύστημα MM. Τέλος, υπάρχουν δομές που αν μεταφραστούν κατά λέξη στη γλώσσα-στόχο, οδηγούν σε μη ισοδύναμη δομή. Για παράδειγμα, η ιταλική δομή *fare una scommessa* δεν θα πρέπει να μεταφραστεί ως *κάνω ένα στοίχημα* αλλά ως *βάζω ένα στοίχημα*. Ως εκ τούτου, τα υποστηρικτικά ρήματα δεν αποδίδονται ορθά και πολλές είναι οι περιπτώσεις που απουσιάζουν από το τελικό μεταφραστικό προϊόν.

Στόχος της παρούσας εργασίας είναι να εξετάσει αφενός αν τα δημοφιλέστερα συστήματα αυτόματης μετάφρασης αναγνωρίζουν δομές με το υποστηρικτικό ρήμα *fare* και αφετέρου αν στο στάδιο της προ-επεξεργασίας της γλώσσας-πηγής συντελεί σημαντικά στη βελτίωση της ποιότητας των μεταφράσεων στην ελληνική γλώσσα. Επιπλέον, εστιάζουμε την προσοχή μας στις παραγόμενες μεταφράσεις και μελετάται αν αυτές αποτελούν κατά λέξη μετάφραση ή παραφράσεις, δίνοντας έμφαση στις ρηματοποιήσεις κατηγορηματικών ουσιαστικών.

Το υπόλοιπο του παρόντος άρθρου είναι δομημένο ως εξής: Στην Ενότητα 2 αναλύονται οι δομές με υποστηρικτικό ρήμα. Στην Ενότητα 3 περιγράφεται η μεθοδολογία της έρευνάς μας, ενώ τα ποιοτικά αποτελέσματα παρουσιάζονται στην Ενότητα 4. Τέλος, στην Ενότητα 5 συνοψίζονται τα κύρια συμπεράσματα και η μελλοντική έρευνα.

2. Δομές με υποστηρικτικό ρήμα

Η βασική ιδέα της χαρισιανής θεωρίας σχετικά με τα κατηγορηματικά ουσιαστικά (Harris 1976) είναι πως ο συνδυασμός ενός μεγάλου αριθμού αφηρημένων ουσιαστικών και ενός απολεξικοποιημένου ρήματος (π.χ. *fare*) μπορεί να ταξινομηθεί και να αναλυθεί σαν κατηγορήμα με υποκείμενο και ενδεχομένως σε ένα ή περισσότερα αντικείμενα (Σφέτσιου 2007: 19). Έτσι, η πρόταση (1) είναι δυνατόν να αναλυθεί σε μια άλλη πρόταση απλούστερης μορφής, διατηρώντας παράλληλα το νόημά της.

1) La telefonata di Ugo a Floria.

2) Ugo fa una telefonata a Floria.

Δεδομένου αυτού του μετασχηματισμού, παρατηρούμε πως η (2), παρ' όλη την απλοποίηση που υπέστη, διατήρησε τα ίδια ορίσματα με την αρχική πρόταση (υποκείμενο και αντικείμενο). Στο παράδειγμα μας, το κατηγορηματικό ουσιαστικό που αναπαρίσταται με το Npred, είναι το *telefonata* και το υποστηρικτικό ρήμα που αναπαρίσταται με το Vsup είναι το *fare*. Τα Vsup εμπεριέχονται σε προτάσεις της μορφής: N₀ [Vsup Npred] (1)

Η έρευνα του Barrier (2006: 5) καταλήγει στο συμπέρασμα πως αυτός ο πολυλεκτικός συνδυασμός εκλαμβάνεται ως παγιωμένη έκφραση, δεδομένου ότι αποτυπώνει μια μοναδική έννοια. Μάλιστα αναφέρει πως αυτοί οι συνδυασμοί κατανοούνται ευκολότερα σε αντίθεση με τις παγιωμένες εκφράσεις (*κάνω μπαμ, θα σε κάνω μαύρο στο ζύλο, κάνω θέατρο, fare quattro chiacchiere, fare del bene, fare orecchie da mercante, fare l'avvocato del diavolo* κ.λπ.).

Εντούτοις, ο συνδυασμός «Vsup Npred» δεν μπορεί να θεωρηθεί παγιωμένη έκφραση αφού επιδέχεται συντακτική διάσπαση. Για παράδειγμα:

3) Giorgio fa una (E + bella + breve + rapida) telefonata a Maria.

4) * fare molta fretta.

η πρόταση (3) μπορεί να διασπαστεί από τα επίθετα *bella, breve* ή *rapida*. Ωστόσο, δεν συμβαίνει το ίδιο με την πρόταση (4). Αυτή η διαπίστωση είναι εφαρμόσιμη και στην ελληνική γλώσσα. Ως παραδείγματα δανειζόμαστε προτάσεις από την έρευνα των Kyriacopoulou και Sfetsiou (2002: 165):

- 5) Ο Πέτρος κάνει έναν (E + ωραίο) περίπατο κάθε μέρα.
- 6) *Ο ουρανός βρέχει πολύ καρεκλοπόδαρα.

Στη βιβλιογραφία προκύπτει πως, οι φράσεις τύπου «Vsup Npred» αναφέρονται ως *support verb constructions*, όρος που στην ελληνική γλώσσα αποδίδεται ως *δομές με υποστηρικτικό ρήμα*, ενώ στην ιταλική γλώσσα ως *costrutti con verbo supporto*¹ (Cantarini 1999). Αξίζει να σημειωθεί πως η ορολογία που χρησιμοποιείται προέρχεται από μετάφραση του γαλλικού όρου *construction à verbe support*.²

3. Μεθοδολογία

Είναι γνωστό ότι η μεθοδολογία που πρέπει να ακολουθείται είναι στενά προσδιοριζόμενη από τις ιδιαιτερότητες και τους επιμέρους στόχους της εκάστοτε έρευνας (Parlett και Hamilton 1977: 13, Burgess 1994:10). Στο πλαίσιο της παρούσας έρευνας προσανατολιστήκαμε στην ποιοτική ανάλυση των μεταφραστικών προϊόντων που παρήχθησαν από υβριδικά συστήματα, τα οποία οι χρήστες του διαδικτύου μπορούν να χρησιμοποιήσουν ελεύθερα. Ειδικότερα, χρησιμοποιήθηκαν τα συστήματα Google Translate³ (GT) και Bing Translator⁴ (BT) διότι είναι υβριδικά και ευρέως διαδεδομένα.

Το μεθοδολογικό μας πλαίσιο ολοκληρώνεται σε δυο στάδια. Το πρώτο στάδιο αφορά τη μετάφραση προτάσεων από τα ιταλικά στα ελληνικά δίχως καμία παρέμβαση στο πρωτογενές κείμενο, ενώ το δεύτερο στάδιο περιλαμβάνει την προ-επεξεργασία του πρωτογενούς κειμένου με την τεχνική της απαλοιφής των δομών με υποστηρικτικό ρήμα και την αυτόματη αντικατάστασή τους από ένα ρήμα το οποίο είναι ομόρριζο με το κατηγορηματικό ουσιαστικό. Στόχος του συγκεκριμένου σταδίου είναι να μελετηθεί αν το στάδιο της προ-επεξεργασίας επηρεάζει την ποιότητα των παραγόμενων μεταφράσεων. Αναλυτικότερα, η δημιουργία παραφράσεων πραγματοποιείται χρησιμοποιώντας το λογισμικό NooJ και ένα ηλεκτρονικό λεξικό κατηγορηματικών ουσιαστικών και των ομόρριζων ρημάτων τους, που δημιουργήθηκε για το σκοπό αυτό. Η προ-επεξεργασία αφορά την ιταλική γλώσσα και ολοκληρώνεται με βάση τον γράφο⁵ της Εικόνας 1. Η Εικόνα 2 παρουσιάζει τις δομές που εξήχθησαν και τις παραφράσεις τους.

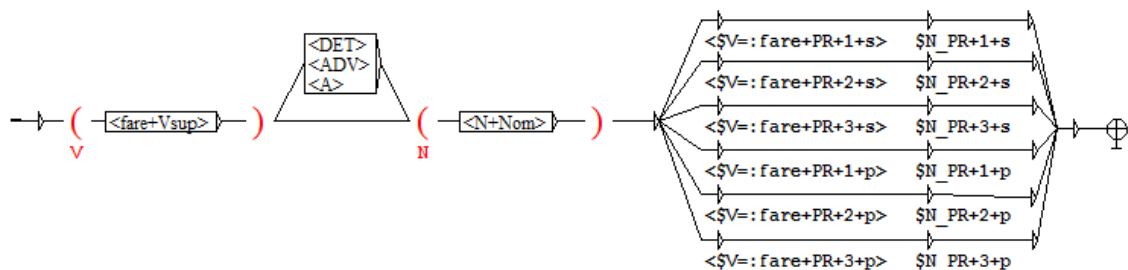
¹ Από άλλους ερευνητές αναφέρονται ως *costrutti a vebo supporto* ή *frasi a vebo supporto, costruzioni a verbo supporto, costruzioni con verbo operatore* ή *costruzioni con verbo supporto* (La Fauci 2003).

² Στην αγγλική ορολογία αναφέρονται ως *support verb constructions* (Danlos 1992, Krenn 2000) ή *stretched verb constructions* (Allerton 2002αβ).

³ <http://translate.google.com/>

⁴ <http://www.bing.com/translator/>

⁵ Ο συγκεκριμένος γράφος αναγνωρίζει και παραφράζει μόνο δομές της οριστικής ενεστώτα. Λόγω οικονομίας χώρου δεν παρουσιάζονται οι υπόλοιποι γράφοι που αφορούν τις υπόλοιπες εγκλίσεις και χρόνους.



Εικόνα 1: Γράφος για τη δημιουργία παραφράσεων δομών με υποστηρικτικό ρήμα

lista e cominciamo ad escludere qualcosa. 2122399	Facciamo una visita/visitiamo	alla sorella? 2126016	Che ne dici
alla sorella? 2126016	facciamo una fermata/fermiamo	sul Monte Fato? 2129074	- Allora faremo
citta' misteriosa e non ti	fai nessuna domanda/domandi	? 2130357	Ma una volta che ti
per me. 2159218	fai una chiamata/chiami	anonima alla polizia. 2159261	Se ti
sono dei lavori nella strada,	facciamo una deviazione/deviamo	. 2165350	Se potessi avere quelle lettere
Ce ne faremmo uno doppio. 2327766	Facciamo una festa/festeggiamo	! 2327773	Beh, me ne fai una
E' l'ultima volta che	faccio una favore/favorisco	a te e ai tuoi	
tantoin tanto faccio una scommessa. 2395233	Faccio una corsa/corro	verso l'aeroporto di San	
una multa a Batman? 2411766	fate una controllo/controllate	approfondito sul fascicolo personale di	

Εικόνα 2: Δομές με υποστηρικτικό ρήμα *fare* σε περιβάλλον NooJ

Τα παραγόμενα μεταφραστικά προϊόντα αξιολογήθηκαν, αναλύθηκαν και κατηγοριοποιήθηκαν με κριτήριο τη ορθότητα της μετάφρασης της δομής με υποστηρικτικό ρήμα που εμπειρεύει η κάθε μια από αυτές. Πρέπει να τονισθεί ότι στην παρούσα έρευνα μας απασχολεί μόνο το τμήμα της πρότασης που περιέχει τη δομή με υποστηρικτικό ρήμα και όχι το υπόλοιπο τμήμα της πρότασης. Επίσης, δεν γίνεται σύγκριση των συστημάτων MM που χρησιμοποιήθηκαν, καθώς δεν είναι αυτός ο κύριος σκοπός της παρούσας μελέτης.

Τέλος, πρέπει να αναφερθεί πως το σώμα αξιολόγησης αποτελείται από 100 προτάσεις που εμπειρεύσαν δομές με υποστηρικτικό ρήμα *fare*. Το δείγμα συστάθηκε τυχαία και αποτελείται από προτάσεις που εξήχθησαν από τα (παράλληλα) σώματα κειμένων Europarl (Koehn 2005), το Corpus e Lessico di Frequenza dell'Italiano Scritto Contemporaneo (Laudanna κ.ά. 1995) καθώς και από κείμενα εφημερίδων (Il Corriere Della Sera,⁶ Repubblica,⁷ La Stampa).⁸

4. Αποτελέσματα

Στα μεταφραστικά προϊόντα των συστημάτων GT και BT έγινε ποιοτική αξιολόγηση και στη συνέχεια αυτά κατηγοριοποιήθηκαν μη αυτόματα από δίγλωσσο γλωσσολόγο. Παρακάτω γίνεται σχολιασμός των μεταφράσεων ανάλογα με την κατηγορία στην οποία ταξινομήθηκαν.

4.1. Αναχώρηση δομών με υποστηρικτικό ρήμα

Στις περισσότερες των περιπτώσεων οι δομές με υποστηρικτικό ρήμα *fare* αναγνωρίζονται από τα συστήματα MM. Σε ότι αφορά τη μετάφραση, άλλοτε αυτές μεταφράζονται κατά λέξη (βλ. παράδειγμα (7)) και άλλοτε με ισοδύναμα (βλ. παράδειγμα (8)). Ωστόσο, υπάρχουν και περιπτώσεις όπου το μετάφρασμα απέχει πολύ από αυτό της γλώσσας-στόχου, παρόλο που αναγνωρίστηκε με ακρίβεια από το σύστημα. Στην περίπτωση αυτή, η προβληματική έγκειται προφανώς στην ενδιάμεση γλώσσα που χρησιμοποιείται, καθώς το ζεύγος ιταλικά-ελληνικά δεν είναι πολύ ισχυρό,

⁶ <http://www.corriere.it/>

⁷ <http://www.repubblica.it/>

⁸ <http://www.lastampa.it/>

κυρίως λόγω των λιγοστών διαθέσιμων πόρων για εκπαίδευση στατιστικών μεταφραστικών μοντέλων. Επίσης, η έλλειψη δίγλωσσων γλωσσολόγων που να μελετούν το συγκεκριμένο γλωσσικό ζεύγος δεν επιτρέπει τη δημιουργία κανόνων με στόχο την επίτευξη ενός καλύτερου μεταφράσματος. Χαρακτηριστικό είναι το παράδειγμα (7) (βλ. Εικόνα 3), όπου η μονάδα *fare una visita* μεταφράζεται στα Ελληνικά ως *πληρώνω μια επίσκεψη*. Μια πιθανή εξήγηση είναι πως προφανώς αποτελεί μετάφραση της αγγλικής δομής *to pay a visit*, είτε κάποιας άλλης γλώσσας. Επιπροσθέτως, παρατηρείται ότι οι περισσότερες από τις προτεινόμενες μεταφράσεις (4/5) δεν είναι αποδεκτές, ενώ μόνο η τελευταία κρίνεται αποδεκτή μετάφραση και μάλιστα αποτελεί προϊόν παράφρασης (βλ. Εικόνα 3) και συγκεκριμένα με την τεχνική της ρηματοποίησης του κατηγορηματικού ουσιαστικού.

7) Era andato a **fare una visita** alla figlia.

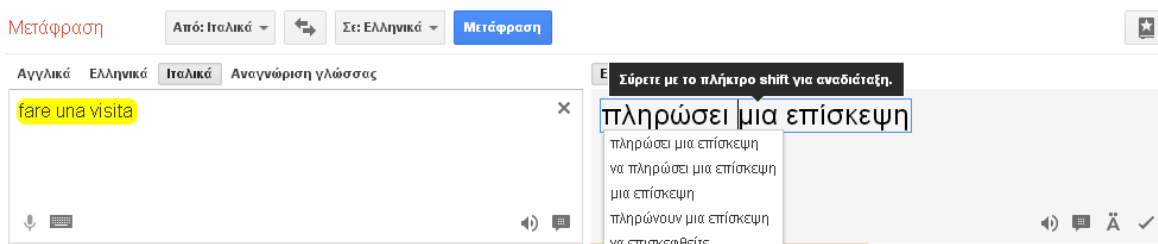
GT: Είχε πάει να **πληρώσει μια επίσκεψη** στην κόρη της.

BT: Πήγε να **πληρώσει μια επίσκεψη** στην κόρη του.

8) ... gli consiglierai, quando parla di buon governo e di mal governo, di **fare una scappata** a Siena, al Palazzo comunale

GT: ... εγώ θα τον συμβούλευα, όταν μιλάει για μια σωστή διακυβέρνηση και την κακή κυβέρνηση, **κάνει ένα ταξίδι** στην Σιένα, στο Δημοτικό Μέγαρο ...

BT: ... συνιστούμε, όταν μιλάμε για καλό και κακό κυβέρνησης, **να κάνει ένα ταξίδι** στην Σιένα, Palazzo comunale ...



Εικόνα 3: Στιγμιότυπο του συστήματος Google Translate (Μάιος 2013)

4.2. Παράφραση (ουσιαστικοποίηση ή ρηματοποίηση)

Όπως αναφέρεται παραπάνω, η ονοματοποίηση και η ρηματοποίηση είναι μετασχηματισμοί στενά συνδεδεμένοι με τις δομές με υποστηρικτικό ρήμα. Δεδομένης αυτής της παρατήρησης, τα συστήματα MM προβαίνουν στην παραγωγή τέτοιου είδους μετασχηματισμών. Πιο συγκεκριμένα, στο παράδειγμα (9) παρατηρούμε πως η ρηματική φράση έχει μετατραπεί σε ονοματική φράση. Αντιστοίχως, στο παράδειγμα (10) το 'κενό' ρήμα έχει αντικατασταθεί από ένα άλλο ρήμα το οποίο είναι ομόριζο του κατηγορηματικού ουσιαστικού. Στις περιπτώσεις αυτές, συμπεραίνουμε/ υποθέτουμε πως στην όλη διαδικασία εμπλέκονται κανόνες.

9) Dovremmo però evitare di **fare uno sbaglio**.

GT: Αλλά εμείς πρέπει να αποφύγουμε **ένα λάθος**.

BT: Ωστόσο, θα πρέπει να αποφύγουμε να **κάνουν λάθος**.

10) ... nel momento in cui l'Unione europea non ha ancora deciso di **fare pulizia** nelle banche europee e di dire ai risparmiatori, ai produttori, all'economia reale, ...

GT: ... σε μια στιγμή που η Ευρωπαϊκή Ένωση δεν έχει ακόμη αποφασίσει να **καθαρίσει** τις τράπεζες και να πω για τους επενδυτές, οι παραγωγοί, η πραγματική οικονομία ...

BT: ...σε μια στιγμή, όταν η Ευρωπαϊκή Ένωση δεν έχει ακόμη αποφασίσει να **κάνει τον καθαρισμό** στις ευρωπαϊκές τράπεζες και να πει επενδυτές, παραγωγών, την πραγματική οικονομία...

4.3. Αναγνώριση άρθρου

Μελετώντας την αναγνώριση και κατά συνέπεια η άρτια απόδοση του άρθρου μιας δομής με υποστηρικτικό ρήμα, μπορούμε εύκολα να κατανοήσουμε πως ακριβώς παράγονται οι μεταφράσεις από υβριδικά συστήματα MM. Όπως παρατηρούμε στο παράδειγμα (11), το άρθρο αρσενικού γένους του πρωτογενούς κειμένου αποδόθηκε άρτια στη γλώσσα-πηγή με άρθρο θηλυκού γένους. Από αυτή την παρατήρηση συμπεραίνουμε πως προφανώς το σύστημα αναγνώρισε τη δομή με υποστηρικτικό ρήμα και απέδωσε την αντίστοιχή της στην ελληνική γλώσσα.

11) Spero che i suoi collaboratori le possano **fare un riassunto** di ciò che di più interessante è stato detto in sua assenza.

GT: Ελπίζω ότι η ομάδα του μπορεί να **κάνει την μια περίληψη** του τι έχει ειπωθεί πιο ενδιαφέρουσα κατά την απουσία του.

BT: Ελπίζω ότι οι συνεργάτες του μπορεί να **κάνει μια περίληψη** των πιο ενδιαφέρουσα τι είχε πει σε περίπτωση απουσίας του.

4.4. Μετάφραση ανάλογα με το συγκείμενο

Ιδιαίτερης σημασίας είναι η διερεύνηση της πιστότητας του συστήματος σε σχέση με το συγκείμενο. Παρατηρούμε λοιπόν στα παραδείγματα των φράσεων (12) και (13), πως αλλάζοντας το συγκείμενο, αλλάζει η απόδοση της δομής με υποστηρικτικό ρήμα. Πιο συγκεκριμένα, στις παρακάτω προτάσεις η μονάδα *fare una domanda* μεταφράζεται ως *έχω μια ερώτηση*, *θέτω μια ερώτηση* ή *κάνω μια ερώτηση*.

12) Vorrei **fare una domanda** complementare sul contesto globale.

GT: Έχω μια ερώτηση σχετικά με το παγκόσμιο πλαίσιο.

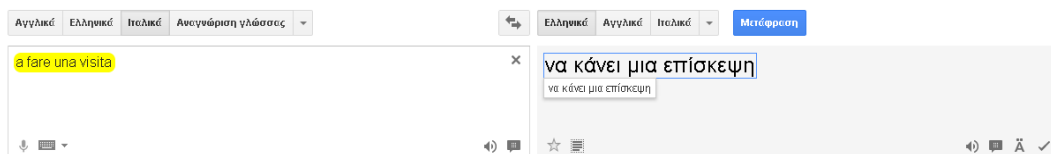
BT: Έχω μια συμπληρωματική ερώτηση σχετικά με το παγκόσμιο πλαίσιο.

13) salve sono una ragazza di 25 anni e vorrei **fare una domanda** io a lavoro condivido il bagno ...

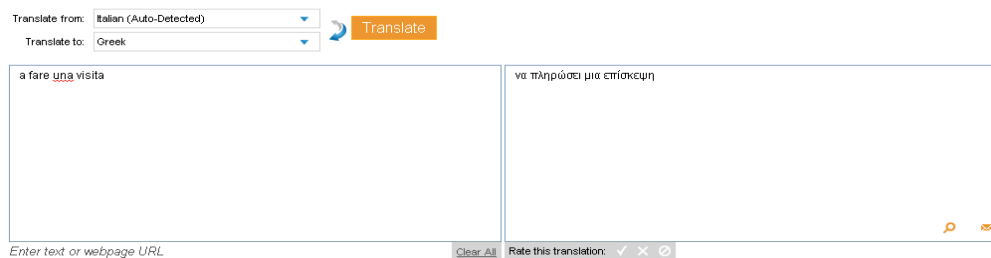
GT: Για σου είμαι ένα κορίτσι 25 ετών και θα ήθελα να **κάνω μια ερώτηση** για τη δουλειά, μοιράζονται το μπάνιο ...

BT: Για σας είμαι 25 χρόνια παλιά κορίτσι, και θα ήθελα να **θέσω ένα ερώτημα**, εγώ δουλεύω με τους άνδρες ...

Τέλος, πρέπει να αναφερθεί πως παρόλο που η μονάδα *fare una visita* μεταφράζεται από το GT ως *πληρώνω μια επίσκεψη*, όταν προηγείται η πρόθεση *a* πριν από αυτή, μεταφράζεται άρτια ως *κάνω μια επίσκεψη*. Το ακριβώς αντίθετο όμως συμβαίνει με το BT (βλ. Εικόνα 4 και Εικόνα 5).



Εικόνα 4: Στιγμιότυπο του συστήματος Google Translate (Μάιος 2013)



Εικόνα 5: Στιγμιότυπο του συστήματος Bing Translator (Μάιος 2013)

4.5. Προ-επεξεργασία δομών με υποστηρικτικό ρήμα

Όπως αναφέρεται παραπάνω, το σώμα αξιολόγησης προ-επεξεργάστηκε έτσι ώστε οι δομές με υποστηρικτικό ρήμα να αντικατασταθούν με ένα ρήμα το οποίο είναι ομόρριζο με το κατηγορηματικό ουσιαστικό και στη συνέχεια υποβλήθηκε για μετάφραση. Στόχος μας ήταν να μελετήσουμε αν η επανεγγραφή της ίδιας πρότασης και η απλοποίηση του πρωτογενούς κειμένου βοηθάει στην βελτίωση της ποιότητας των μεταφράσεων. Τα αποτελέσματα είναι ενθαρρυντικά καθώς πολύ λίγες είναι εκείνες οι προτάσεις που δεν παρουσιάζουν βελτίωση. Ο κύριος λόγος για τη συγκεκριμένη διαπίστωση είναι ότι ενώ η ιταλική πρόταση μπορεί να παραφραστεί με ένα ρήμα, δεν υπάρχει ισοδύναμο ρήμα στα ελληνικά. Για παράδειγμα, *fare uno sconto* > *scontare* > *κάνω/προσφέρω έκπτωση*.

5. Συμπεράσματα

Στην παρούσα έρευνα μελετήσαμε αν τα δημοφιλέστερα συστήματα MM αναγνωρίζουν δομές με το υποστηρικτικό ρήμα και αν στο στάδιο της προ-επεξεργασίας του πρωτογενούς κειμένου συντελεί σημαντικά στη βελτίωση της ποιότητας των μεταφράσεων. Επιπλέον, μελετήσαμε αν τα συστήματα MM μεταφράζουν χρησιμοποιώντας τεχνικές παράφρασης ή κατά λέξη μετάφραση.

Το γενικό συμπέρασμα είναι πως, παρόλο που οι δομές με υποστηρικτικό ρήμα αναγνωρίζονται από τα συστήματα MM, είναι εξαρτώμενες από το συγκεκριμένο και παρουσιάζουν ποικιλομορφία ως προς την απόδοσή τους στην ελληνική γλώσσα. Άλλη μια σημαντική παρατήρηση είναι πως σε πολλές περιπτώσεις το GT και BT αποδίδουν την ίδια λαθεμένη μετάφραση. Αυτό μπορεί να οφείλεται όχι μόνο στη χρήση κοινών παράλληλων σωμάτων εκπαίδευσης, αλλά και «δανεισμό» μεταφραστικών μονάδων από το άλλο σύστημα ή άτυπη συνεργασία των δυο συστημάτων. Εξάλλου, μια από τις πολλές προσεγγίσεις δημιουργίας υβριδικών συστημάτων που έχουν προταθεί ενθαρρύνει την συνεργασία συστημάτων MM με στόχο την παραγωγή μιας βέλτιστης εκδοχής (Heafield και Lavie 2010).

Αν και τα αποτελέσματα της έρευνάς μας είναι ενθαρρυντικά, κρίνεται επιτακτική η ανάγκη δημιουργίας γράφων αναγνώρισης δομών με υποστηρικτικό ρήμα για το γλωσσικό ζεύγος ιταλικά – ελληνικά, καθώς και η δημιουργία νέων κανόνων και ηλεκτρονικών δίγλωσσων λεξικών που θα εμπλουτίσουν τους ήδη υπάρχοντες πόρους (Viteri, 2004). Επίσης, ο σχεδιασμός ενός υβριδικού συστήματος MM που θα χρησιμοποιεί κανόνες για δημιουργία παραφράσεων θα λύσει πολλά μεταφραστικά προβλήματα για το γλωσσικό ζεύγος Ιταλικά-Ελληνικά και θα βελτιώσει την ποιότητα των παραγόμενων μεταφράσεων.

Βιβλιογραφικές αναφορές

- Allerton D.J. (2002a). *Stretched Verb Constructions in English*. London και New York.
 Allerton D.J. (2002β). «Three (or four) levels of word cooccurrences restriction». Στο *Lingua* 63:17-40.
 Barr H. (1994). *Perspectives on Shared Learning*. London: CAIPE.

- Barrier S. και N. Barrier (2003). «Une métagrammaire pour les noms prédicatifs du français». Στο *Actes du colloque TALN 2003*, 303-308.
- Bennett W. και J. Slocum (1985). «The LRC Machine Translation System». Στο *Computational Linguistics* 11(2-3): 111-119.
- Brown P., S.A. Della Pietra, V.J. Della Pietra και R.L. Mercer (1993). «The mathematics of statistical machine translation: Parameter estimation». Στο *Computational Linguistics* 19(2): 263-311.
- Cantarini S. (1999). *Costrutti con verbo supporto: una descrizione contrastiva italiano-tedesco*. Διδακτορική Διατριβή, Università di Pavia.
- Danlos L. (1992). «Support verb constructions: linguistic properties, representation, translation». Στο *Journal of French Language Studies* 2 (1): 1-32.
- Harris Z.S. (1976). *Notes du cours de syntaxe*. Paris: Editions du Seuil.
- Heafield K. και A. Lavie (2010). «Combining Machine Translation Output with Open Source: The Carnegie Mellon Multi-Engine Machine Translation Scheme». Στο *The Prague Bulletin of Mathematical Linguistics* 93: 27–36.
- Hutchins, W.J. και H.L. Somers (1992). *An Introduction to Machine Translation Array*. <http://www.hutchinsweb.me.uk/IntroMT-TOC.htm>.
- Koehn P. (2005). «Europarl: A Parallel Corpus for Statistical Machine Translation». Στο *Proceedings of MT Summit 2005*. Thailand. <http://mt-archive.info/MTS-2005-TOC.htm>
- Koehn P. (2010). *Statistical Machine Translation*, Cambridge University Press.
- Krenn B. (2000). *The Usual Suspects: Data Oriented Models for the Identification and Representation of Lexical Collocations*, Διδακτορική Διατριβή, DFKI & Universität des Saarlandes.
- Kyriakopoulou T. και V. Sfetsiou (2002). «Les constructions nominales à verbe support en grec moderne». Στο S. Nuccorini (επιμ.) *Linguistic Insights, Phrases and Phraseology – Data and Descriptions*. 5: 163-181.
- Laudanna A., A.M. Thornton, G. Brown, C. Burani, και L. Marconi (1995). «Un corpus dell'italiano scritto contemporaneo dalla parte del ricevente». Στο S. Bolasco, L. Lebart και A. Salem (επιμ.), *III Giornate internazionali di Analisi Statistica dei Dati Testuali*. I, Roma: Cisu, 103-109.
- La Fauci N. και I.M. Mirto (2003). *Fare. Elementi di sintassi*. Pisa: ETS
- Mitamura, T., E. Nyberg και J. Carbonell (1991). «An Efficient Interlingua Translation System for Multilingual Document Production». Στο *Proceedings of Machine Translation Summit III*, Washington, DC, 1-8.
- Nagao M. (1984). «A Framework of Mechanical Translation between Japanese and English by Analogy Principle». Στο A. Elithorn (επιμ.) *Artificial and Human Intelligence*. North-Holland, 173-180.
- Parlett M. και D. Hamilton (1977). «Evaluation as Illumination: a new approach to the study of innovatory programmes». Στο *Hamilton D. κ.ά. (επιμ.) Beyond the numbers game*. Exeter: Macmillan Education Ltd.
- Σφέτσιου Β. (2007). *Κατηγορηματικά ονόματα: μέθοδος ανάλυσής τους για ηλεκτρονικές εφαρμογές*. Διδακτορική διατριβή. Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
- Silberstein M. (2005). «NooJ's Dictionaries». Στο *Proceedings of the 2nd Language and Technology Conference*, Poznan University.
- Vietri S. (2004). *Lessico-grammatica dell'italiano: metodi, descrizioni e applicazioni*, UTET Università, Torino.