

"ΛΟΓΟΠΛΟΗΓΗΣΗ"

Δεκέμβριος 2001

Τεύχος 8

Επιστημονικός Υπεύθυνος:

Καθηγητής Γιώργος Καραγιάννης

Υπεύθυνοι Έκδοσης:

Γιάννης Δολόγλου

Βάσω Παναγοπούλου

Συνεργάτες:

Αναστάσιος Πατρικάκος

Γραφίστας:

Άρτεμις Γλάρου

Διεύθυνση:

Ινστιτούτο Επεξεργασίας του Λόγου

Αρτέμιδος 6 & Επιδάουρου

151 25 Παράδεισος Αμαρουσίου

τηλ.: 010 6875300 • fax: 010 6854270

e-mail: www@ilsp.gr<http://www.ilsp.gr>

Την ευθύνη των κειμένων έχουν οι συγγραφείς.

Η χρηματοδότηση της έκδοσης αυτής έγινε από το έργο EUROMAP-Language Technologies της Ευρωπαϊκής Επιτροπής.

Η "Λογοπλοήγηση" διανέμεται δωρεάν.
<http://www.ilsp.gr/logoplo.html>

**"LogoNavigation"**

December 2001

Issue No 8

Scientific Director:

Professor George Carayannis

Edition Executives:

Ioannis Dologlou

Vaso Panagopoulou

Collaborators:

Anastasios Patikakos

Graphics Designer:

Artemis Glarou

Address:Institute for Language and Speech Processing
Artemidos 6 & Epidavrou Str.

151 25 Marousi

Athens, Greece

tel.: +3010-6875300 • fax: +3010-6854270

e-mail: www@ilsp.gr<http://www.ilsp.gr>

The authors are responsible for text content.

Funding for this issue was provided by the EUROMAP-Language Technologies Project of the European Commission.

"LogoNavigation" is distributed free of charge.

<http://www.ilsp.gr/logoplo.html>

Πίνακας Περιεχομένων / Table of Contents

	σελ./pg.
Εισαγωγικό Σημείωμα / Introductory Note	2
I. ΓΕΓΟΝΟΤΑ	3
A. Ευρωπαϊκό Έτος Γλωσσών – Γλωσσική εκπαίδευση στην Κοινωνία της Πληροφορίας»	3
1. Το Ευρωπαϊκό Έτος Γλωσσών	3
2. Γλώσσα και τεχνολογία	3
3. Πρακτικά Ημερίδας	5
B. "European Language Technologies: Actions and Prospects"	6
Programme and speakers	6
Participants	6
Results	6
Third International Conference on Language Resources and Evaluation	7
Conference Aims	7
Conference Topics	8
Conference Programme Committee	10
International Advisory Committee	10
Local Committee	11
Conference Addresses	11
3^ο Συνέδριο για την Ορολογία	12
II. ΣΥΝΕΙΣΦΟΡΕΣ ΕΡΕΥΝΗΤΩΝ	14
Δυνατότητες συστημάτων πολυμέσων για εκπαιδευτικές εφαρμογές – Τεχνολογικές τάσεις και εργαλεία	14
1. Εισαγωγή	14
2. Σύνδεση των Πληροφοριών Διαφορετικών Μέσων	15
3. Ανάκτηση των Πληροφοριών από Συστήματα Πολυμέσων	16
4. Παιδαγωγικές Αρχές	16
5. Αποθήκες Γνώσεων	18
6. Λειτουργικότητα και Αισθητική	18
7. Αξιολόγηση Εκπαιδευτικών Τίτλων	19
Euromap Language Technologies at European Language Technologies: Actions and Prospects	20
Language Technology: Present and Future of Automatic Speech Recognition	23
Προσδιορισμός της σχέσης αριθμού λέξεων που περιέχει ένα κείμενο και του εύρους αντιπροσώπευσης των γλωσσικών μονάδων του	26
1. Εισαγωγή	26
2. Καθορισμός ελάχιστου αριθμού λέξεων ανά κειμενικό δείγμα	27
3. Αποτελέσματα	29
4. Συμπεράσματα	32
Μεθοδολογικές Παρατηρήσεις για την Διδασκαλία της Μετάφρασης του Επιστημονικού και Τεχνικού Λόγου	33
III. ΓΛΩΣΣΑΡΙΟ ΟΡΩΝ ΓΛΩΣΣΙΚΗΣ ΤΕΧΝΟΛΟΓΙΑΣ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ	40
IV. ΤΑ ΝΕΑ ΤΟΥ ΙΕΛ	41
Η Φιλογλωσσία+ στην κορυφαία θέση του εκπαιδευτικού λογισμικού για την Ελληνική Γλώσσα	41
Νέο Βιβλίο	41
Νέο Προϊόν	42
Πρόγραμμα Μεταπτυχιακών Σπουδών στην Γλωσσική Τεχνολογία "Τεχνολογία"	42
EuroMaT	42

Εισαγωγικό Σημείωμα

To 2001 κυκλοφορεί μόνον ένα τεύχος της “Λογοπλοήγησης” αντί δύο τευχών ετησίως που είναι η συνήθης ροή. Λόγω καθυστερήσεων στις χρηματοδοτήσεις δημιουργήθηκε αυτό το εκδοτικό κενό. Ελπίζεται ότι το 2002 θα υπάρξουν 2 τεύχη.

Γίνεται προσπάθεια να διατηρείται η φυσιογνωμία της “Λογοπλοήγησης” που έχει στόχους:

- a) Να αποτυπώνει τα σημαντικά γεγονότα στην χώρα μας, αλλά και διεθνώς, στην θεματική περιοχή της γλωσσικής τεχνολογίας.*
- β) Να εκφράζει τις τάσεις στο αντικείμενο αυτό τόσο από την τεχνολογική πλευρά, όσο και από την πλευρά της ερευνητικής πολιτικής.*
- γ) Να δίνει την δυνατότητα στους ερευνητές του κλάδου να δημοσιεύουν συνεισφορές που θεωρούν είτε καινοτομικές είτε πρωτοποριακές είτε ενδιαφέρουσες για τους άλλους ερευνητές.*
- δ) Να αποτελέσει ένα βήμα συζήτησης αναδυόμενων όρων στον χώρο της γλώσσας και πληροφορικής ώστε να προετοιμαστεί το έδαφος για την αποδοχή των πλέον επιτυχημένων.*

Η γλωσσική τεχνολογία μετά την πρόσκαιρη μετονομασία της σε τεχνολογία του ανθρωπίνου λόγου, στο 5ο πρόγραμμα πλαίσιο, έχει ενταχθεί στο 6ο πρόγραμμα πλαίσιο μαζί με τις τεχνολογίες της γνώσης ως αναπόσπαστο μέρος αυτών των τεχνολογιών για καλύτερευση του περιεχομένου, για αυξημένη διαδραστικότητα και εξασφάλιση πολυγλωσσικής προσπέλασης.

Introductory Note

In 2001, only one issue of ‘Logonavigation’ came into circulation, instead of the annual standard rate of two. Due to funding delays, a publication gap has been created. We hope that in 2002 there will be 2 issues of the Newsletter. We strive to maintain the features of ‘Logonavigation’, which has the following aims:

- a. To keep track of the important events in the field of language technology that take place both in our country as well as internationally.*
- b. To express the trends in the above mentioned field from a technological as well as from a research policy point of view.*
- c. To provide the researchers of the field with the opportunity to publish innovative, pioneering or stimulating contributions.*
- d. To become the forum, where terms emerging in the field of Language and Information Technology are debated, in order to prepare the ground for the standardisation of the most successful ones.*

“Language Technology”, after its temporary renaming into “Human Language Technologies” in the 5th Framework Programme, has been integrated into the 6th Framework Programme along with the other knowledge and interface technologies as an irreplaceable part of the programme because it contributes to the improvement of the content, enhances interactivity and secures multilingual access.

I. ΓΕΓΟΝΟΤΑ

A. Ευρωπαϊκό Έτος Γλωσσών – Γλωσσική εκπαίδευση στην Κοινωνία της Πληροφορίας»

Γιάννα Αντωνίου-Κρητικού

1. Το Ευρωπαϊκό Έτος Γλωσσών

Η Ευρωπαϊκή Ένωση σε συνεργασία με το Συμβούλιο της Ευρώπης ανέλαβαν τη διοργάνωση εκδηλώσεων στο πλαίσιο του Ευρωπαϊκού Έτους Γλωσσών 2001 με στόχο την προώθηση της γλωσσομάθειας και την εδραίωση της γλωσσικής πολυμορφίας της Ευρώπης. Η φιλοσοφία του όλου εγχειρήματος βασίζεται στις ακόλουθες δύο θεμελιώδεις αρχές:

«Πρώτον, η Ευρώπη του μέλλοντος, όπως και αυτή του παρελθόντος και του παρόντος, θα είναι μια Ευρώπη με γλωσσική πολυμορφία. Η πολυμορφία αυτή είναι ένα από τα ισχυρά σημεία της Ευρώπης.

Δεύτερον, καθένας στην Ευρώπη πρέπει να έχει τη δυνατότητα να μαθαίνει γλώσσες, σε οποιαδήποτε στιγμή της ζωής του. Ο κάθε πολίτης αξίζει να έχει την ευκαιρία να ωφεληθεί από τα πολιτιστικά και οικονομικά πλεονεκτήματα που προσφέρουν οι γλωσσικές δεξιότητες. Η εκμάθηση γλωσσών βοηθά επίσης στην ανάπτυξη της ανεκτικότητας και της αμοιβαίας κατανόησης μεταξύ των λαών με διαφορετικό γλωσσικό και πολιτισμικό υπόβαθρο»¹.

Έτσι με γνώμονα τις αρχές αυτές, «το Ευρωπαϊκό Έτος Γλωσσών πλαισιώνεται από ειδικές δράσεις που στοχεύουν στη γλωσσική εκπαίδευση των ευρωπαϊών πολιτών. Η προώθηση της διδασκαλίας όλων ανεξαιρέτως των γλωσσών, η υποστήριξη των λιγότερο ομιλουμένων και η ευαισθητοποίηση στις πολιτισμικές ιδιαιτερότητες συμβάλλουν στη διαφύλαξη του ευρωπαϊκού γλωσσικού και πολιτισμικού πλουραλισμού. Ως εκ τούτου, η ενίσχυση της γλωσσικής/πολιτισμι-

κής εκπαίδευσης αποτελεί το μέσο διαφύλαξης της πολυγλωσσικής και πολυπολιτισμικής Ευρώπης, αφού γλώσσες και πολιτισμοί συνθέτουν το βάθρο πάνω στο οποίο στηρίζεται το οικοδόμημα της Ενωμένης Ευρώπης»².

2. Γλώσσα και Τεχνολογία

2.1 Η ημερίδα του ΙΕΛ

Το Ινστιτούτο Επεξεργασίας του Λόγου (ΙΕΛ) συμμετέχοντας στον εορτασμό του Ευρωπαϊκού Έτους Γλωσσών διοργάνωσε, στις 24 Οκτωβρίου 2001, ημερίδα με θέμα: «Γλωσσική εκπαίδευση στην Κοινωνία της Πληροφορίας»³.

Οι Νέες Τεχνολογίες αποτελούν πεδίο στο οποίο οφείλει να εξοικειωθεί ο ευρωπαίος πολίτης, προκειμένου να αντιμετωπίσει αποτελεσματικά και «ανταγωνιστικά» τις καθημερινές του ανάγκες, είτε αυτές αφορούν τον χώρο της εργασίας, είτε την προσωπική του ζωή. Είναι γεγονός ότι υπάρχει άμεση ανάγκη για αφύπνιση και ευαισθητοποίηση του κοινού σε τομείς εφαρμογών των Νέων Τεχνολογιών.

Στα προγράμματα της Ευρωπαϊκής Ένωσης, η γλωσσική τεχνολογία διεκδικεί πρωταρχική θέση, όχι μόνον για τις πολλά υποσχόμενες προοπτικές της σε επίπεδο ερευνητικών κατακτήσεων, αλλά κυρίως για τις δυνατότητές της σε επίπεδο εφαρμογών, δεδομένου ότι μπορεί να υποστηρίξει ποικιλία επικοινωνιακών πρακτικών στις οποίες εμπλέκεται η φυσική γλώσσα των χρηστών.

Η γλωσσική εκπαίδευση σε συνδυασμό με τις Νέες Τεχνολογίες μπορεί να λειτουργήσει υποστηρικτικά στην γλωσσική κατάρτιση των ευρωπαίων πολιτών και στην δια βίου εκπαίδευση, προσφέροντας ίσες ευκαιρίες στην μάθηση σε παιδιά και ενήλικες, χωρίς χωροχρονικές δεσμεύσεις. Γιατί «η γλωσσική εκπαίδευση με χρήση ψηφιακών και διαδικτυακών τεχνολογιών δεν συμβάλλει απλώς στην επικοινωνία λαών, κοινωνιών και ατόμων, αλλά, συγχρόνως, επιτρέπει τον τε-

¹ Απόσπασμα των V. Reding και W. Schwimmer για το Ευρωπαϊκό Έτος Γλωσσών 2001, ενημερωτικό φυλλάδιο, Γενική Διεύθυνση Εκπαίδευσης και Πολιτιστικών Θεμάτων, Ευρωπαϊκή Επιτροπή.

² Ι. Αντωνίου-Κρητικού, Αξιοποίηση των Νέων Τεχνολογιών στη γλωσσική εκπαίδευση: παραδείγματα εφαρμογών, Γλωσσική εκπαίδευση στην Κοινωνία της Πληροφορίας, Πρακτικά Ημερίδας, Ινστιτούτο Επεξεργασίας του Λόγου, Αθήνα 2001, σ. 65.

³ Τα Πρακτικά της Ημερίδας διατίθενται από το Ινστιτούτο Επεξεργασίας του Λόγου.

χνολογικό “εγγραμματισμό”⁴ και καθιστά ελκυστική και ανταποδοτική την δια βίου εκπαίδευση, αφού μπορεί να προσεγγίσει «όλα τα μέλη της κοινωνίας όσο μακριά και αν βρίσκονται από την εκπαίδευση και την κατάρτιση, επιτρέποντάς τους να αναπτύξουν τις δεξιότητές τους και να τις εκμεταλλευτούν με το βέλτιστο τρόπο» (COM 2001, 59 τελικό).

Στόχος της Ημερίδας «Γλωσσική εκπαίδευση στην Κοινωνία της Πληροφορίας» που πραγματοποιήθηκε στο κτίριο του ΙΕΛ ήταν η ενημέρωση του ευρύτερου κοινού σε θέματα σύγχρονων μεθόδων γλωσσικής εκπαίδευσης με την βοήθεια των Νέων Τεχνολογιών.

2.2 Θέματα της ημερίδας

Την Ημερίδα προλόγησε ο Διευθυντής του ΙΕΛ, καθηγητής Γ. Καραγιάννης, και στη συνέχεια απηύθυναν χαιρετισμό ο πρόεδρος του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών, καθηγητής Γεώργιος Μπαμπινιώτης, ο πρόεδρος του Εθνικού Μετσοβίου Πολυτεχνείου, καθηγητής Θ. Ξανθόπουλος, και ο πρόεδρος του Πανεπιστημίου Κρήτης, καθηγητής Χ. Νικολάου.

Οι εργασίες της Ημερίδας περιελάμβαναν:

- Ομιλίες
- Συζήτηση
- Επίδειξη εκπαιδευτικού λογισμικού που έχει αναπτυχθεί από το ΙΕΛ και αφορά στην γλωσσική διδασκαλία

Οι εισηγήσεις των ομιλητών εντάχθηκαν σε τρεις κύκλους και αφορούσαν τόσο σε θέματα γενικού προβληματισμού για την γλωσσική εκπαίδευση όσο και σε δραστηριότητες που αναπτύσσονται στον ελληνικό χώρο, σε επίπεδο προσανατολισμού θεωρίας και εφαρμογών των Νέων Τεχνολογιών στην διδασκαλία της ελληνικής γλώσσας.

Ο Διευθυντής του ΙΕΛ καθηγητής Γ. Καραγιάννης ξεκίνησε τον πρώτο κύκλο των ομιλιών αναφερόμενος στις τάσεις που υπάρχουν σε θέματα ανάπτυξης εργαλείων γλωσσικής τεχνολογίας και στους επόμενους στόχους της τεχνολογικής έρευνας. Η καθηγήτρια Β. Τοκατλίδου προσέγγισε την γλωσσική διδασκαλία με την ευαισθησία της ειδικού σε θέματα Διδακτικής των Γλωσσών και ο καθηγητής Χ. Χαραλαμπίδης έκανε

μία εμπειριστατωμένη περιήγηση στις γλωσσικές ποικιλίες στην Ευρώπη.

Στον δεύτερο κύκλο των εισηγήσεων, ο Γ. Σταϊνχάουερ, ερευνητής του ΙΕΛ, παρουσίασε διάφορα εργαλεία επεξεργασίας φωνής και κειμένου που μπορούν να ενσωματωθούν σε περιβάλλοντα ηλεκτρονικής μάθησης. Η Ι. Αντωνίου-Κρητικού, ερευνήτρια του ΙΕΛ, επικέντρωσε την ομιλία της στις δυνατότητες των Νέων Τεχνολογιών, αναφέροντας παραδείγματα πολυμεσικών εφαρμογών του ΙΕΛ για την γλωσσική διδασκαλία. Ο Γ. Παπαδόπουλος, σύμβουλος του ΠΙ, παρουσίασε τις δραστηριότητες του ΠΙ σε ό,τι αφορά τις διαδικασίες και τα κριτήρια που εφαρμόζονται για τον έλεγχο ποιότητας εκπαιδευτικού λογισμικού. Τον δεύτερο κύκλο ομιλιών έκλεισε η Β. Καγκά, σύμβουλος του ΠΙ, η οποία αναφέρθηκε σε θέματα εκμάθησης ξένων γλωσσών και ειδικότερα, στους στόχους του Ευρωπαϊκού Portfolio Γλωσσών.

Ο τρίτος κύκλος εισηγήσεων αφορούσε στην γλωσσική εκπαίδευση κάτω από το πρίσμα συγκεκριμένων εφαρμογών του ΙΕΛ για την διδασκαλία της ελληνικής ως μητρικής, ως δεύτερης και ως ξένης γλώσσας, καθώς και για τη διδασκαλία των αρχαίων ελληνικών. Συγκεκριμένα, η ομιλία του Β. Κατσούρου, ερευνητή του ΙΕΛ, επικεντρώθηκε στο ευρωπαϊκό έργο «MYTHE». Η Μ. Τζεβελέκου, ερευνήτρια του ΙΕΛ, αναφέρθηκε στο ευαίσθητο θέμα της γλωσσικής εκπαίδευσης σε μειονοτικά σχολεία παρουσιάζοντας τα προγράμματα «Ένας ιππότης στο κάστρο των γραμμάτων» και «Ένας ιππότης στο κάστρο των λέξεων». Η Φ. Χαραλαμπούλου, ερευνήτρια του ΙΕΛ, αναφέρθηκε στην ανάπτυξη εκπαιδευτικού λογισμικού για την διδασκαλία της Ελληνικής ως ξένης γλώσσας δίνοντας έμφαση στην «Φιλογλωσσία» το ειδικευμένο προϊόν του ΙΕΛ. Τον τρίτο κύκλο ομιλιών έκλεισε η Ν. Οικονόμου, ερευνήτρια του ΙΕΛ, η οποία αναφέρθηκε σε πολυμεσικές εφαρμογές που έχουν αναπτυχθεί από το ΙΕΛ για την διδασκαλία των Αρχαίων Ελληνικών στο Γυμνάσιο και το Λύκειο.

Οι εργασίες της ημερίδας ολοκληρώθηκαν με μια εποικοδομητική συζήτηση που παρατάθηκε χρονικά εξ αιτίας των πολλών ερωτήσεων και παρεμβάσεων των συμμετεχόντων. Μετά την ολοκλήρωση των εργασιών της Ημερίδας, οι παριστάμενοι είχαν την ευκαιρία να

⁴ Ι. Αντωνίου-Κρητικού, *op.cit.* σ. 66.

παρακολουθήσουν επίδειξη λογισμικού στο χώρο έκθεσης των προϊόντων του ΙΕΛ, να συνομιλήσουν με τους ερευνητές του ΙΕΛ και να ανταλλάξουν απόψεις. Πολλοί ήταν αυτοί που τίμησαν με την παρουσία τους την Ημερίδα του ΙΕΛ και μάλιστα, από διάφορα μέρη της Ελλάδας (Λαμία, Άμφισσα, Ξάνθη). Ευχή όλων ήταν η εκδήλωση αυτή να συμβάλει ουσιαστικά στην προώθηση της έρευνας για την παιδαγωγική και μαθησιακή αξιοποίηση των Τεχνολογιών της Πληροφορίας και της Επικοινωνίας, καθώς και στον συντονισμό των φορέων για περαιτέρω ανάπτυξη και αξιοποίηση των μέχρι στιγμής επιτευγμάτων στον συγκεκριμένο χώρο.

3. Πρακτικά Ημερίδας

ISBN: 960-8191-14-9

160 σελίδες

Περιεχόμενα

- Πρόλογος
- Δυνατότητες συστημάτων πολυμέσων για εκπαιδευτικές εφαρμογές – Τεχνολογικές τάσεις και εργαλεία
Γιώργος Καραγιάννης



- Κοινωνία της πληροφορίας και γλωσσική εκπαίδευση. Στόχοι και προσεγγίσεις
Βάσω Τοκατλίδου
- Εκπαίδευση και γλωσσική ποικιλία στην Ευρώπη
Χριστόφορος Χαραλαμπίδης
- Νέες τεχνολογίες και υπολογιστικά εργαλεία στη γλωσσική εκπαίδευση
Γρηγόρης Σταϊνχάουερ, Στέλιος Μπακαμίδης, Γεώργιος Καραγιάννης, Ιωάννα Αντωνίου, Φρίντα Χαραλαμποπούλου, Νάντια Οικονόμου
- Αξιοποίηση των Νέων Τεχνολογιών στη γλωσσική διδασκαλία: παραδείγματα εφαρμογών
Ιωάννα Αντωνίου-Κρητικού
- Έλεγχος Ποιότητας Εκπαιδευτικού Λογισμικού
Γιώργος Παπαδόπουλος
- Ανάπτυξη γλωσσικών δεξιοτήτων: προϋπόθεση για την πρόσβαση στην Κοινωνία της γνώσης και της Πληροφορίας
Βάλια Καγκά
- ΜΥΤΗΕ – Πολυγλωσσική πλατφόρμα για την παράλληλη διδασκαλία μητρικής και δεύτερης γλώσσας
Βασίλης Κατσούρος
- Γλωσσικός και υπολογιστικός εγγραμματισμός: ένα υπολογιστικό σύστημα για τη διδασκαλία της ελληνικής ως δεύτερης γλώσσας
Μαρία Τζεβελέκου
- Αρχές σχεδιασμού και ανάπτυξης εκπαιδευτικού λογισμικού για τη διδασκαλία της Ελληνικής ως ξένης γλώσσας
Φρίντα Χαραλαμποπούλου
- Διδασκαλία Αρχαίων Ελληνικών στο Γυμνάσιο και Λύκειο: Παιδαγωγικοί Στόχοι και Εκπαιδευτικό Λογισμικό
Νάντια Οικονόμου
- Πίνακας ομιλητών

B. "European Language Technologies: Actions and Prospects"

Athina Papatheodorou, *ILSP*

The Greek National Focal Point (NFP) for the EUROMAP Language Technologies project organised the second in a series of three large scale events for the wider promotion of Human Language Technologies in Greece. The project consists a European funded initiative aiming at the creation of opportunities to increase demand for results of national and European R&D in the field of HLT.

The event, which took place on October 24th 2001 at the ILSP premises, was part of a bigger celebration organised by ILSP in the framework of the European Year of Languages. The underlying concept was to have a double event, where participants from various backgrounds would have the chance to receive information on the latest trends in language teaching within a new technologies environment (morning session) and on the current status of European Commission HLT projects and prospects as well as on trends and developments in the Greek HLT scene in terms of new services and technology transfer opportunities (afternoon session), as presented by eminent speakers. Demonstrations of language technology tools and applications were especially arranged for the audience, during the break between the morning and afternoon session.

Programme and speakers

Anastasios Patrikakos chaired the session, which included the following presentations:

- **Opening Speech** by George Carayannis, ILSP Director
- **EUROMAP Language Technologies**
Bente Maegaard, CST Director, Denmark
- **EUROMAP Language Technologies - The Greek NFP**
Athina Papatheodorou, ILSP
- **Language Technology: Achievements & Visions**
Jean Paul Haton, Loria Institute, France
- **The IST Programme: Current Status**
Xenophon Tsilibaris, GRNET

- **The eContent Programme: Current Status**
Stelios Piperidis, ILSP
- **VoiceWeb: Building the phonetic Internet - B2C and B2B voice applications**
Nikos Patsis, VoiceWeb S.A.

The session ended with a round table discussion.

Participants

The total of 83 participants who attended one or both events belonged to different communities, with the morning session attracting mostly people involved in the educational system. The Euromap session, carried out in the afternoon, had less participants, as was expected, but more representative ones, holding senior posts in their organisations. It is remarkable to notice this wide range of participants: from leading academics (e.g. from the National Technical University of Athens and the University of Ioannina), and researchers (e.g. from "Demokritos") to representatives from key players (e.g. the Hellenic Telecommunications Organisation and Intrasoft International S.A.) and IT companies (e.g. EXODUS S.A.), as well as the National Documentation Center represented by its Director. This proved to be a significant factor for the success of the event, especially for the interesting discussion that followed the presentations.

Results

The organisation of this National Event provided an excellent opportunity for the NFP to better identify and establish contacts with current and future key players of the Information Society in Greece. Even further than that, all participants could capitalise on the variety and representativeness of the audience, to develop or strengthen interaction and cooperation.

The session started with the presentation of the EUROMAP project. It was a pleasure to welcome Bente Maegaard, the project coordinator, who strongly presented the goals, the parallel actions and the allocated resources of EUROMAP, emphasising at the same time the successes and the aspirations for the future. The EUROMAP actions in a

local level and the accomplishments of the Greek NFP were presented immediately after by Athina Papatheodorou.

Shifting towards more specialised fields of Language Technology, the invited speaker from Loria Institute in France, Prof. Jean Paul Haton, talked about speech applications and related restrictions, current research and estimated prospects.

In the third part, in terms of content, useful information on EU projects and related statistics, data and figures, trends and prospects were presented, with great emphasis given to the IST and eContent initiatives and opportunities. Xenophon Tsilibaris, former coordinator of the Greek EUROMAP NFP and current IST representative in E.C., discussed the project Action Lines and Calls and presented interesting data regarding Greek participation and possibilities for Greek companies to enter a consortium. Stelios Piperidis, eContent representative to the E.C., carried on the discussion on EC funding opportunities in the framework of the upcoming eContent Calls for proposals.

Nikos Patsis from VoiceWeb, a young but dynamic company from the private sector, made a very interesting demonstration of the voice portal developed by his company. The voice driven access to different kinds of information with a simple dial of (+3010)-810-8000 captured the attention of the participants and made a statement of a user-friendly every-day language technology application.

The spontaneous, in a relaxed tone but highly informative and interesting discussion that followed the presentations gave a distinct character to the event and led naturally to its closing. Topics of discussion included the state of HLT in the 6th Framework Programme and the role of SMEs in future HLT and eContent projects. Bente Maegaard put the closing remark.

Third International Conference on Language Resources and Evaluation

LREC 2002

Las Palmas, Canary Islands (Spain)
27th May - 2nd June 2002
<http://www.lrec-conf.org/>



Hosted by
**The University of Las Palmas
de Gran Canaria,
Facultad de Informatica**

Main Conference:

29-30-31 May 2002

Workshops:

27-28 May and 1-2 June 2002

With support of TELEFONICA Foundation (of Spain) and support sought from the Commission of the EU and other institutions.

We are pleased to announce that the third International Conference on Language Resources and Evaluation (LREC) will take place in Las Pas, Canary Islands - Spain.

The Third International Conference on Language Resources and Evaluation has been launched by ELRA and is organised in cooperation with other Associations and Consortia, including ACL, AFNLPA, ALLC, CLASS, COCODA, ORIENTAL COCO-SDA, EAFT, EAGLES/ISLE, ELSNET, ENABLER, EURALEX, FRANCIL, ISCA, LDC, ONTOWEB, PAROLE, etc., and with major national and international organisations, including the European Commission - DG XIII, ARPA, NSF, and the Japanese Project for International Coordination of East-Asian Spoken Language Resources and Evaluation. Co-operation with other institutions is currently being sought.

CONFERENCE AIMS

In the framework of the Information Society, the pervasive character of Human Language Technologies

(HLT) and their relevance to practically all the fields of Information Society Technologies (IST) has been widely recognised.

Two issues are considered particularly relevant: the availability of language resources and the methods for the evaluation of resources, technologies, products and applications. Substantial mutual benefits can be expected from addressing these issues through international cooperation.

The term language resources (LR) refers to sets of language data and descriptions in machine readable form, used e.g. for: building and evaluating natural language, speech and multimodal algorithms or systems, software localisation industries and language services, language enabled information and communication services, natural interactivity, knowledge management, electronic commerce, electronic publishing, language studies, subject-area specialists and end users.

Examples of linguistic resources are written, spoken and multimodal corpora and lexica, grammars, terminology databases, multimedia databases, basic software tools for the acquisition, preparation, collection, management, customisation and use of these and other resources.

The relevance of evaluation for Language Technologies is increasingly recognised. This involves assessment of the state-of-the-art for a given technology, measuring the progress achieved within a programme, comparing different approaches to a given problem and choosing the best solution, knowing its advantages and disadvantages, assessment of the availability of technologies for a given application, product benchmarking, and assessment of system usability and user satisfaction.

In the recent past, language engineering and research and development in language technologies have led to important advances in various aspects of written, spoken and multimodal language processing. Although the evaluation paradigm has been studied and used in large national and international programmes, including the US DARPA HLT programme, the EU HLT programme under FP5-IST,

the Francophone AUF programme and others, particularly in the localisation industry (LISA and LRC), it is still subject to substantial unsolved basic research problems. The European 6th Framework program (FP6), planned for a start in 2003, includes multilingual and multisensorial communication as one of the major R&D issue, and the evaluation of technologies appears as a specific item in the Integrated Project instrument presentation.

The aim of this Conference is to provide an overview of the state-of-the-art, discuss problems and opportunities, exchange information regarding language resources, their applications, ongoing and planned activities, industrial use and requirements, discuss evaluation methodologies and demonstrate evaluation tools, explore possibilities and promote initiatives for international cooperation in the areas mentioned above.

CONFERENCE TOPICS

The following non-exhaustive list gives some examples of topics which could be addressed by papers submitted to the Conference:

Issues in the design, construction and use of Language Resources (LR)

- Guidelines, standards, specifications, models and best practices for LR
- Methods, tools, procedures for the acquisition, creation, management, access, distribution, use of LR
- Organisational issues in the construction, distribution and use of LR
- Legal aspects and problems in the construction, access and use of LR
- Availability and use of generic vs. task/domain specific LR
- Methods for the extraction and acquisition of knowledge (e.g. terms, lexical information, language modelling) from LR
- Monolingual and multilingual LR
- Multimodal and multimedia LR
- Integration of various modalities in LR (speech, vision, language)
- Documentation and archiving of languages,

including minority and endangered languages

- Ontological aspects of creation and use of LR
- LR for psycholinguistic and sociolinguistic research in human-machine communication
- Exploitation of LR in different types of applications (information extraction, information retrieval, vocal and multisensorial interfaces, translation, summarisation, www services, etc.)
- Industrial LR requirements and community's response
- Industrial production of LR
- Industrial use of LR
- Analysis of user needs for LR
- Internet-accessible metadata descriptions of LR
- Mechanisms of LR distribution and marketing
- Economics of LR

Issues in Human Language Technologies evaluation

- Evaluation, validation, quality assurance of LR
- Benchmarking of systems and products; resources for benchmarking and evaluation
- Evaluation in written language processing (text retrieval, terminology extraction, message understanding, text alignment, machine translation, morphosyntactic tagging, parsing, semantic tagging, word sense disambiguation, text understanding, summarization, localization, etc.)
- Evaluation in spoken language processing (speech recognition and understanding, voice dictation, oral dialog, speech synthesis, speech coding, speaker and language recognition, spoken translation, etc.)
- Evaluation of document processing (document recognition, on-line and off-line machine and hand-written character recognition, etc.)
- Evaluation of (multimedia) document retrieval and search systems (including detection, indexing, filtering, alert, question answering, etc)
- Evaluation of multimodal systems
- Qualitative and perceptive evaluation
- Evaluation of products and applications, benchmarking
- Blackbox, glassbox and diagnostic evaluation of systems
- Situated evaluation of applications
- Evaluation methodologies, protocols and measures
- From evaluation to standardisation of LR

General issues

- National and international activities and projects
- LR and the needs/opportunities of the emerging multimedia cultural industry
- Priorities, perspectives, strategies in the field of LR national and international policies
- Needs, possibilities, forms, initiatives of/for international cooperation
- Open architectures for LR

Pre-conference workshops:

- **26/05/2002**
 1. International Workshop on Resources and Tools in Field Linguistics
- **27/05/2002**
 1. International Workshop on Resources and Tools in Field Linguistics (Continuing)
 2. OntoLex 2002: Ontologies and Lexical Knowledge Bases
 3. Machine Translation Evaluation: Human Evaluators Meet Automated Metrics
 4. Annotation Standards for Temporal Information in Natural Language
- **28/05/2002**
 5. Customizing knowledge in NLP applications
 6. Question Answering: Strategy and Resources
 7. Language Resources in Translation Work and Research
 8. International Standards of Terminology and Language Resources Management
 9. Workshop on Wordnet Structures and Standardization, and how these affect Wordnet Applications and Evaluation

Post-conference workshops:

- **01/06/2002**
 10. First International Workshop on UNL, other Interlinguas and their Applications
 11. Arabic Language Resources (LR) and Evaluation: Status and Prospects
 12. Multimodal Resources and Multimodal Systems Evaluation
 13. Portability Issues in Human Language Technologies (HLT)
 14. Linguistic Knowledge Acquisition and Representation: Bootstrapping Annotated Language Data

- **02/06/2002**

15. Using Semantics for Information Retrieval and Filtering: State of the Art and Future Research
16. Towards a Roadmap for Multimodal Language Resources and Evaluation
17. Event Modelling for Multilingual Document Linking
18. Beyond PARSEVAL: Towards Improved Evaluation Measures for Parsing Systems

Pre-conference workshops:

- **26/05/2002**

1. International Workshop on Resources and Tools in Field Linguistics

- **27/05/2002**

1. International Workshop on Resources and Tools in Field Linguistics (Continuing)
2. OntoLex 2002: Ontologies and Lexical Knowledge Bases
3. Machine Translation Evaluation: Human Evaluators Meet Automated Metrics
4. Annotation Standards for Temporal Information in Natural Language

- **28/05/2002**

5. Customizing knowledge in NLP applications
6. Question Answering: Strategy and Resources
7. Language Resources in Translation Work and Research
8. International Standards of Terminology and Language Resources Management
9. Workshop on Wordnet Structures and Standardization, and how these affect Wordnet Applications and Evaluation

Post-conference workshops:

- **01/06/2002**

10. First International Workshop on UNL, other Interlinguas and their Applications
11. Arabic Language Resources (LR) and Evaluation: Status and Prospects
12. Multimodal Resources and Multimodal Systems Evaluation
13. Portability Issues in Human Language Technologies (HLT)

14. Linguistic Knowledge Acquisition and Representation: Bootstrapping Annotated Language Data

- **02/06/2002**

15. Using Semantics for Information Retrieval and Filtering: State of the Art and Future Research
16. Towards a Roadmap for Multimodal Language Resources and Evaluation
17. Event Modelling for Multilingual Document Linking
18. Beyond PARSEVAL: Towards Improved Evaluation Measures for Parsing Systems

CONFERENCE PROGRAMME COMMITTEE

Nicoletta Calzolari, Istituto di Linguistica Computazionale,
Pisa, Italy

Khalid Choukri, ELRA, Paris, France

Bente Maegaard, CST, Copenhagen, Denmark

Joseph Mariani, LIMSI-CNRS, Orsay, France

Angel Martin Municio, President of the Real Academia
de Ciencias, Madrid, Spain

Daniel Tapias, TELEFONICA I+D, Madrid, Spain

Antonio Zampolli, Istituto di Linguistica Computazionale,
Pisa, Italy (Conference chair)

INTERNATIONAL ADVISORY COMMITTEE

Sture Allen, professor, former permanent secretary of
the Swedish Academy, Sweden

Souguil Ann, Seoul National University, Korea

Georges Carayannis, Institute for Language and
Speech Processing, Athens, Greece

Roberto Cencioni, Commission of the EU, Information
Society DG, Luxembourg

Zhiwei Feng, The State Language Commission of
China, Beijing, China

Hiroya Fujisaki, Science University of Tokyo, Japan

Mark Maybury, MITRE Corporation, Boston, USA

Bernard Quemada, Conseil Supérieur de la Langue
Française, Paris, France

Gary Strong, NSF, Washington D.C., USA

Piet G.J. Van Sterkenburg, International Permanent

Committee of Linguists, Leiden, The Netherlands
Giovanni Battista Varile, Commission of the EU,
Information Society DG, Luxembourg
Charles Wayne, DARPA, Washington D.C., USA

35017 Las Palmas de Gran Canaria
Tel.: +34 928 451865
Fax: +34 928 458760
Email: ocon@cicei.ulpgc.es
<http://www.cicei.ulpgc.es>

LOCAL COMMITTEE

Manual Gonzales Rodriguez (University of Las Palmas de Gran Canaria).
Carmen Paz Suarez Araujo (University of Las Palmas de Gran Canaria).
Antonio Ocon Carreras (University of Las Palmas de Gran Canaria).
Luis Jollanes Aguilar (Universidad Pontificia de Salamanca).
Angel Martin Municio (President of the Real Academia de Ciencias, Madrid).
A representative of Telefonica Foundation.

CONFERENCE ADDRESSES

The **Secretariat of the Conference**, which will provide general information on the Conference is:

LREC 2002 Secretariat
Universidad Las Palmas de Gran Canaria
Facultad de Informatica
Edificio Informatica y Matematicas
Campus de Tafira
35017 Las Palmas de Gran Canaria
Tel: +34 928 458719
Fax: +34 928 458760
E-mail: LREC@elda.fr
<http://www.fi.ulpgc.es>

The **Conference Chair** is:

Antonio Zampolli
Istituto di Linguistica Computazionale del CNR
Area della Ricerca di Pisa
Via G. Moruzzi, 1
56124 Pisa - Italy

The **demonstration organiser** is:

Antonio Ocon Carreras
Universidad Las Palmas de Gran Canaria
Facultad de Informatica
Edificio Informatica y Matematicas
Campus de Tafira

3^ο ΣΥΝΕΔΡΙΟ ΓΙΑ ΤΗΝ ΟΡΟΛΟΓΙΑ

Βασίλης Α. Φιλόπουλος, Κώστας Ε. Βαλεοντής

Την 1η και στις 2 και 3 Νοεμβρίου 2001 διεξήχθη στην Αθήνα το **3ο Συνέδριο «Ελληνική Γλώσσα και Ορολογία»**. Αυτό το Συνέδριο θα μπορούσε να χαρακτηριστεί Πανελλήνιο, καθώς στη διοργάνωσή του συμμετείχαν δέκα φορείς που καλύπτουν σχεδόν το σύνολο όσων ασχολούνται και εργάζονται για την Ορολογία στην Ελλάδα. Συγκεκριμένα το 3ο Συνέδριο διοργανώθηκε από την Ελληνική Εταιρεία Ορολογίας (**ΕΛΕΤΟ**), με συμμετοχή του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών (**ΕΚΠΑ**), του Αριστοτέλειου Πανεπιστημίου Θεσσαλονίκης (**ΑΠΘ**), του Πανεπιστημίου Πατρών (**ΠΠ**), του Ιόνιου Πανεπιστημίου (**ΙΠ**), του Γαλλικού Ινστιτούτου Αθηνών (**ΙΦΑ**), του Ελληνικού Οργανισμού Τυποποίησης (**ΕΛΟΤ**), του Ινστιτούτου Επεξεργασίας του Λόγου (**ΙΕΛ**), του Οργανισμού για τη Διεθνοποίηση της Ελληνικής Γλώσσας (**ΟΔΕΓ**) και του Πανελληνίου Συλλόγου Επαγγελματιών Μεταφραστών (**ΠΣΕΜ**).

Με την πραγματοποίηση του 3ου Συνεδρίου βεβαιώνεται ότι τα συνέδρια για την Ορολογία αποτελούν πια θεσμό στην Ελλάδα. Εκτός από το **1ο Συνέδριο** το 1997 και το **2ο Συνέδριο** το 1999, **πρόδρομο συνέδριο** αποκλειστικά για την Ορολογία ήταν το Συνέδριο «Τυποποίηση Ορολογίας» που διοργανώθηκε από το Τεχνικό Επιμελητήριο της Ελλάδας, το 1992. Δηλαδή την ίδια εποχή που ιδρύθηκε η Ελληνική Εταιρεία Ορολογίας. Δεν υπάρχει εμφανώς καμιά σχέση ανάμεσα στα δύο ταυτόχρονα γεγονότα. Όμως, η πλειονότητα των μελών της Οργανωτικής Επιτροπής και των ομιλητών ήταν ιδρυτικά μέλη της ΕΛΕΤΟ. Η ιδέα για την πραγματοποίηση εκείνου του Συνεδρίου ήταν της Μαίρης Παπαδάκη, υπαλλήλου του ΤΕΕ, ιδρυτικού μέλους της ΕΛΕΤΟ και Τεχνικής Υπευθύνου της Τεχνικής Επιτροπής ΕΛΟΤ/ΤΕ21 «Τεχνική Ορολογία». Δυστυχώς η Μαίρη Παπαδάκη δεν πρόφτασε να λάβει μέρος στο 3ο Συνέδριο, αφού μας άφησε λίγες μέρες πιο πριν. Η μεγάλη επιτυχία του συνεδρίου «Τυποποίηση Ορολογίας» έδειξε το ενδιαφέρον για θέματα Ορολογίας και οδήγησε στην πραγματοποίηση των επόμενων συνεδρίων και στην προσπάθεια ανόδου της στάθμης τους.

Η **εναρκτήρια συνεδρίαση** του 3ου Συνεδρίου έγινε την Πέμπτη 1η Νοεμβρίου, στη **Μεγάλη Αίθουσα Τε-**



λετών του ιστορικού κτιρίου του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών, παρουσία προσκεκλημένων μελών της πολιτικής ηγεσίας, καθηγητών πανεπιστημίων και άλλων διακεκριμένων προσώπων.

Μετά την υποδοχή από τον αντιπρύτανη του Πανεπιστημίου Αθηνών καθ. Αντ. Κουτσελίνη, έλαβαν το λόγο οι πρόεδροι της Οργανωτικής και της Επιστημονικής Επιτροπής και, στη συνέχεια, απηύθυναν χαιρετισμούς οι εκπρόσωποι των συνοργανωτών.

Καθώς το 3ο Συνέδριο ήταν **αφιερωμένο στο Σωκράτη**, για τα 2400 έτη από το θάνατό του, τον μεγάλο Έλληνα φιλόσοφο που μέσα από τον πλατωνικό διάλογο «Κρατύλος» μίλησε πρώτος για την αξία της γνώσης των ονομάτων και τη σχέση τους με την πραγματικότητα. Με θέμα «Η παρουσία του Σωκράτη σήμερα» ο καθηγητής Ε. Μικρογιαννάκης εκφώνησε τον πανηγυρικό της ημέρας.

Η πανηγυρική συνεδρίαση ολοκληρώθηκε με δεξίωση που προσφέρθηκε από την ΕΛΕΤΟ στους συνέδρους και στους προσκεκλημένους.

Το **3ο Συνέδριο**, όπως και τα προηγούμενα συνέδρια για την Ορολογία, αποσκοπούσε στην **παρουσίαση της Ελληνικής Γλώσσας και Ορολογίας στις σημερινές συνθήκες**, όπως διαμορφώνονται στο ευρωπαϊκό και παγκόσμιο πολύγλωσσο περιβάλλον. Η μεγάλη απήχηση που έχουν πια τα συνέδρια για την Ορολογία εκδηλώθηκε με την αθρόα υποβολή ανακοινώσεων στο 3ο Συνέδριο που ανήρχοντο σε 46. Αυτό προβλημάτισε πάρα πολύ την Επιστημονική Επιτροπή για την επιλογή των πιο αξιόλογων. Το αποτέλεσμα ήταν να επιλεγούν 33 και να αποκλειστούν 13 όχι μόνο για

λόγους χαμηλής επιστημονικής στάθμης αλλά και επειδή δεν ενέπιπταν στη θεματολογία που είχε καθοριστεί. Οι υψηλής στάθμης προδιαγραφές και η απαίτηση παράδοσης του πλήρους κειμένου κάθε ανακοίνωσης – και μάλιστα σε συγκεκριμένη ηλεκτρονική μορφή – σε απaráβατη χρονική προθεσμία, αρκετά πριν από το Συνέδριο, ώστε η ανακοίνωση να ενταχθεί στο πρόγραμμα του Συνεδρίου και το κείμενο να συμπεριληφθεί στον τόμο των ανακοινώσεων, συνετέλεσαν ώστε στο πρόγραμμα να ενταχθούν, τελικά, **30 ανακοινώσεις**.

Οι **εργασίες του 3ου Συνεδρίου**, στις επόμενες δύο ημέρες, έγιναν στο **αμφιθέατρο του Γαλλικού Ινστιτούτου Αθηνών**. Οι 30 ανακοινώσεις ήταν κατανεμημένες σε **έξι ενότητες**, σύμφωνα με τη θεματολογία, ως εξής:

- **Ιστορική θεώρηση και Γλωσσολογικές-Οντολογικές Αρχές:** 7 ανακοινώσεις

Έγινε σύντομη ιστορική αναδρομή στην ανάπτυξη του κλάδου της Ορολογίας, εξετάστηκε η συμβολή της αρχαίας Ελληνικής και της Λατινικής στη διαμόρφωση της επιστημονικής ορολογίας των ευρωπαϊκών γλωσσών, διερευνήθηκε η σύνθεση και παραγωγή στην Ελληνική Γλώσσα για τον σχηματισμό μονολεκτικών σύμπλοκων όρων, επισημάνθηκε το φαινόμενο του «αναβαπτισμού» παλαιών εννοιών με εμφάνιση νέων ειδών κατά την ανάπτυξη της γνώσης, εξετάστηκαν ορισμένα προβλήματα απόδοσης γλωσσολογικών όρων, εξετάστηκε η σύγχρονη προσέγγιση παραγωγής, εκφοράς και χρήσης όρων με αναφορά στο έργο του Wüster και εξετάστηκε η γλώσσα της λογοτεχνικής κριτικής στη μέση εκπαίδευση.

- **Λεξικογραφία:** 3 ανακοινώσεις

Αναλύθηκε η σύνταξη και δομή δίγλωσσου γλωσσολογικού νομικών όρων του πεδίου της βιομηχανικής ιδιοκτησίας, δόθηκαν κριτήρια ανάλυσης δίγλωσσων και πολύγλωσσων λεξικών που περιλαμβάνουν την ελληνική και γαλλική, και παρουσιάστηκε ερμηνευτικό λεξικό μετεωρολογικών και κλιματολογικών όρων.

- **Τυποποίηση Ορολογίας:** 7 ανακοινώσεις

Διερευνήθηκε η απόδοση του όρου «design» και τα προβλήματά της, εξετάστηκε η απόδοση στα ελληνικά των γαλλικών όρων που αφορούν την μπαρόκ

σουίτα και γενικότερα τη μουσική ορολογία, έγινε μια διαδρομή στην ορολογία του χορού στην ελληνική πραγματικότητα, εκτέθηκαν προβληματισμοί σχετικά με την τυποποίηση της ορολογίας των νέων προϊόντων ξύλου, έγιναν προτάσεις για αναθεώρηση όρων της στατιστικής, επισημάνθηκαν διάφορες εννοιολογικές δυσχέρειες βασικών όρων της Χημείας και προτάθηκαν νέες αποδόσεις για μερικούς όρους της Ποιότητας.

- **Νέες Τεχνολογίες και Ορολογία:** 5 ανακοινώσεις

Περιγράφηκε τρόπος δημιουργίας δίγλωσσων τραπεζών όρων από «παράλληλα κείμενα» στα ελληνικά και σε άλλη γλώσσα της Ευρωπαϊκής Ένωσης με εφαρμογή στα Πορτογαλικά, εξετάστηκαν αποκλίσεις στην αυτόματη εξαγωγή πολυλεκτικών όρων από κείμενα, έγινε αναφορά στη διδακτική της μετάφρασης ειδικών πραγματολογικών κειμένων και τη χρήση νέων τεχνολογιών, περιγράφηκε ειδικό περιβάλλον διαχείρισης ορολογικών πόρων στο πλαίσιο συγκεκριμένου Έργου και παρουσιάστηκαν ηλεκτρονικά λεξικά της ελληνικής γλώσσας.

- **Διερμηνεία και Μετάφραση:** 5 ανακοινώσεις

Εξετάστηκε ο ρόλος του μεταφραστή ως δημιουργού νεολογισμών, διερευνήθηκε το φαινόμενο των συνωνύμων στις ειδικές γλώσσες και το πρόβλημα των αποδόσεών τους, έγινε αναφορά στην αρχιτεκτονική ορολογία, εξετάστηκε η ορολογική διασπαση της μεταφραστικής πρακτικής και εκτέθηκαν μέθοδοι αξιολόγησης και προοπτικές της διδακτικής της μετάφρασης.

- **Όργανα Ορολογίας:** 3 ανακοινώσεις

Έγινε αναδρομή στις πρόσφατες εξελίξεις για τη δημιουργία ελληνικών όρων και στο ιστορικό της ίδρυσης και ανάπτυξης της Ελληνικής Εταιρείας Ορολογίας και ανασκόπηση του μέχρι σήμερα επιτελεσθέντος έργου παραγωγής ελληνικών όρων.

Περιγράφηκε η στρατηγική μελέτη καθώς και η μεθοδολογία και οι στόχοι του Εθνικού Προγράμματος Ορολογικού Συντονισμού (**ΕΠΟΣ**) και αναλύθηκε το Επιχειρησιακό Πρόγραμμα του ΕΠΟΣ.

Μετά από κάθε ανακοίνωση οι σύνεδροι υπέβαλλαν ερωτήσεις και ακολουθούσε γόνιμη συζήτηση.

Οι εργασίες του Συνεδρίου ολοκληρώθηκαν με Ανοικτή Συζήτηση που είχε ως θέμα το «Εθνικό Πρόγραμμα Ορολογικού Συντονισμού» (ΕΠΟΣ).

Οι 30 ανακοινώσεις που ακούστηκαν στο 3ο Συνέδριο είχαν εκδοθεί πριν από το Συνέδριο σε καλαίσθητο βιβλίο 360 σελίδων. Το βιβλίο αυτό διανεμήθηκε στους συνέδρους κατά την έναρξη των κυρίως εργασιών του Συνεδρίου (2 Νοεμβρίου), έχει αποσταλεί σε βιβλιοθήκες πανεπιστημίων και άλλων ιδρυμάτων και διατίθεται σε κάθε ενδιαφερόμενο έναντι χαμηλού τιμήματος.

Κλείνοντας το 3ο Συνέδριο, ο Πρόεδρος της Οργανωτικής Επιτροπής ευχαρίστησε τους συνέδρους, και όσους συνέβαλαν στην επιτυχία του Συνεδρίου και εξήγγειλε τη διοργάνωση του 4ου Συνεδρίου «Ελληνική Γλώσσα και Ορολογία» για το έτος 2003.



Β. Α. Φιλόπουλος

Χημικός-Μηχανικός, τέως διευθυντικό στέλεχος του ΕΛΟΤ
Μέλος της επιτροπής ΕΛΟΤ/ΤΕ21,
Πρόεδρος της ΕΛΕΤΟ και
Μέλος του Γενικού Επιστημονικού Συμβουλίου (ΓΕΣΥ) της ΕΛΕΤΟ.

Διεύθυνση: Σωκράτη Τσάκωνα 5, 15236 ΠΕΝΤΕΛΗ
Τηλ.: 010 8042313
Τηλεομ.: 010 8042385

Κ. Βαλεοντής

Φυσικός-Ηλεκτρονικός, Τηλεπικοινωνιακός Μηχανικός ΟΤΕ,
Υπεύθυνος των Ομάδων: ΜΟΤΟ, ΕΛΟΤ/ΤΕ2/ΟΕ5, ΕΛΟΤ/ΤΕ48/ΟΕ1
και ΕΛΟΤ/ΤΕ80/ΟΕ1
Γενικός Γραμματέας της ΕΛΕΤΟ και
Πρόεδρος του Γενικού Επιστημονικού Συμβουλίου (ΓΕΣΥ) της ΕΛΕΤΟ.

Διεύθυνση: Καραμανλάκη 18, 112 53 ΑΘΗΝΑ
Τηλ.: 010 6118986, 010 8619521
Τηλεομ.: 010 8068299
Ηλ.Ταχ.: valeonti@otenet.gr

II. ΣΥΝΕΙΣΦΟΡΕΣ ΕΡΕΥΝΗΤΩΝ

Δυνατότητες συστημάτων πολυμέσων για εκπαιδευτικές εφαρμογές – Τεχνολογικές τάσεις και εργαλεία

Γιώργος Καραγιάννης
Ινστιτούτο Επεξεργασίας του Λόγου

1. Εισαγωγή

Η τεχνολογία των πολυμέσων αποτελεί μία νέα τεχνολογική διάσταση που δίνει την δυνατότητα να συνθέσει κανείς σε ενιαία σύνολα τις γραφικές τέχνες, την τυπογραφία, τον κινηματογράφο είτε στην ρεαλιστική μορφή του είτε υπό μορφή κινουμένων σχεδίων, όλα αυτά κάτω από τον έλεγχο του χρήστη. Επομένως μοιάζει να είναι μία 8η τέχνη. Είναι ταυτόχρονα η σύγχρονη εκδοτική τέχνη. Η τεχνολογία αυτή έχει δημιουργήσει εργαλεία λογισμικού συγγραφής στην διάθεση τόσο του μεμονωμένου δημιουργού, όσο και ομάδων δημιουργών που έχουν την διάθεση να αφιερώσουν τον χρόνο τους σε σύνθετες δημιουργίες. Το αποτέλεσμα είναι πολλές φορές εξαιρετικά δυναμικά περιβάλλοντα με μεγάλη χρησιμότητα στους τομείς: της εκπαίδευσης, της διασκέδασης, του τουρισμού, της διαφήμισης, της φιλικής επικοινωνίας με την μηχανή, της παροχής υπηρεσιών μέσω ολοκληρωμένων πληροφοριακών συστημάτων.

Τα σημερινά εκπαιδευτικά λογισμικά έχουν ποικιλία δυνατοτήτων και ξεκινούν από απλές εφαρμογές που διεξέρχεται κανείς σύντομα μέχρι εξαιρετικά πολύπλοκες δημιουργίες που μπορούν να χαρακτηρισθούν ως «ολόκληροι κόσμοι» περιέχοντες όλη την πληροφορία σε ένα θέμα, ενώ μπορούν να απασχολήσουν τον χρήστη πολλά χρόνια. Σε επίπεδο κατασκευής είναι δυνατόν ένα σύστημα να παραχθεί σε χρόνους που ποικίλουν από 2-3 ημέρες μέχρι 5-10 χρόνια. Έτσι υπάρχει ποικιλία στην επένδυση κόπου και πόρων. Στην δεύτερη περίπτωση των μεγάλων κατασκευών έχει ο κατασκευαστής το μειονέκτημα ότι πρέπει να αντιμετωπίσει την αλλαγή των τεχνολογιών των εργαλείων συγγραφής. Επομένως με άλλα εργαλεία “authoring” πιθανόν να ξεκινήσει και με άλλα να τελειώσει. Για τους λόγους αυτούς χρειάζεται μεγάλη

προσοχή στην σχεδίαση του λογισμικού. Όπως στην TV έτσι και στα πολυμέσα έχει αρχίσει η μόδα των "serials". Η δημιουργία ενός CD-ROM π.χ. κάθε μήνα με καινούργιο υλικό είναι εφικτή, όταν υπάρχει ή πλατφόρμα και έτσι μπορεί να αποτελεί μέρος μίας μηνιαίας έκδοσης ενός περιοδικού. Ένα εκπαιδευτικό σύστημα πολυμέσων μπορεί να αποτελεί ταυτόχρονα και μία αποθήκη πληροφοριών ή μία συλλογή ηλεκτρονικών αντικειμένων.

Μία ενδιαφέρουσα νοοτροπία θέλει ένα μέρος ενός CD-ROM να υπάρχει και στο διαδίκτυο. Αυτό αποτελεί ταυτόχρονα και διαφήμιση του προϊόντος, αλλά βοηθάει και τον χρήστη να καταλάβει αν τον ενδιαφέρει ένα προϊόν χωρίς να χάνει χρόνο αναζήτησης, επίδειξης κ.λπ. Τρία είναι τα κύρια κριτήρια στην επιλογή ενός CD-ROM:

- α) Η Χρησιμότητα
- β) Η Λειτουργικότητα
- γ) Η Αισθητική

Τελικά ο χρήστης τις περισσότερες φορές ξεγελιέται και επιλέγει ένα CD-ROM αποκλειστικά με κριτήριο την αισθητική που εντυπωσιάζει. Έτσι πολλές φορές η αισθητική μπορεί να δημιουργήσει μία «μάσκα», αν είναι πολύ προσεγμένη, σε λειτουργικές ατέλειες. Αν κάποιος έχει πεισθεί για την χρησιμότητα, το βασικότερο που πρέπει να εξετάσει είναι η λειτουργικότητα. Η λειτουργικότητα έχει σχέση και με την συστηματοποίηση. Δυστυχώς μέχρι σήμερα ακόμη υπάρχουν παραγωγές που θα μπορούσε κανείς να χαρακτηρίσει τύπου «λερναίας ύδρας» λόγω ακριβώς της έλλειψης συστηματοποίησης είτε της περιεχόμενης γνώσης είτε των λειτουργιών, ενώ έχουν αρχίσει να δημιουργούνται και παραγωγές με μεγάλη συστηματικότητα.

2. Σύνδεση των Πληροφοριών Διαφορετικών Μέσων

Η δυνατότητα σύνδεσης των πληροφοριών είναι όπως γνωρίζουμε ένα από τα μεγάλα πλεονεκτήματα των πολυμέσων. Η σύνδεση βοηθάει στην καλύτερη ιεράρχηση των γνώσεων που είναι αποθηκευμένες στις βάσεις δεδομένων της εφαρμογής και επιτρέπει στον χρήστη την πλοήγηση με ένα εκλογικευμένο και επαγωγικό τρόπο. Πέραν από τα συμβατικά "υπερκεί-

μενα", η σύνδεση των πληροφοριών είναι αντικείμενο έρευνας σήμερα ώστε να προσδιορισθούν εργονομίες σύνδεσης μεταξύ διαφορετικών μέσων που να έχουν αυξημένη λειτουργικότητα και να επιτρέπουν εύκολα την πρόσβαση από το ένα μέσο στο άλλο.

Πολλές κατηγορίες εκπαιδευτικού λογισμικού έχουν εξ ορισμού κειμενοκεντρική υπόσταση γιατί πρέπει ο μαθητής να μπορεί να χρησιμοποιεί και να επεξεργάζεται τα κείμενα. Αυτό συμβαίνει με την γλωσσική εκπαίδευση. Οι υπόλοιπες πληροφορίες (εικόνες, video, κινούμενα σχέδια) διανθίζουν τα κείμενα, αλλά είναι λιγότερο κεντρικές. Η σύνδεση των πληροφοριών γίνεται με βάση τα κείμενα.

Σε άλλες περιπτώσεις εκπαιδευτικού λογισμικού η παρουσίαση με video είναι χρήσιμη και τα κείμενα λειτουργούν συνοδευτικά, μόνο για την καλύτερη αποτύπωση των εννοιών ή στο πλαίσιο κάποιας ανάγκης επιβεβαίωσης. Όταν το video παίζει τον κύριο ρόλο είναι δυσκολότερο να ιεραρχήσει κανείς την υπάρχουσα γνώση γιατί το video επιβάλλει σε ένα βαθμό μία ακολουθιακή δομή που περιορίζει τις δυνατότητες ενδιαμέσων επιλογών. Υπάρχουν όμως κάποιες μεθοδολογίες για να σπάσει κανείς την ακολουθιακή δομή και να ακολουθήσει μία ιεραρχική προσέγγιση στην κατάκτηση της γνώσης.

Μία ειδική περίπτωση σύνδεσης κειμένου και video είναι αυτή της αμφιμονοσήμαντης αντιστοιχίας μεταξύ τους. Ειδικά στην περίπτωση διαλόγων που χρησιμοποιούνται στην γλωσσική εκπαίδευση, αλλά και σε άλλες περιπτώσεις, η ένα-προς-ένα αντιστοιχία είναι δύσκολο και δαπανηρό έργο, αλλά έχει μεγάλο αντίκτυπο στην λειτουργικότητα και προσφέρει στον χρήστη δυνατότητες ανάκτησης του video μέσα από κείμενο ή του κειμένου μέσα από το video. Όταν έχει κανείς να αντιμετωπίσει μεγάλες ακολουθίες video που θέλει να επαναχρησιμοποιήσει στην δημιουργία εκπαιδευτικών τίτλων, μία τεχνολογία που θα μπορούσε να βοηθήσει είναι η αναγνώριση φωνής. Θα μπορούσε με τον τρόπο αυτό να δημιουργηθεί το κείμενο και να γίνει η αντιστοιχία αυτόματα.

Η αποτελεσματική σύνδεση των πληροφοριών μεταξύ τους επηρεάζει πολύ και την διαλογικότητα ενός συστήματος. Έτσι όσο καλύτερη είναι η σύνδεση των

ποικίλων ειδών πληροφορίας σε ένα σύστημα πολυμέσων, τόσο περισσότερο αυξημένες δυνατότητες υλοποίησης της διαλογικότητας παρουσιάζονται.

Μία ειδική περίπτωση είναι και η στενή σύνδεση κειμένων μεταξύ τους. Αυτή η περίπτωση παρουσιάζεται σε κειμενοκεντρικές προσεγγίσεις διδασκαλίας γλωσσών όπου το ένα κείμενο πρέπει να υποστηρίζει το άλλο. Έτσι τα κείμενα αυτά μπορούν αν συνδεθούν με μία τεχνική γνωστή ως των παραλλήλων – στοιχισμένων κειμένων που βοηθάει στην αποκατάσταση αντιστοιχιών σε επίπεδο πρότασης και αν χρειασθεί και κάτω της πρότασης. Αυτές οι αντιστοιχίες μπορούν να πραγματοποιηθούν πρακτικά με χειρωνακτικά, ημιαυτόματα ή και πλήρως αυτόματα εργαλεία που δανείζει η γλωσσική τεχνολογία.

Η σύνδεση κειμένων με φωνή ώστε να υπάρχει πλήρης αντιστοίχιση είναι επίσης κάτι που στοιχίζει σε προσπάθεια και πόρους, ειδικά αν θέλουμε να υπάρχει πλήρης συγχρονισμός μεταξύ των δύο. Τότε πρέπει να λάβει κανείς υπ' όψη και τις πιθανές σιωπές που θα εισαχθούν στην διαδικασία αφήγησης. Υπάρχουν σήμερα εργαλεία αυτόματης επεξεργασίας για την εξαγωγή ή την μείωση της διάρκειας των σιωπών. Η λεπτομερής αντιστοίχιση κειμένου και φωνής παραμένει κάτι δύσκολο αν θέλει κανείς να επιτύχει αυξημένη διαλογικότητα π.χ. με ένα κλικ σε οποιοδήποτε σημείο του κειμένου να μπορεί να ακούσει το συγκεκριμένο σημείο. Η οριστική λύση που θα συντελέσει και στην δραματική μείωση του κόστους αυτής της διαδικασίας είναι η χρήση συστημάτων μετατροπής κειμένου σε φωνή με απαιτήσεις φυσικής ποιότητας συνθετικής φωνής. Πολύ σύντομα θα υπάρχει διαθεσιμότητα φυσικής ποιότητας συνθετικής φωνής στις περισσότερες ευρωπαϊκές γλώσσες αφού ωρίμασε πλέον αυτή η τεχνολογία και ετοιμάζεται τα προσεχή χρόνια να βγει από το εργαστήριο. Θα είναι επομένως διαθέσιμη στους παραγωγούς πολυμέσων οι οποίοι θα διευκολύνουν χάρη σ' αυτά τα εργαλεία την όλη διαδικασία μειώνοντας σοβαρά το κόστος.

3. Ανάκτηση των Πληροφοριών από Συστήματα Πολυμέσων

Η ανάκτηση της πληροφορίας σχετίζεται με την οργάνωση της πληροφορίας στο εκπαιδευτικό λογισμικό. Συνήθως υπάρχουν τα μαθήματα που αντιμετωπι-

ζουν το καθένα μία συγκεκριμένη ύλη και είναι οργανωμένα σε κατάλληλα σενάρια και η υπόλοιπη ύλη που χρησιμεύει στην αντιμετώπιση πρόσθετων αναγκών του μαθητή. Η επί πλέον αυτή ύλη μπορεί να οργανωθεί σε βάσεις δεδομένων, επομένως να τεθούν στην διάθεση του μαθητή όλες οι δυνατότητες ανάκτησης μέσω βάσεων δεδομένων ώστε να μπορεί να αναζητήσει τις πληροφορίες που του είναι απαραίτητες.

Πέραν των κλασικών μεθοδολογιών ανάκτησης τα συστήματα πολυμέσων μπορούν να επωφεληθούν από αριθμό σύγχρονων μεθόδων που βασίζονται κυρίως στις κατακτήσεις της γλωσσικής τεχνολογίας. Μία ενδιαφέρουσα μεθοδολογία που μπορεί να υποστηρίξει την ανάκτηση σε κειμενικά περιβάλλοντα με ισχυρή μορφολογική ποικιλία, όπως συμβαίνει στην ελληνική γλώσσα, είναι αυτή του «ασαφούς ταιριάσματος». Μπορεί έτσι να ανακτήσει κανείς παραγράφους ή φράσεις που περιέχουν συγκεκριμένες γλωσσικές πολυλεκτικές πραγματώσεις οι οποίες είναι πιθανόν αντικείμενο αναζήτησης, ακόμη και όταν υπάρχουν αποκλίσεις στην καταγραφή ή/και λαθών.

Η μεθοδολογία αυτή μπορεί να χρησιμοποιηθεί για ανάκτηση κειμενικής πληροφορίας και μετά από αναγνώριση φωνής, αν ο χρήστης θέλει να αποφύγει την πληκτρολόγηση. Η αναγνώριση φωνής μπορεί επίσης να χρησιμοποιηθεί για την ανάκτηση φωνητικής πληροφορίας που είναι αποθηκευμένη στα φωνητικά αρχεία της εφαρμογής. Μία ακόμη δυνατότητα που θα μπορούσε να είναι χρήσιμη στα συστήματα πολυμέσων όταν αυτά περιέχουν πληροφορίες έντασης video, είναι η ανάκτηση μετά από συνδυασμένη αναγνώριση φωνής και εικόνας για να δίνεται στον χρήστη η δυνατότητα να εντοπίσει π.χ. έναν ομιλητή την στιγμή που αναφέρεται σε κάποιο συγκεκριμένο θέμα.

4. Παιδαγωγικές Αρχές

Είναι πια γνωστό πόσο και πώς μπορούν να βοηθήσουν τα πολυμέσα στην επίτευξη συγκεκριμένων εκπαιδευτικών στόχων. Επειδή κεντρίζουν τις αισθήσεις και ανεξαρτητοποιούν τον μαθητή, δίνοντας του την δυνατότητα να συμμετέχει ενεργά λόγω της διαλογικότητας αντί να είναι παθητικός δέκτης, ξεπερνούν σε δυνατότητες εμπέδωσης τις συμβατικές εκπαιδευτικές μεθοδολογίες. Η ανεξαρτητοποίηση του μαθητή έχει πολλές διαστάσεις. Κατ' αρχάς κατάργηση της πίεσης του χρόνου, στην συνέχεια δυνατότητες χρήσης του συ-

στήματος την κατάλληλη στιγμή όταν υπάρχει έφεση στον μαθητή, επίσης μεγάλη ευκολία στην επανάληψη και τέλος αφαίρεση της επιβολής που σε μερικές κατηγορίες μαθητών μπορεί να δημιουργεί διασπάσεις. Ειδικά τα παιδιά που έχουν δυσκολίες στην συγκέντρωση, είναι ιδιαίτερα ενεργητικά, ή τα παιδιά που είναι φοβισμένα μπορούν να επωφεληθούν από την χρήση των πολυμέσων σε αυτοδιδασκαλία συμπληρωματικά με την σχολική εκπαίδευση.

Το εκπαιδευτικό λογισμικό πολυμέσων πρέπει να είναι ιδιαίτερα προσεγμένο επειδή ακριβώς δίνει την δυνατότητα βιωματικής μάθησης χάρη στην ταυτόχρονη στράτευση των αισθήσεων δημιουργώντας αξέχαστες εντυπώσεις στα παιδιά. Έτσι δεν πρέπει να υπάρχουν καθόλου λάθη στο υλικό ενώ η διαλογικότητα πρέπει να σχεδιάζεται με την βοήθεια ειδικών παιδαγωγών και ανθρώπων που έχουν μελετήσει την διδακτική πλευρά ενός εκπαιδευτικού αντικειμένου.

Πολλά σύγχρονα συστήματα αποτελούν πλατφόρμες εργασίας ικανές να συντροφεύσουν τον μαθητή πολλά χρόνια της μαθητικής ζωής του, εάν είναι σχεδιασμένα σαν συστήματα ανοικτού ορίζοντα, πέραν από τους περιορισμούς ενός σχολικού βιβλίου. Τα συστήματα ανοικτού ορίζοντα επιτρέπουν στους μαθητές που ενδιαφέρονται να μάθουν να ξεπεράσουν γρήγορα τα όρια που έχουν χαραχθεί για την ηλικία τους επιταχύνοντας έτσι την κατάκτηση της γνώσης. Μία σοβαρή δυνατότητα των συστημάτων πολυμέσων είναι ότι επιτρέπουν εύκολα την ανάμειξη της εκπαίδευσης με το παιχνίδι και διευκολύνουν έτσι την προσέγγιση με τον μαθητή, αφού στις ηλικίες μεταξύ 8 – 16 ετών, το κανάλι του παιχνιδιού παραμένει πάντα ανοικτό. Με τον τρόπο αυτό μπορεί να συνδυάσει κανείς την εκπαίδευση και το παιχνίδι με την επιβράβευση και να οδηγήσει τον μαθητή σε μία παράλληλη δεύτερη εκπαίδευση ευαισθητοποιώντας τον π.χ. σε πολιτιστικά θέματα, αν επιλεγούν τα «ηλεκτρονικά βραβεία» από τον πολιτιστικό χώρο. Τα εκπαιδευτικά συστήματα πολυμέσων, σχεδιασμένα κατάλληλα, μπορούν να αποτελέσουν ένα πραγματικό βοηθό του παιδιού στο διάβασμα και έτσι να κερδίσουν τους μαθητές.

Τα πολυμέσα δεν είναι υποκατάστατα του δασκάλου, αλλά βοηθοί του δασκάλου. Ο ρόλος του διδάσκοντα μπορεί να αναβαθμιστεί χάρη στην τεχνολογία αυτή και να γίνει περισσότερο επιτελικός. Δεν χρειάζεται

να ασχολείται με την λεπτομέρεια, αφού αυτήν την αναλύει κατάλληλα το σύστημα και την παρουσιάζει στα παιδιά. Ο δάσκαλος σχεδιάζει καινούργιες ασκήσεις, εμπλουτίζει το σύστημα, επιβλέπει τα παιδιά, επισημαίνει τις αδυναμίες και αξιολογεί. Το εκπαιδευτικό λογισμικό δεν είναι απαραίτητο για όλα τα μαθήματα. Είναι χρήσιμο εκεί που υπάρχει μεγάλη αναλυτικότητα και η επιστράτευση των αισθήσεων θα βοηθήσει στην καλύτερη κατανόηση π.χ. γλώσσα, γεωγραφία, φυσική, χημεία, μαθηματικά. Στα μαθήματα όμως όπου κυριαρχεί ο λόγος και χρειάζεται να αναπτυχθεί η φαντασία, μπορεί να είναι χρήσιμο κυρίως ως προς την λειτουργία της ενεργοποίησης του μαθητή.

Κρίσιμο σημείο στις δυνατότητες των πολυμέσων είναι και η εξομίωση. Μπορεί κανείς με μηδενική δαπάνη να φέρει το εργαστήριο στο σπίτι. Με τις σύγχρονες τεχνολογίες πολυμέσων, τις τρεις διαστάσεις και την εικονική πραγματικότητα, το εικονικό εργαστήριο γίνεται αποτελεσματικότερο του πραγματικού γιατί οδηγεί σε περισσότερο πολύπλοκο πειραματισμό και σε μεγάλη επαναληψιμότητα των πειραμάτων, που δεν είναι εύκολο να υλοποιηθούν σε πραγματικό εργαστήριο.

Στην συνέχεια δίνεται ένα τέτοιο παράδειγμα πειραματισμού που θα μπορούσε δύσκολα να εκτελεστεί σε ένα πολύπλοκο εργαστήριο «ελεύθερης πτώσης». Γίνεται επεξήγηση της απόκτησης οριακής ταχύτητας σε σώμα εν πτώσει. Η εξομίωση στον Η/Υ του φαινομένου πραγματοποιείται με κατάλληλη επιλογή μάζας όγκου αρχικής ταχύτητας, πυκνότητας αέρος. Μαθηματική προσέγγιση του φαινομένου επιτρέπει την εποπτεία των τιμών των μεγεθών σε κάθε πείραμα. Γίνεται υλοποίηση της πτώσης διερευνητικά μέχρι να επιτευχθεί οριακή ταχύτητα. Προτείνεται πειραματισμός με ποικιλία γεωμετριών. Γίνεται οπτικοποίηση των δυνάμεων που δρουν στο σώμα κατά την πτώση και επεξήγηση πρακτικών εφαρμογών με οριακή ταχύτητα, καθώς και περιπτώσεων όπου η φύση υιοθετεί την οριακή ταχύτητα κατά την πτώση (αλεξίπτωτο, χιονόπτωση).

Η τεχνολογία των πολυμέσων δίνει ευκαιρίες σχεδιασμού της εκπαιδευτικής διαδικασίας σε νέες βάσεις. Με δεδομένη και την ύπαρξη του παγκόσμιου ιστού μπορεί να σχεδιάσει κανείς την εκπαίδευση ώστε να είναι:

1. Πιο κοντά στις εφαρμογές μέσα από την πρόσβαση σε πραγματικό υλικό.

2. Περισσότερο διαθεματική, αφού το μείγμα των μέσων δημιουργεί άνοιγμα προς αυτήν την κατεύθυνση.
3. Περισσότερο διερευνητική, αφού οι πληροφορίες υπάρχουν και πρέπει να τις αναζητήσει κανείς και να τις οργανώσει.

Με βάση αυτές τις αρχές πρέπει να περιμένει κανείς από αυτήν την τεχνολογία να βοηθήσει τους νέους να αποκτήσουν τις ιδιότητες που απαιτεί η σύγχρονη ζωή, όπως:

- Αυξημένη μεθοδικότητα και συγκρότηση
- Αυξημένη τεχνογνωσία και καλύτερη εποπτεία του σύγχρονου κόσμου
- Αυξημένη κοινωνικότητα
- Αυξημένη ευαισθησία στις πολιτισμικές αξίες
- Ικανότητα για καλύτερες επιλογές της πληροφορίας με κριτήρια την αντικειμενικότητα και την διαχρονική αξία.

5. Αποθήκες Γνώσεων

Η τάση σήμερα, πέρα από τα εκπαιδευτικά σενάρια, τα μαθήματα που είναι αρχικά περιορισμένα, τις ασκήσεις που και αυτές είναι περιορισμένες, είναι στην χρήση αυτού που λέμε ανοικτά συστήματα, δηλαδή συστήματα τα οποία εύκολα θα μπορέσει κανείς να συνεχίζει να εμπλουτίζει με ασκήσεις και θεωρία που θα πηγάζουν στα εκπαιδευτήρια, εκεί όπου διεξάγεται η εκπαιδευτική διαδικασία και που όλοι ξέρουμε ότι από την αλληλοτροφοδότηση μεταξύ διδασκόντων και διδασκομένων δημιουργούνται νέες προσεγγίσεις στα συστήματα, τις δυνατότητες επέκτασης από τον δάσκαλο και ίσως τον ίδιο τον μαθητή. Αυτή η διαδικασία είναι ενδιαφέρουσα, και δίνει κάποια αποτελέσματα, όμως έχει και κάποια όρια. Τα πιθανά όρια σχετίζονται με την τεχνολογική πολυπλοκότητα. Έτσι όσο πιο απλό είναι στην δομή του ένα εκπαιδευτικό λογισμικό τόσο πιο εύκολα επεκτείνεται, όσο πιο πολύπλοκο είναι στην δομή του και τεχνολογικά εξελιγμένο τόσο η επέκταση από τους χρήστες είναι δυσκολότερη. Έτσι δεν είναι πάντα εύκολο να μετατρέψει κανείς ένα εκπαιδευτικό λογισμικό σε αποθήκη γνώσεων χωρίς τον κατασκευαστή.

Σε κάποιους τομείς όμως όπως η γλωσσική εκπαίδευση, λόγω της ανάπτυξης της γλωσσικής τεχνολογίας, είναι δυνατό να δημιουργήσει κανείς πραγματικές «αποθήκες γνώσεων» όπου θα μπορεί να ανατρέχει

έναν μαθητή. Έτσι π.χ. αν διαθέτει ένα μορφολογικό λεξικό θα μπορεί να βοηθήσει τον μαθητή για οποιαδήποτε λέξη συναντάει να μπορεί να ξέρει τι μέρος του λόγου είναι και πώς κλίνεται σε όλη την κλητική ποικιλία που μπορεί να έχει σε μία γλώσσα σαν την ελληνική με την γνωστή πολυτυπία της. Με ειδικά εργαλεία επίσης, για κάποιον που αρχίζει να μαθαίνει μία ξένη γλώσσα, μπορεί να τον βοηθήσει ένα σύστημα να μεταγράψει οποιαδήποτε λέξη με χρήση του φωνητικού αλφαβήτου, ώστε να μπορέσει να την κατανοήσει γρήγορα από πλευράς φωνητικού περιεχομένου. Η απόδοση στην συνέχεια μπορεί να γίνει με συνθετική ομιλία ώστε ακούγοντάς την πολλές φορές να είναι σε θέση ο εκπαιδευόμενος να την προφέρει. Στην συνέχεια, πάλι χάρη σε ειδικά εργαλεία να μπορεί το σύστημα να αξιολογήσει με κατάλληλο “score” την προφορά του ομιλητή κ.ο.κ. Έτσι τα συστήματα πολυμέσων εφοδιάζονται με μία λειτουργικότητα που ξεπερνάει οποιοδήποτε συμβατικό σύστημα.

6. Λειτουργικότητα και Αισθητική

Η λειτουργικότητα και η αισθητική δεν δρουν ανταγωνιστικά στα συστήματα πολυμέσων. Η αυξημένη αισθητική όμως δρα ως «μάσκα» κρύβοντας πολλές φορές λειτουργικές ατέλειες. Ποια όμως είναι τα στοιχεία που επηρεάζουν την λειτουργικότητα και ποια είναι εκείνα που επηρεάζουν την αισθητική;

Η αισθητική έχει να κάνει με την ποιότητα των γραμματοσειρών, την ποιότητα των εικόνων και του video, ενώ στην περίπτωση κινουμένων σχεδίων έχει μεγάλη σημασία και ο χρονισμός των κινήσεων. Το ίδιο συμβαίνει στις προσεγγίσεις εικονικής πραγματικότητας όπου η ταχύτητα αντίδρασης του περιβάλλοντος είναι σημαντικός παράγων.

Επίσης η διαρρύθμιση των οθονών έχει να κάνει με την αισθητική, αλλά και με την λειτουργικότητα σε ένα βαθμό.

Η λειτουργικότητα συνδέεται περισσότερο με τη σωστή ιεράρχηση των πληροφοριών, την ευκολία στη πρόσβαση και την ανάκτηση της πληροφορίας, τις ευκολίες στην πλοήγηση και την επιστροφή. Η συστηματοποίηση είναι σημαντικό βήμα για την αυξημένη λειτουργικότητα. Σημαντική είναι επίσης και η χαρτογράφηση των πληροφοριών που υπάρχουν στο σύστημα και η δυνατότητα απεικόνισης με πίνακες πε-

ριεχομένων συμβατικούς και μη. Είναι χρήσιμο για αυξημένη λειτουργικότητα να υπάρχει συνεχής ανατροφοδότηση συστήματος – χρήστη, ώστε ο χρήστης κάθε στιγμή να «αισθάνεται το σύστημα».

Όλες αυτές οι αξιώσεις δημιουργούν δυσκολίες στην ανάπτυξη και υψηλά κόστη. Πρέπει τότε κάποια ομάδα που θα αναπτύξει ένα καινούργιο σύστημα να αποφασίσει αν θα γείρει προς την πλευρά της αυξημένης λειτουργικότητας ή της αυξημένης αισθητικής γιατί συνήθως το κόστος είναι απαγορευτικό για να έχει και τα δύο.

7. Αξιολόγηση Εκπαιδευτικών Τίτλων

Η τελική αξιολόγηση γίνεται από τους χρήστες. Η καλύτερη μεθοδολογία είναι να δημιουργείται ένα εκπαιδευτικό λογισμικό πιλοτικά και μετά την χρήση να ανασχεδιάζεται ώστε να ληφθούν υπ' όψη οι απαιτήσεις των χρηστών. Πολλές φορές αυτός ο ανασχεδιασμός κοστίζει ακριβά.

Στην συνέχεια σημειώνονται επιγραμματικά τα κριτήρια τα οποία χρησιμοποιήθηκαν κατά την κρίση των τίτλων πολυμέσων από την κριτική επιτροπή του Ευρωπαϊκού Διαγωνισμού που χορηγεί τα βραβεία EUROPRIX.

1. Ποιότητα, κατανοητότητα και έκταση του περιεχόμενου
2. Ευκολία στην χρήση: λειτουργικότητα, πλοήγηση, προσανατολισμός
3. Προστιθέμενη αξία χάρη στην διαλογικότητα
4. Ελκυστικότητα: αισθητική αξία, σχεδιασμός οθονών
5. Τεχνολογία
6. Ευρωπαϊκή Διάσταση

Το Παιδαγωγικό Ινστιτούτο αξιολογεί επίσης το εκπαιδευτικό λογισμικό για τα σχολεία με ιδιαίτερη προσοχή. Έχει τα εξής γενικά κριτήρια που βαθμολογούνται με λεπτομέρεια:

1. Περιεχόμενο (15 υπό-κριτήρια)
2. Διδακτική και Παιδαγωγική Μεθοδολογία (9 υπό-κριτήρια)
3. Τεχνικά Χαρακτηριστικά – Λειτουργία (7 υπό-κριτήρια)
4. Περιβάλλον Διεπαφής / Αλληλεπίδραση (38 υπό-κριτήρια)

Η αξιολόγηση είναι μία λεπτή διαδικασία είτε γίνεται

εν όψει της πιστοποίησης (Π.Ι.), είτε εν όψει της βράβευσης (EUROPRIX). Φυσικά η αξιολόγηση εν όψει πιστοποίησης είναι περισσότερο λεπτομερής. Αξίζει να συγκρίνει κανείς τις δύο προσεγγίσεις του Europrix και του ΠΙ. Αμφότερες οι λειτουργίες, τόσο η βράβευση όσο και η πιστοποίηση είναι εξαιρετικά χρήσιμες και διαμορφώνουν την αγορά του εκπαιδευτικού λογισμικού σε ποιοτικότερη κατεύθυνση.

Τώρα το εκπαιδευτικό λογισμικό κινείται ώστε να υπάρξουν πλατφόρμες που θα προσεγγίσουν σε ποιότητα και ταχύτητα αυτές των ηλεκτρονικών παιχνιδιών. Ίσως είναι η επόμενη πρόκληση.

Euromap Language Technologies at European Language Technologies: Actions and Prospects

Στην ημερίδα με θέμα "Γλωσσική Εκπαίδευση στην Κοινωνία της Πληροφορίας & Ευρωπαϊκή Τεχνολογία Ανθρωπίνου Λόγου: Δράσεις & Προοπτικές" που πραγματοποιήθηκε στο ΙΕΛ στις 24 Οκτωβρίου 2001 (βλ. παρόν τεύχος, σελίδα 3) συμμετείχαν ως προσκεκλημένοι, μεταξύ άλλων, η Καθ. κα Bente Maegaard, διευθύντρια του Ινστιτούτου Γλωσσικής Τεχνολογίας (CST) με έδρα την Κοπεγχάγη της Δανίας και συντονίστρια του ευρωπαϊκού έργου EUROMAP Language Technologies, και ο Καθ. κ. Jean-Paul Haton, διευθυντής του τμήματος Αναγνώρισης Προτύπων και Τεχνητής Νοημοσύνης του Ινστιτούτου LORIA/INRIA στη Νανσύ της Γαλλίας. Οι παρουσιάσεις τους κατά τη διάρκεια της απογευματινής εκδήλωσης, ακολουθούν:



EUROMAP Language Technologies
at
European Language Technologies: Actions and Prospects, Athens October 24, 2001

Bente Maegaard
EUROMAP co-ordinator
Centre for Language Technology, Denmark
bente@cst.dk



EUROMAP Language Technologies

What is it?

- EU project
- 11 National members
- Knowledge building and Awareness
- Language Technology
 - EU support
 - Its path to market



11 partners in 11 countries

AT	GR
DE	IT
DK	NL/B
ES	UK
FI	BG
FR	and contacts to other countries

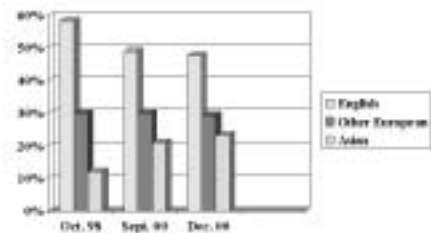


Examples of language technology

- Speech recognition, understanding
- Dialogues
- Machine translation
- Documentation
- Lively and multilingual web sites
- Educational software
- etc.



Less than 50% is in English



How will it be?

- Tourists - small tools for interpretation
- Greek web pages will be translated by a click - to a number of languages
- Automatic abstracting for Greek
- Household machines
- Toys - Greek speaking dolls



Human Language Technologies



IST programme - possibilities for support

- Support for projects (R&D, demonstration projects etc.)
- Multilinguality
- Natural Interactivity
- Cross-lingual information management



eContent



Connected to eEurope - possibilities for support

- European content industries
- Publishers, web-publishers, TV, telecom, news agencies etc



EUROMAP Language Technologies



At three interacting levels

- National level
- Web level
- International level



EUROMAP at the National level



See next presentation

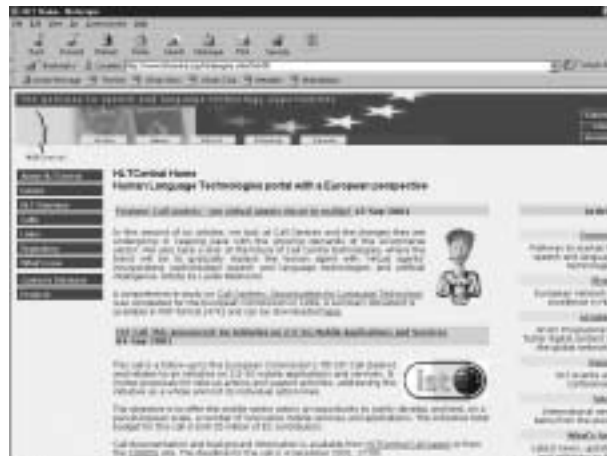


EUROMAP at the Web level



www.HLTCentral.org

- The HLT website
- News
- Events
- Articles
- Etc.
- The EUROMAP site
- www.HLTCentral.org/euromap



EUROMAP Newsletter



EUROMAP monthly newsletter

- Gives a fast overview of the news of the month
- News
- Events
- New content on the web site
- Subscribe at
- www.HLTCentral.org/euromap



Knowledge base



- For each country
 - Producers, products
 - Researchers, institutions
 - Internet, PC penetration
 - Telecom
 - Users, user organisations
 - Path to market, technology transfer



Central and Eastern Europe and other countries



Special focus on CEEC, Malta, Cyprus etc.

- Important to
 - Create a knowledge base covering these countries
 - Spread the word, promote language technology
 - Starting a network



All knowledge bases will be available on the websites



National, CEEC and international ones

- Gradual updates
- For the time being Success stories
- Then products, producers etc.
- CEEC and other countries
- Towards the end (2003): Analysis of the data



International event



Berlin, 26-27 September 2002

- European Language Solutions 2002
 - Conference
 - Exhibition
 - Venture capital, SMEs



Path to market - White paper



Final report

- The final report will build on the knowledge base
 - Analysis of the data made available
 - By market analyst



Our vision



- Europe will take advantage of its rich research base
- Europe will build on its multi-linguality
- Europe will understand the path to market and use it
- Here Europe also covers the EU accession countries



Our vision



- I hope you will be part of it!



Language Technology: Present and Future of Automatic Speech Recognition

Spoken language interfaces to computers is a topic that has fascinated engineers and speech scientists for over five decades. For many, the ability to converse freely with a machine represents the ultimate challenge to our understanding of the production and perception processes involved in human speech communication. In addition to being an interesting topic, spoken language interfaces are rapidly becoming a necessity. In the near future, interactive networks will provide easy access to a wealth of information and services that will fundamentally affect how people work, play and conduct their daily affairs. Advances in human language technology are needed for the average citizen to communicate with networks using natural communication skills and everyday devices, such as telephones and televisions.

Language Technology : Present and future of automatic speech recognition

Jean-Paul Haton
LORIA-INRIA
Université Henri Poincaré, Nancy 1
Institut Universitaire de France
jph@loria.fr

ILSP Oct. 2001 Athens

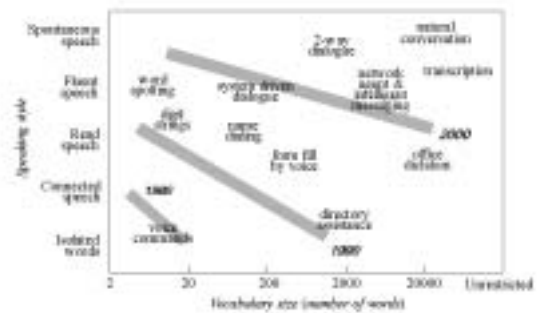
Outline

- Evolution of ASR
- Basic principles
- A major problem: robustness
- State of the art
- Future trends and perspectives

Applications of ASR

- Data entry: voice dictation, etc.
- Access to information: travel, banking, portals
- Commands: avionics, automobile, etc.
- Speech transcription
- Handicapped
- Speech-to-speech translation

ILSP Athens, Oct. 2001 Jean-Paul Haton 3



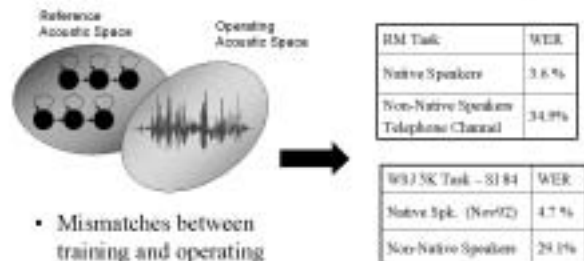
Speech recognition technology
ILSP Athens, Oct. 2001 Jean-Paul Haton 4

Speech modeled by HMMs

- The speech production process is supposed to be Markovian

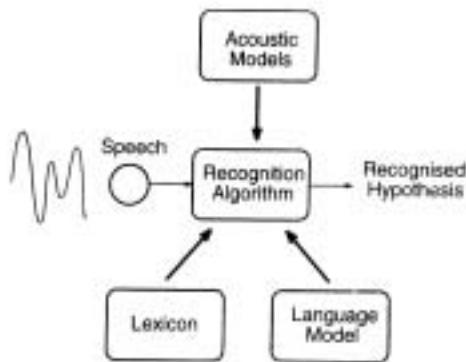


Robustness issues in ASR

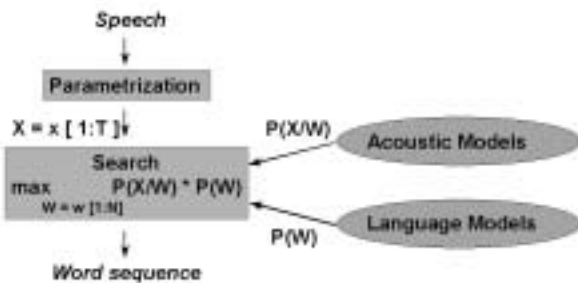


- Mismatches between training and operating conditions can lead to a significant loss in performance

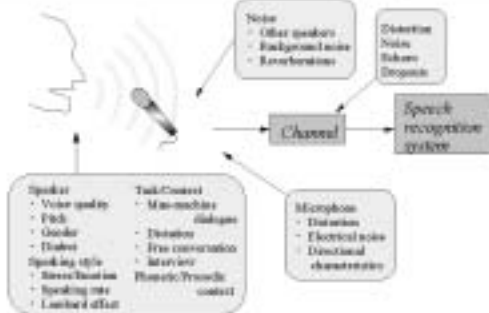
Principle of Automatic Speech Recognition



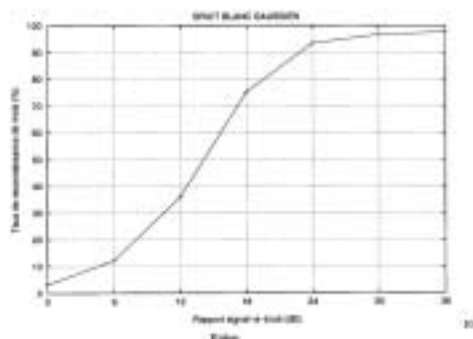
Bayes Decision Rule



Main causes of acoustic variation in speech



Influence of noise on recognition



Approaches to robustness

- Three non-exclusive approaches
 1. Signal acquisition and denoising
 2. Robust parameterization (less sensitive to noise variations)
 3. Model modification : adaptation

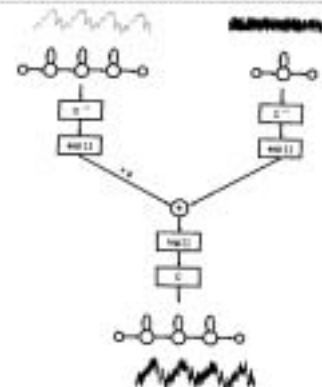
Speech Parametrization

- Use of knowledge from perception, psychoacoustics, neurology, etc. :
 - ear models
 - noise masking
 - dynamic parameters
 - analyses PLP, RASTA, RASTA-PLP, etc.
- Cepstral Normalization

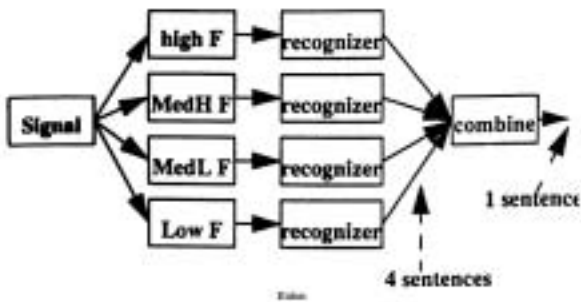
Examples of adaptation methods

- Combination of models: PMC (Parallel Model Combination)
- Linear transformations of models. Example : MLLR (Maximum Likelihood Linear Regression) : $\hat{\mu} = A\mu + b$
- MAP (Maximum a posteriori): the *a priori* probability $P(\lambda)$ of models is not uniform
- Hybridations MLLR/MAP
- Model grouping into classes (eigenvoices, etc.)
- Non-linear transforms (neural networks, etc.)

Parallel Model Combination (PMC): Gales, 1995

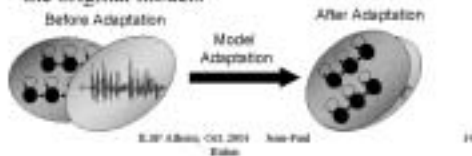


Multibands recognition

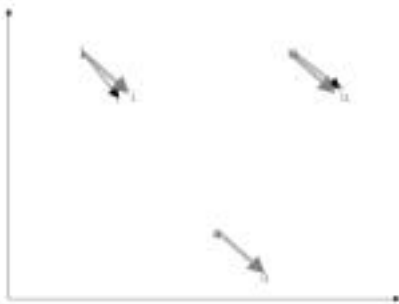


Acoustic Model Adaptation

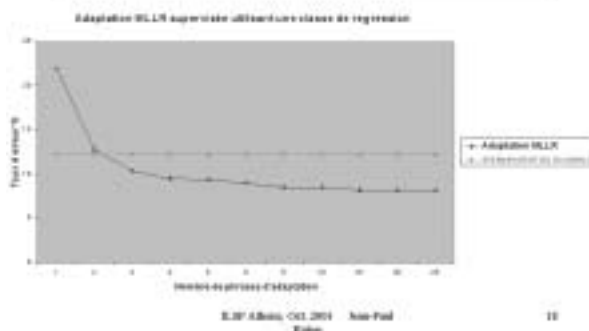
- Objective
 - To derive a new set of models that better match the operating environment
- Principle
 - To use a *small* amount of *adaptation* data reflecting the operating environment to modify the original models



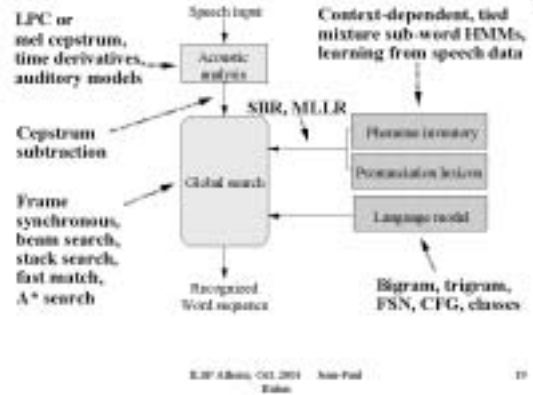
MLLR: example



MLLR : speaker adaptation



State-of-the-art algorithms in speech recognition



Future research needs

- Better modelling of speech variability (structural and non-structural)
- Better knowledge on speech perception and adaptation by humans
- Beyond recognition: speech understanding

Προσδιορισμός της σχέσης αριθμού λέξεων που περιέχει ένα κείμενο και του εύρους αντιπροσώπησης των γλωσσικών μονάδων του.

Γεώργιος Κ. Μικρός
Πανεπιστήμιο Αθηνών & Ινστιτούτο Επεξεργασίας
του Λόγου

1. Εισαγωγή

1.1 Ποσοτικός σχεδιασμός στα ΗΣΚ

Τα Ηλεκτρονικά Σώματα Κειμένων (ΗΣΚ) αποτελούν αναμφισβήτητα μια από τις σημαντικότερες εξελίξεις στον χώρο της γλωσσικής τεχνολογίας τα τελευταία χρόνια. Η ανάπτυξη ιεραρχικά δομημένων συλλογών ηλεκτρονικών κειμένων ήρθε σύντομα αντιμέτωπη με ορισμένα βασικά ερωτήματα δειγματοληψίας. Δεδομένου ότι ένας σημαντικός αριθμός ΗΣΚ αναπτύχθηκαν με στόχο να αποτελέσουν αντιπροσωπευτικά δείγματα της συγχρονικής χρήσης κάποιας γλώσσας σύντομα τέθηκε το θέμα της αντιπροσωπευτικότητας στο προσκήνιο. Η ποσόστωση διαφορετικών κειμενικών κατηγοριών αποτελεί ακόμα και σήμερα ένα σημαντικό θέμα συζήτησης με πολλές και διαφορετικές μεταξύ τους προσεγγίσεις. Συναφές θέμα με αυτό της αντιπροσωπευτικότητας των κειμενικών κατηγοριών είναι η ποσοτική σύσταση του ΗΣΚ. Δεδομένου ότι έχουμε καταλήξει στην κατάλληλη ποσόστωση των κειμενικών κατηγοριών με βάση κάποιο από τα πολλά προτεινόμενα κριτήρια στη βιβλιογραφία, ερχόμαστε αντιμέτωποι με το μέγεθος και τον αριθμό των κειμένων που θα συμπεριλάβουμε κάτω από κάθε κειμενική κατηγορία.

Η σημαντικότερη επίδραση στον ποσοτικό σχεδιασμό ενός ΗΣΚ προέρχεται από τη δομή και τις επιλογές που έγιναν στην κατάρτιση του Brown Corpus (Kucera & Frances 1967). Τα βασικότερα σημεία της μπορούν να συνοψιστούν παρακάτω:

- α) 1 εκ. λέξεις
- β) Περίπου ισόποση κατανομή μεταξύ των κειμενικών ειδών
- γ) 500 κειμενικά δείγματα
- δ) 2.000 λέξεις για κάθε κειμενικό δείγμα

- ε) γραπτές πηγές που έχουν εκδοθεί

Αν και το συγκεκριμένο ΗΣΚ αποτελεί σημείο αναφοράς ακόμα και σήμερα πολλές από τις επιλογές σχεδιασμού του αντιπροσωπεύουν συγκεκριμένες ανάγκες / περιορισμούς της δεκαετίας που αναπτύχθηκε (Eagles 1996). Ειδικότερα:

- α) Το 1 εκ. λέξεις ήταν τεράστια ποσότητα ηλεκτρονικού κειμένου για τη δεκαετία του 1960. Ωστόσο, η τεχνολογική εξέλιξη κατέστησε το συγκεκριμένο αριθμό εύκολα επιτεύξιμο μετέπειτα. Ήδη από τα μέσα της δεκαετίας του 1970 ο αριθμός λέξεων είχε ανέβει σε κάποια ΗΣΚ στα 10 εκ. λέξεων φτάνοντας το 1985 στον αριθμό 20 εκ. λέξεων (The Birmingham Collection of English Texts). Ήδη από τα μέσα του 1990 κάποια ΗΣΚ αγγίζουν τα 200 εκ. λέξεις (The Bank of English).
- β) Με την ίση κατανομή ανάμεσα σε κειμενικά είδη προσεγγίστηκε το θέμα αντιπροσωπευτικότητας.
- γ) Θα πρέπει να υπάρχει μια ενδιάμεση κατηγοριοποίηση μεταξύ του συνολικού ΗΣΚ και των μεμονωμένων κειμένων. Το ρόλο αυτό τον παίζει η κατηγοριοποίηση των κειμενικών ειδών.
- δ) Το θέμα του ίσου μεγέθους ανά κειμενικό δείγμα είναι υπό αμφισβήτηση και δεν θεωρείται υποχρεωτικό.

Μελέτες που έχουν γίνει για την αγγλική γλώσσα (Haan 1992) δείχνουν ότι τα συχνά γλωσσικά χαρ/κά παρουσιάζουν σχετική σταθερότητα εμφάνισης ακόμα και σε μικρά δείγματα, ενώ τα πιο σπάνια παρουσιάζουν μεγάλη ποικιλία και χρειάζονται μεγαλύτερα δείγματα για να αντιπροσωπευθούν σε ένα ΗΣΚ. Μια μέση λύση είναι ο αριθμός των 1000 λέξεων ανά κείμενο, αφού αυτό το μέγεθος επιδεικνύει καλή αντιπροσωπευτικότητα σε μια ποικιλία χαρ/κών (Biber & Finegan 1991).

1.2 Η έννοια του λειτουργικού εύρους ενός ΗΣΚ

Οι προτάσεις στη διεθνή βιβλιογραφία σχετικά με τις ποσοτικές παραμέτρους ενός ΗΣΚ έχουν βασιστεί σε δεδομένα της Αγγλικής γλώσσας. Είναι απαραίτητο τέτοιες μετρήσεις να επαναληφθούν σε ΗΣΚ της Ελληνικής για να επιβεβαιώσουν ή να διαφοροποιήσουν τα συμπεράσματα για την ποσοτική σύσταση ενός ΗΣΚ.

Στο παρόν άρθρο θα ασχοληθούμε με τον καθορισμό του ελάχιστου μεγέθους αριθμού λέξεων που θα πρέπει να περιέχει ένα κειμενικό δείγμα για να συμπεριληφθεί σε ένα ΗΣΚ. Ειδικότερα θα προσπαθήσουμε να αναλύσουμε τη σχέση που έχει το μέγεθος των κειμενικών αρχείων με την έννοια του λειτουργικού εύρους ενός ΗΣΚ. Με τον όρο λειτουργικό εύρος προσδιορίζουμε τον αριθμό των αρχείων ενός ΗΣΚ τα οποία παρέχουν πραγματώσεις σε μια ποικιλία γλωσσικών χαρακτηριστικών που αναζητούνται σε αυτό το ΗΣΚ. Η έννοια αυτή σχετίζεται με την αξιοποίηση των κειμένων που συμπεριλαμβάνονται σε ένα ΗΣΚ. Ένα ΗΣΚ που χαρακτηρίζεται από μεγάλο λειτουργικό εύρος περιέχει κείμενα στα οποία εμφανίζονται γλωσσικά χαρακτηριστικά όλων των συχνοτήτων, από τα πολύ συχνά έως τα πολύ σπάνια. Φυσικά, το λειτουργικό εύρος καθορίζεται κυρίως από το πόσο αξιόπιστα απεικονίζεται η πραγμάτωση σπάνιων γλωσσικών μονάδων. Η δυνατότητα αυτή επιτρέπει την οικονομικότερη αξιοποίησή των γλωσσικών πόρων που συμπεριλαμβάνουμε σε ένα ΗΣΚ. Έτσι ένα μικρό σε μέγεθος ΗΣΚ με μεγάλο λειτουργικό εύρος μπορεί να καλύψει τη χρήση γλωσσικών στοιχείων που ένα ΗΣΚ χαμηλού λειτουργικού εύρους για να καλύψει θα χρειαζόταν πολλαπλάσιο μέγεθος. Κάτω από το συγκεκριμένο πλαίσιο το λειτουργικό εύρος ενός ΗΣΚ σχετίζεται άμεσα με το εύρος των κειμενικών ειδών και γενών που συμπεριλαμβάνονται. Όσο πιο ευρεία αντιπροσώπευση πετύχουμε, τόσο μεγαλύτερες πιθανότητες θα έχουμε να συμπεριλάβουμε γλωσσικές μονάδες χαμηλής συχνότητας.

2. Καθορισμός ελάχιστου αριθμού λέξεων ανά κειμενικό δείγμα

2.1 Δείκτες προσδιορισμού της επίδρασης του κειμενικού μεγέθους στην αναζήτηση γλωσσικών χαρακτηριστικών σε ένα ΗΣΚ

Η μεθοδολογία για να αξιολογηθεί η αξιοποίηση ενός ΗΣΚ στηρίζεται στην εξέταση συγκεκριμένων γλωσσικών χαρ/κών σε ομάδες κειμενικών δειγμάτων το μέγεθος των οποίων αυξάνει διαδοχικά. Θα χρησιμοποιήσουμε επομένως την έννοια του λειτουργικού εύρους όχι ως κριτήριο σύγκρισης δύο διαφορετικών ΗΣΚ, άλλα ως μέτρο αξιολόγησης της επίπτωσης που

έχει η αύξηση του κειμενικού μεγέθους στην εμφάνιση συγκεκριμένων γλωσσικών μονάδων. Τα γλωσσικά χαρ/κά που εξετάζονται ανήκουν σε μια ποικιλία γλωσσικών επιπέδων (φωνολογικό, μορφολογικό, συντακτικό, λεξιλογικό) και χωρίζονται σε δύο βασικές κατηγορίες:

- α) Γλωσσικά χαρ/κά υψηλής συχνότητας
- β) Γλωσσικά χαρ/κά μεσαίας συχνότητας
- γ) Γλωσσικά χαρ/κά χαμηλής συχνότητας

Η διάκριση αυτή κρίνεται απαραίτητη καθώς τα γλωσσικά χαρ/κά υψηλής συχνότητας παρουσιάζουν σχετική σταθερότητα εμφάνισης ανεξαρτήτως μεγέθους κειμένου, ενώ τα χαρ/κά χαμηλής συχνότητας παρουσιάζουν υψηλές διακυμάνσεις στις εμφανίσεις τους.

Τα συμπεράσματα στα οποία οδηγείται κανείς από την εξέταση των παραμέτρων της κατανομής των υπό εξέταση γλωσσικών χαρ/κών δεν είναι τελεσίδικα, αφού η έννοια του ελάχιστου μεγέθους κειμένου δεν μπορεί να οριστεί στατιστικά. Έγινε προσπάθεια να βελτιστοποιηθεί η σχέση δύο αντίρροπων απαιτήσεων που παρουσιάζει ο ποσοτικός σχεδιασμός ενός ΗΣΚ:

- α) Κείμενα πολλών λέξεων για να καλύπτουν την εμφάνιση των περισσότερων γλωσσικών χαρ/κών.
- β) Κείμενα λίγων λέξεων γιατί είναι πολύ περισσότερα και συλλέγονται πιο εύκολα.

2.2 ΗΣΚ ελέγχου

Για να ελεγχθεί η επιρροή του μεγέθους των κειμενικών δειγμάτων στην κατανομή των γλωσσικών χαρ/κών επιλέχθηκε η ανάλυση κειμένων πολλαπλού μεγέθους και η εξέταση της κατανομής που παρουσιάζει η συχνότητα λεξικών και γραμματικών δεδομένων υψηλής και χαμηλής συχνότητας εμφάνισης. Το ΗΣΚ που χρησιμοποιήθηκε αποτελείται από ένα υπόσυνολο του ΕΘΕΓ. Ειδικότερα χρησιμοποιήθηκε ένα ΗΣΚ από τις εφημερίδες «ΤΟ ΒΗΜΑ» και «Ελευθεροτυπία» το οποίο αριθμεί συνολικά 651.900 λέξεις και περιλαμβάνει 485 κειμενικά δείγματα τα οποία καλύπτουν ένα μεγάλο φάσμα κειμενικών γενών και θεμάτων. Η κατανομή του ΗΣΚ που χρησιμοποιήθηκε δίνεται στους παρακάτω πίνακες:

Πίνακας 1

Κατανομή των λέξεων του ΗΣΚ ανά εκδότη και κειμενικό είδος

PAROLE-GENRE	ΕΛΕΥΘΕΡΟΤΥΠΙΑ	ΤΟ ΒΗΜΑ	Γενικό Σύνολο
ADV	928		928
DIS	63387	98747	162134
FEA	66630	139624	206254
INF	162561	102912	265473
INS		11528	11528
NON	3274		3274
OFF	1211		1211
PRI	1098		1098
<i>Γενικό σύνολο</i>	299089	352811	651900

Πίνακας 2

Κατανομή των λέξεων του ΗΣΚ ανά εκδότη και κειμενικό θέμα

PAROLE-TOPIC	ΕΛΕΥΘΕΡΟΤΥΠΙΑ	ΤΟ ΒΗΜΑ	Γενικό Σύνολο
BUS	50482	52913	103395
GEO	8172	21889	30061
HEA	48494	50213	98707
HIS	9098	41106	50204
HUM	50203	50072	100275
LEI	49151	50627	99778
SCI	32723	37518	70241
SOC	50766	48473	99239
<i>Γενικό Σύνολο</i>	299089	352811	651900

2.3 Γλωσσικά χαρακτηριστικά

Για λόγους ευελιξίας στη διαχείριση των μετρήσεων δημιουργήθηκαν δύο κατηγοριοποιήσεις. Η πρώτη είναι αναλυτική και κάθε κατηγορία μεγέθους καλύπτει από 100 έως 200 λέξεις. Σχηματίστηκαν 15 κατηγορίες μεγέθους. Κάθε κατηγορία διαμορφώθηκε έχοντας το λιγότερα 25 κείμενα για να ικανοποιηθεί το θεώρημα του κεντρικού ορίου (ενδ. Lehman 1991: 237). Η δεύτερη κατηγοριοποίηση είναι πιο συγκεκριμένη και χωρίζει τα υπάρχοντα αρχεία σε 6 μεγάλες κατηγορίες μεγέθους. Η κάθε κατηγορία από αυτές καλύπτει 500 λέξεις. Και οι δύο κατηγοριοποιήσεις σε σχέση με τον αριθμό των κειμένων που κάθε μια περιλαμβάνει δίνονται στον παρακάτω πίνακα:

Πίνακας 3

Κατηγοριοποιητικό σχήμα του ΗΣΚ βάσει του μεγέθους των κειμένων

Αναλυτική Κατηγοριοποίηση			Συγκεντρωτική Κατηγοριοποίηση		
Μέγεθος (λέξεις)	Αριθμός κειμένων	Κατηγορία Μεγέθους	Μέγεθος (λέξεις)	Αριθμός κειμένων	Κατηγορία Μεγέθους
100-300	24	1	100-500	86	1
300-400	39	2	500-1000	95	2
400-550	33	3	1000-1500	154	3

550-750	30	4	1500-2000	67	4
750-900	30	5	2000-3000	58	5
900-1000	25	6	3000<	25	6
1000-1100	45	7			
1100-1200	28	8			
1200-1300	48	9			
1300-1500	33	10			
1500-1750	36	11			
1750-2000	31	12			
2000-2300	31	13			
2300-3000	27	14			
3000<	25	15			

Η επιλογή των γλωσσικών χαρ/κών των οποίων η κατανομή εξετάστηκε έγινε με γνώμονα την αντιπροσώπευση λεξικών δειγμάτων που ανήκουν τόσο στις λέξεις με υψηλή εμφάνιση, όσο και στις λέξεις που θεωρούνται χαμηλής συχνότητας. Για να επιλεγούν τα παραπάνω λεξικά στοιχεία δημιουργήθηκε μια λίστα συχνότητας του λεξιλογίου που περιείχε το ΗΣΚ που εξετάσαμε. Στη συνέχεια η λίστα χωρίστηκε σε 12 περιοχές συχνότητας. Για κάθε μια από αυτές επιλέχθηκαν με τυχαίο τρόπο 2 λέξεις με τον μοναδικό περιορισμό να μην ανήκουν στο ίδιο μέρος του λόγου.

Πίνακας 4

Λέξεις που αναζητήθηκαν και η συχνότητα εμφάνισής τους

A/A	Περιοχή Συχνότητας	Σειρά εμφάνισης	Λέξη	Συχνότητα στο ΗΣΚ	% επί του συνόλου των λέξεων του ΗΣΚ
1	1	4	του	14684	2,252
2		5	να	13787	2,115
3	2	13	για	7757	1,190
4		14	είναι	7371	1,131
5	3	29	τις	3413	0,524
6		32	ένα	2537	0,389
7	4	42	ήταν	1856	0,285
8		44	μου	1839	0,282
9	5	64	πρέπει	1054	0,162
10		96	χρόνια	906	0,139
11	6	161	υπάρχουν	472	0,072
12		168	πολιτική	433	0,066
13	7	288	κυβέρνηση	224	0,034
14		297	αποτελεί	213	0,033
15	8	635	δώσει	100	0,015
16		637	θάλασσα	100	0,015
17	9	1293	σημειωθεί	50	0,008
18		1294	σκέψη	50	0,008

Στη συνέχεια επιλέχθηκαν κάποια φωνολογικά και μορφολογικά χαρακτηριστικά τα οποία επιλέχθηκαν με τρόπο ώστε να παρουσιάζουν ομοιόμορφη κατανομή σε σχέση με τη συχνότητα εμφάνισής τους. Τα χαρακτηριστικά δίνονται στον παρακάτω πίνακα:

Πίνακας 5

Φωνολογικά και μορφολογικά χαρ/κά που αναζητήθηκαν και η συχνότητα εμφάνισής τους στο ΗΣΚ

2.3.1 Φωνολογία		
Χαρ/κό	Περιοχή συχνότητας	Συχνότητα
θε	1	13487
sm	2	7069
kl	3	3288
sθ	4	1849
xθ	5	904
lt	6	466
nst	7	196
dl	8	101
sn	9	51

Μορφολογία		
Χαρ/κό	Περιοχή συχνότητας	Συχνότητα
ους	1	11885
ής	2	6672
ός	3	4945
ουμε	4	2861
ετε	5	1071
εως	6	421
μένης	7	187
σ(Ξ)(Ψ)ιμο	8	136
ούσας	9	50

Θα πρέπει να σημειωθεί ότι τα χαρακτηριστικά τα οποία επιλέχθηκαν δεν παρουσιάζουν κάποια ιδιαίτερη εξάρτηση από το κειμενικό θέμα. Αυτό επιτρέπει την εστίαση της ανάλυσης μόνο στην επίδραση του μεγέθους του κειμένου ως ανεξάρτητης ερευνητικής μεταβλητής.

3. Αποτελέσματα

Αναζητήθηκαν τα χαρ/κά και υπολογίστηκε το ποσοστό των κειμένων στα οποία δεν εμφανίστηκε το κάθε χαρ/κό. Η σχέση μείωσης του μεγέθους και απουσίας εμφάνισης ενός χαρακτηριστικού δίνεται στον παρακάτω πίνακα:

Πίνακας 6

Κείμενα που δεν εμφανίστηκε καθόλου το λεξιλογικό χαρ/κό που αναζητήθηκε

Λέξη	Περιοχή Συχνότητας	Κείμενα στα οποία δεν εμφανίστηκε καθόλου το χαρ/κό	Ποσοστό % επί του συνόλου των κειμένων (N=485)
του να	1	0 4	0 0,82
για είναι	2	4 20	0,82 4,12
τις ένα	3	20 78	4,12 16,08
ήταν μου	4	164 315	33,81 64,94
πρέπει χρόνια	5	185 171	38,14 35,27
υπάρχουν πολιτική	6	264 362	54,43 74,63
κυβέρνηση αποτελεί	7	400 338	82,47 69,69
δώσει θάλασσα	8	411 448	84,74 92,37
σημειωθεί σκέψη	9	449 450	92,57 92,78

Πίνακας 7

Κείμενα στα οποία δεν εμφανίστηκε καθόλου το φωνολογικό και μορφολογικό χαρ/κό που αναζητήθηκε

Λέξη	Περιοχή Συχνότητας	Κείμενα στα οποία δεν εμφανίστηκε καθόλου το χαρ/κό	Ποσοστό % επί του συνόλου των κειμένων (N=485)	Λέξη	Περιοχή Συχνότητας	Κείμενα στα οποία δεν εμφανίστηκε καθόλου το χαρ/κό	Ποσοστό % επί του συνόλου των κειμένων (N=485)
θε	1	0	0	-ους	1	6	1,23
sm	2	2	0,41	-ής	2	7	1,44
kl	3	31	6,39	-ός	3	21	4,32
sθ	4	101	20,82	-ούμε	4	149	30,72
xθ	5	155	31,95	-έτε	5	323	66,59
lt	6	268	55,25	-έως	6	346	71,34
nst	7	371	76,49	-μένης	7	355	73,19
dl	8	419	86,39	-σ(Ξ)(Ψ)ιμο	8	390	80,41
sn	9	451	92,98	-ούσας	9	442	92,37

Μια συνολική εικόνα της σχέσης μεταξύ σπανιότητας ενός χαρ/κού και μεγέθους του κειμενικού δείγματος δίνεται στον παρακάτω πίνακα:

Πίνακας 8

Αριθμός κειμένων που δεν εμφάνισαν καθόλου το γλωσσικό χαρ/κό που αναζητήθηκε βάσει του κειμενικού μεγέθους

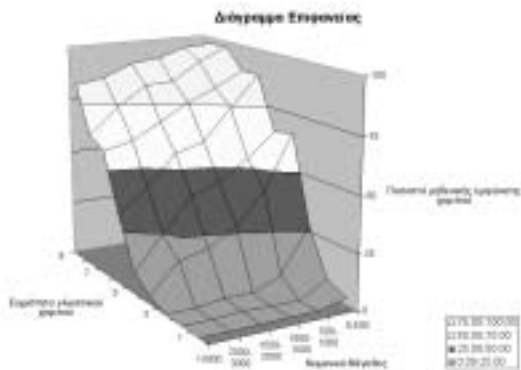
Χαρακτηριστικό	Περιοχή Συχνότητας	Μέγεθος Κειμένου											
		0-500 (N=86)		500-1000 (N=95)		1000-1500 (N=154)		1500-2000 (N=67)		2000-3000 (N=58)		>3000 (N=25)	
		N	%	N	%	N	%	N	%	N	%	N	%
του	1	0	0	0	0	0	0	0	0	0	0	0	0
να		4	4,6	0	0	0	0	0	0	0	0	0	0
για	2	3	3,5	0	0	0	0	1	1,5	0	0	0	0
είναι		17	19,7	0	0	1	0,6	2	3	0	0	0	0
τις	3	15	17,4	4	4,2	1	0,6	0	0	0	0	0	0
ένα		47	54,6	18	18,9	8	5,2	5	7,5	0	0	0	0
ήταν	4	60	69,7	45	47,4	36	23,4	14	20,9	8	13,8	1	4
μου		84	97,6	82	86,3	90	58,4	29	43,2	26	44,8	4	16
πρέπει	5	61	70,9	42	44,2	49	31,8	20	29,8	12	20,7	1	4
χρόνια		57	66,2	37	38,9	50	32,5	17	25,4	9	15,5	1	4
υπάρχουν	6	70	81,4	57	60	81	52,6	32	47,8	18	31	6	24
πολιτική		79	91,8	83	87,3	112	72,7	50	74,6	28	48,3	10	40
κυβέρνηση	7	79	91,9	84	88,4	124	80,5	57	85,1	39	67,2	17	68
αποτελεί		75	87,2	79	83,1	101	65,6	41	61,2	30	51,7	12	48
δώσει	8	84	97,7	84	88,4	124	80,5	55	82,1	45	77,6	19	76
θάλασσα		84	97,7	92	96,8	144	93,5	61	91	48	82,7	19	76
σημειωθεί	9	84	97,7	89	93,7	144	93,5	58	86,6	51	87,9	23	92
σκέψη		85	98,8	94	99	142	92,2	63	94	52	89,7	14	56
the	1	0	0	0	0	0	0	0	0	0	0	0	0
sm	2	2	2,3	0	0	0	0	0	0	0	0	0	0
kl	3	23	26,7	6	6,3	2	1,3	0	0	0	0	0	0
sθ	4	37	43	18	19	38	24,7	3	4,5	5	8,6	0	0
xθ	5	25	29,1	38	40	57	37	25	37,3	8	13,8	2	8
lt	6	72	83,7	65	68,4	77	50	26	38,8	24	41,4	4	16
nst	7	72	83,7	80	84,2	124	80,5	41	61,1	40	69	14	56
dl	8	83	96,5	87	91,6	133	86,4	53	79,1	44	75,9	19	76
sn	9	85	98,8	92	96,8	144	93,5	61	91	48	82,7	21	84
-ους	1	6	7	0	0	0	0	0	0	0	0	0	0
-ής	2	7	8,1	0	0	0	0	0	0	0	0	0	0
-ός	3	17	19,8	3	3,1	0	0	1	1,5	0	0	0	0
-ούμε	4	62	72	38	40	26	16,9	16	23,9	7	12	0	0
-έτε	5	82	95,3	76	80	99	64,3	38	56,7	24	41,4	4	16
-έως	6	81	94,1	76	80	104	67,5	46	68,6	27	46,5	12	48
-μένης	7	78	90,7	70	73,7	113	73,4	47	70,1	32	55,1	15	60
-σ(ξ)(ψ)ιμο	8	84	97,7	86	90,5	118	76,6	55	82	38	65,5	9	36
-ούσας	9	81	94,2	91	95,7	140	90,9	140	90,9	49	84,5	23	92

Ο παραπάνω πίνακας μας επιτρέπει να αξιολογήσουμε την αλληλεπίδραση τριών βασικών μεγεθών σημαντικών για την λήψη αποφάσεων για την ποσοτική διάσταση ενός ΗΣΚ:

α) Το κειμενικό μέγεθος. Στην παρούσα εκτίμηση για λόγους απεικονιστικής απλότητας χρησιμοποιήσαμε την δεύτερη κατηγοριοποίηση κειμενικού μεγέθους (5 κατηγορίες).

- β) Τη σπανιότητα του γλωσσικού χαρ/κού που μελετάται. Τα γλωσσικά χαρακτηριστικά όπως είδαμε και παραπάνω έχουν χωριστεί σε 9 περιοχές συχνότητας με την πρώτη περιοχή να περιλαμβάνει τα πιο συχνά και την 9 τα πιο σπάνια.
- γ) Το ποσοστό των κειμένων όπου δεν εμφανίστηκε καθόλου το γλωσσικό χαρ/κό που αναζητήθηκε.

Διάγραμμα 1



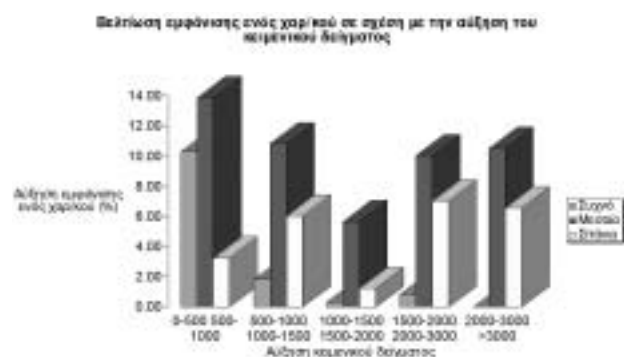
Διάγραμμα επιφανείας που αποτυπώνει τη σχέση μεταξύ της σπανιότητας ενός γλωσσικού χαρ/κού και του κειμενικού μεγέθους στο οποίο αναζητάται

Η εξέταση του διαγράμματος επιφανείας αποκαλύπτει ορισμένα χαρακτηριστικά της αλληλεπίδρασης των τριών παραμέτρων που προαναφέραμε:

1. Τα συχνά γλωσσικά χαρ/κά (μέχρι και την 3η περιοχή συχνότητας) εμφανίζουν σταθερότητα σε κείμενα μεγαλύτερα των 500 λέξεων. Παρουσιάζουν ομοιομορφία εμφάνισης και η απουσία τους δεν σχετίζεται με το μέγεθος του κειμενικού δείγματος.
2. Τα μεσαίως συχνότητας εμφάνισης γλωσσικά χαρ/κά (από την 4η έως και την 6η περιοχή συχνότητας) εμφανίζουν ισχυρή εξάρτηση από μέγεθος του κειμενικού δείγματος. Η πιθανότητα να τα αναζητήσουμε και να μην βρεθούν σε κάποιο κείμενο αυξάνει γραμμικά σε σχέση με το μέγεθος του κειμένου στο οποίο τα ψάχνουμε.
3. Τα σπάνια γλωσσικά χαρ/κά (από την 7η έως και την 9η περιοχή συχνότητας) εμφανίζουν παρόμοια συμπεριφορά με τα χαρ/κά υψηλής συχνότητας εμφάνισης. Παρουσιάζουν σχετική ανεξαρτησία από το μέγεθος κειμένου, αφού η σπανιότητα πραγμάτωσής τους τεκμηριώνεται και στα κειμενικά δείγματα μεγάλου μεγέθους (>3000 λέξεων).

Από την ανάλυση του παραπάνω διαγράμματος προκύπτει ότι τα μεσαίως συχνότητας εμφάνισης χαρ/κά είναι εκείνα που παρουσιάζουν τη μεγαλύτερη ευαισθησία στο μέγεθος του κειμενικού δείγματος. Για να προσδιορίσουμε περαιτέρω το πρόβλημα θα πρέπει να καθορίσουμε ποια αύξηση στο μέγεθος του κειμενικού δείγματος βελτιώνει περισσότερο τα ποσοστά εμφάνισης των γλωσσικών χαρ/κών. Για να απαντηθεί το παραπάνω ερώτημα πρέπει να συγκριθούν τα ποσοστά εμφάνισης των γλωσσικών χαρ/κών σε δύο γειτονικές κατηγορίες μεγέθους κειμενικών δειγμάτων για να δούμε πότε παρουσιάζεται η μεγαλύτερη αύξηση. Η σύγκριση αυτή αποτυπώνεται στο παρακάτω διάγραμμα:

Διάγραμμα 2



Βαθμός επίδρασης της αύξησης του κειμενικού μεγέθους στην αύξηση εμφάνισης ενός γλωσσικού χαρ/κού

Το παραπάνω διάγραμμα δείχνει ότι η μεγαλύτερη βελτίωση στην αντιπροσώπευση συχνών και μεσαίως συχνότητας χαρ/κών σε κείμενα επιτυγχάνεται όταν αυξηθεί το κειμενικό δείγμα από 500 σε 1000 λέξεις. Ένα δεύτερο συμπέρασμα που συνάγεται είναι ότι την χαμηλότερη βελτίωση στην αντιπροσώπευση των χαρ/κών την επιτυγχάνουμε όταν αυξήσουμε το μέγεθος του κειμενικού δείγματος από 1000 – 1500 λέξεις σε 1500 – 2000 λέξεις. Από αυτό συνάγεται ότι τα κειμενικά δείγματα από τις 1000 έως τις 2000 λέξεις παρουσιάζουν σχετική σταθερότητα και ομοιογένεια στην ποσοτική αντιπροσώπευση των γλωσσικών χαρ/κών ανεξαρτήτου συχνότητάς τους. Τα σπάνια χαρ/κά παρουσιάζουν διαφορετική συμπεριφορά. Τις υψηλότερες τιμές αύξησης τις εμφανίζουν στην μετάβαση από τα δείγματα των 1500 - 2000 στα δείγματα των 2000 – 3000 λέξεων. Ωστόσο σημαντική αύξηση παρατηρείται και στην μετάβαση από 500 – 1000 λέξεις σε 1000 – 1500 λέξεις.

4. Συμπεράσματα

Η μέχρι τώρα ανάλυση έδειξε ότι τα συχνά γλωσσικά χαρ/κά αντιπροσωπεύονται ικανοποιητικά από ένα ΗΣΚ ανεξαρτήτως του μεγέθους των επιμέρους κειμένων που το απαρτίζουν. Η αναζήτηση μεσαίας συχνότητας χαρ/κών σχετίζεται άμεσα με το μέγεθος του κειμενικού δείγματος από το οποίο αποτελείται το ΗΣΚ. Σε γενικές γραμμές όταν χρησιμοποιούνται κείμενα ίσα ή μεγαλύτερα των 1000 λέξεων επιτυγχάνεται ικανοποιητική αντιπροσώπηση γεγονός που δεν επιτυγχάνεται σε κειμενικά μεγέθη μικρότερα των 1000 λέξεων. Τα σπάνια γλωσσικά χαρ/κά εμφανίζουν και αυτά σταθερότητα εμφάνισης. Για την ικανοποιητική αντιπροσώπυσή τους χρειάζονται κείμενα μεγάλου μεγέθους (>3000 λέξεων). Ωστόσο, εξετάζοντας τον δείκτη βελτίωσης της εμφάνισης των χαρ/κών αυτών σε διαδοχικές αυξήσεις κειμενικών μεγεθών (βλ Διάγραμμα 2) παρατηρούμε ότι η αύξηση από δείγματα 500 – 1000 λέξεων σε δείγματα 1000 – 1500 λέξεων επιφέρει δραστική βελτίωση της αντιπροσωπευτικότητας αυτών των χαρ/κών.

Εξετάζοντας συνδυαστικά τα παραπάνω δεδομένα μπορούμε να προτείνουμε τη χρήση κειμενικών μεγεθών 1000 – 1500 λέξεων ως την βέλτιστη επιλογή για την διατήρηση υψηλού λειτουργικού εύρους του ΗΣΚ που αναπτύσσουμε. Η λύση αυτή βέβαια προτείνεται μόνο για τα κειμενικά είδη που διαθέτουν τέτοια μεγέθη κειμένων και δεν ισχύει για κειμενικά είδη που εγγενώς είναι πολύ μικρότερα όπως, αγγελίες, δελτία καιρού, ανακοινώσεις κ.ά. Παράλληλα η παραπάνω ποσοτική πρόταση δείχνει την αναγκαιότητα ξεχωριστών ποσοτικών εκτιμήσεων για κάθε γλώσσα. Η υιοθέτηση μετρήσεων που έχουν προταθεί για την Αγγλική γλώσσα δεν εγγυάται τη βέλτιστη εφαρμογή τους σε άλλες γλώσσες, ειδικά όταν αυτές είναι σημαντικά διαφοροποιημένες ως προς τα δομικά χαρ/κά τους, όπως συμβαίνει με τη Νέα Ελληνική.

Βιβλιογραφικές Αναφορές

Aijmer K. & B. Altenberg, (eds), 1991. *English corpus linguistics*. London: Longman.
 Biber, D. & E. Finegan, 1991. "On the exploitation of computerized corpora in variation studies". Στο

Aijmer K. & B. Altenberg (eds), 204 – 220.
 de Haan, P. 1992. "The optimum corpus sample size?". Στο G. Leitner, (ed.), 3 – 19.
 EAGLES, 1996. Preliminary recommendations on Corpus Typology. EAGLES Document, EAG-TCWG-CTYP/P.
 {Ku era, H & Frances, W.N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
 Leitner, G. (ed.), 1992. *New dimensions in English language corpora*. Berlin: Mouton de Gruyter.
 Lehman, R.S. 1991. *Statistics and research design in the behavioral sciences*. California: Wadsworth Publishing Company.

ΜΕΘΟΔΟΛΟΓΙΚΕΣ ΠΑΡΑΤΗΡΗΣΕΙΣ ΓΙΑ ΤΗ ΔΙΔΑΣΚΑΛΙΑ ΤΗΣ ΜΕΤΑΦΡΑΣΗΣ ΤΟΥ ΕΠΙΣΤΗΜΟΝΙΚΟΥ ΚΑΙ ΤΕΧΝΙΚΟΥ ΛΟΓΟΥ*

Μαριάννα Κατσογιάννου - Διαμαντούλα Κόρδα-Σάββα

Το άρθρο αυτό αναφέρεται στη διδασκαλία του μαθήματος της μετάφρασης, ακριβέστερα της μετάφρασης κειμένων όπως τα παρακάτω:

Hematogenous tuberculosis may appear in various morphologic forms: classical miliary tubercles, sizable nodules, necrotic foci simulating abscesses, massive tuberculomas simulating neoplasms and even caseous pneumonia not at all suggestive of a hematogenous origin. (Rubin, Eli H., 1948: Diseases of the Chest, Philadelphia and London, W.B. Saunders Company, p. 194)

Le sommet européen qui s'ouvrira le 7 décembre a d'abord pour objectif [...] de trouver un compromis sur les quatre points restés en suspens: le nombre de commissaires, la «repondération» des voix (un meilleur équilibre démographique), l'extension du vote à la majorité qualifiée et les «coopérations renforcées». (L'Express du 7/12/2000)

Τα παραδείγματα που δίνουμε προέρχονται από διαφορετικές γλώσσες, εφόσον το ζήτημα που μας απασχολεί δεν είναι η διδασκαλία της μετάφρασης από συγκεκριμένη γλώσσα, αλλά η μεθοδολογία που πρέπει να ακολουθείται για την επίλυση προβλημάτων που συνδέονται με συγκεκριμένα είδη κειμένων. Πριν προχωρήσουμε όμως στην παρουσίαση του θέματος, είναι απαραίτητο να αναφέρουμε ορισμένα στοιχεία σχετικά με τις ιδιαιτερότητες των κειμένων αυτών και με τις ερευνητικές και διδακτικές ανάγκες που μας οδήγησαν στην μελέτη τους.

Είναι πλέον σαφές ότι η μετάφραση του επιστημονικού και τεχνικού λόγου αποτελεί ιδιαίτερο κλάδο, ο

οποίος απαιτεί αυτόνομη ερευνητική και διδακτική προσέγγιση. Χρησιμοποιούμε τον χαρακτηρισμό «επιστημονικός και τεχνικός λόγος» για να αποδώσουμε τις έννοιες *sublanguage* και *LSP (language for special purposes)* ή, όπως λέγονται στα γαλλικά *langue de spécialité* ή *langue spécialisée*, όταν αντιπαρατίθεται στη «γενική» γλώσσα. Με τον όρο αυτό αναφερόμαστε σε ένα είδος λόγου που ορίζεται με βάση τον ιδιαίτερο επικοινωνιακό του στόχο: την μεταβίβαση ορισμένης κάθε φορά γνώσης, σχετικής με διάφορους τομείς των επιστημονικών ή τεχνικών δραστηριοτήτων. Η εμφανέστερη συνιστώσα αυτού του είδους λόγου, είναι η χρήση ειδικής ορολογίας. Για τον σχετικό μεταφραστικό κλάδο, θα χρησιμοποιήσουμε την ονομασία τεχνική μετάφραση. Με τη γενική του έννοια, ο όρος «τεχνική μετάφραση» δεν αναφέρεται αποκλειστικά σε κείμενα τεχνικού ή τεχνολογικού περιεχομένου, αλλά και σε κάθε είδος κειμένου το οποίο:

- α) ανήκει σε ορισμένο πεδίο της ανθρώπινης γνώσης και εμπειρίας (διοίκηση, επιστήμη, οικονομία, τέχνη, ...), και
- β) εμπεριέχει ορολογία και συνήθως διαθέτει καθορισμένη φρασεολογία.

Η έννοια της τεχνικής μετάφρασης αντιπαρατίθεται σε αυτήν της λογοτεχνικής. Η σημαντικότερη διάκριση μεταξύ τεχνικής και λογοτεχνικής μετάφρασης μπορεί να αναλυθεί με βάση την παραδοσιακή σχέση μορφής/περιεχομένου. Για το λογοτεχνικό κείμενο η μορφή, τα γλωσσικά μέσα που χρησιμοποιεί ο συγγραφέας, αποτελούν κεντρικό στοιχείο, στο οποίο συχνά βασίζεται και η αισθητική αξία του κειμένου. Το αποτέλεσμα είναι ότι ο μεταφραστής αντιμετωπίζει κυρίως προβλήματα που συναρτώνται με την γλώσσα και το ύφος του συγγραφέα – χωρίς αυτό να σημαίνει ότι άλλοι παράγοντες (όπως ο χρόνος παραγωγής του πρωτότυπου κειμένου, το κοινό στο οποίο απευθυνόταν ή το κοινωνικό περιβάλλον στο οποίο παρουσιάστηκε) δεν λαμβάνονται υπόψη κατά τη διαδικασία της μετάφρασης. Αντίθετα, η τεχνική μετάφραση αποτελεί ένα είδος ερμηνείας του κειμένου, με την έννοια ότι το αποτέλεσμά της (το μεταφρασμένο κείμενο) περιστρέφεται σχεδόν ολοκληρωτικά γύρω από την πληροφορία που πρέπει να μεταφερθεί στη γλώσ-

* Το κείμενο αυτό προέρχεται από ανακοίνωση των συγγραφέων στο 5ο Διεθνές Γλωσσολογικό Συνέδριο για την ελληνική γλώσσα (Παρίσι, Σεπτέμβριος 2001). Μία άλλη μορφή του άρθρου είναι υπό δημοσίευση στο περιοδικό Terminologie et traduction.

σα-στόχο. Η μορφή του κειμένου δεν έχει συνήθως κάποια αισθητική αξία ούτε προσδίδει κάτι το ιδιαίτερο όπως γίνεται στη λογοτεχνία. Αυτό που προέχει είναι το περιεχόμενο, η μεταφορά της πληροφορίας.

Η μεταφορά ενός τεχνικού κειμένου σε άλλη γλώσσα αφενός απευθύνεται σε ένα αναγνωστικό κοινό με διαφορετική γνωστική υποδομή και αντίληψη του θέματος, και αφετέρου δεν υλοποιείται πάντοτε στο πλαίσιο του ίδιου συστήματος εννοιών. Ένα γνωστό παράδειγμα είναι τα νομικά κείμενα, στα οποία οι έννοιες ορίζονται διαφορετικά σε κάθε γλώσσα και ορισμένες φορές δεν έχουν μεταφραστικό αντίστοιχο: για παράδειγμα, η αγγλική δεν διαθέτει όρο κατάλληλο για την μετάφραση του ελληνικού *ανακριτής* ή του γαλλικού *juge d'instruction*, δεδομένου ότι ο ίδιος ο θεσμός είναι ανύπαρκτος και ότι πρόκειται για νομικά συστήματα που λειτουργούν σε διαφορετικές βάσεις. Ανάλογα προβλήματα μπορούν να ανακύψουν σε όλους τους κλάδους, ιδιαίτερα σε σχέση με γνωστικά αντικείμενα των οποίων η ανάπτυξη διαφέρει από χώρα σε χώρα. Στις ιδιαιτερότητες αυτές προστίθεται και το ζήτημα της ορολογίας, το οποίο μάλιστα βλέπουμε συχνά να ταυτίζεται με την ίδια την τεχνική μετάφραση. Στην πραγματικότητα πρόκειται για ένα *επιπλέον* πρόβλημα το οποίο καλείται να αντιμετωπίσει ο μεταφραστής: η ορολογία των επιστημονικών και τεχνικών κειμένων αποτελεί ένα ιδιαίτερο «στρώμα» λόγου (Κουτσιβίτης, 1994: 175), ένα «εργαλείο» επικοινωνίας με σημαντικό ρόλο στην παραγωγή και στην μεταφορά της γνώσης. Έτσι, οι ανάγκες για μετάφραση της ορολογίας αυξάνουν τις απαιτήσεις ως προς την αναζήτηση και αξιοποίηση της τεκμηρίωσης, ενώ παράλληλα διαφοροποιούν την μεθοδολογία προσέγγισης του κειμένου - φορέα ειδικής γνώσης. Για να δώσουμε ένα παράδειγμα από τα ελληνικά δεδομένα, μπορούμε να αναφέρουμε το χώρο της πληροφορικής και ιδιαίτερα των νέων τεχνολογιών, όπου η ορολογία δεν έχει ακόμη παγιωθεί και η μετάφρασή της δεν μπορεί να βασιστεί σε προϋπάρχουσα εμπειρία. Εδώ ανακύπτει το ζήτημα της εξ ολοκλήρου ανάπτυξης ενός εννοιολογικού και γλωσσικού πλαισίου στο οποίο θα εκφραστούν οι νέες σημασίες. Έτσι πολύ συχνά υπάρχουν περισσότερα από ένα μεταφραστικά ισοδύναμα (*διαπροσωπία* και *διεπαφή* για

το *interface*, *μνήμη άμεσης πρόσβασης* και *μνήμη τυχαίας πρόσβασης* για το *random access memory*), από τα οποία κανένα δεν γίνεται ευρέως αποδεκτό. Εκτός από το γεγονός ότι οι μεταφραστές δεν καταλήγουν εύκολα σε κοινές προτάσεις απόδοσης των όρων, η πολυμορφία που παρατηρείται οφείλεται και στην έλλειψη τυποποίησης. Αντιπροσωπευτικό είναι το παράδειγμα του *compact disk* που μεταφράζεται ως *οπτικός δίσκος*, *σύμπυκνος δίσκος*, *συμπαγής δίσκος*, *ψηφιακός δίσκος*, *ψηφιακός συμπαγής δίσκος*, *δίσκος ακτίνας*, ενώ σε όλες τις περιπτώσεις αναφέρεται στο γνωστό μας... *σιντί* (ή, κατ' άλλους, CD). Ένα ειδικότερο πρόβλημα στην περίπτωση της ελληνικής αποτελεί και η γραφή. Πολλοί όροι που μεταφράζονται από άλλες γλώσσες δεν συντομογραφούνται, με αποτέλεσμα τα ξενόγλωσσα ακρώνυμα που τους αντιστοιχούν να επιβάλλουν τη χρήση λατινικών χαρακτήρων: άσχετα με τη μετάφραση που θα επιλέξουμε, το αγγλικό *random access memory* διατηρεί την συντομογραφία RAM, όχι μόνο στον προφορικό, αλλά και στο γραπτό λόγο.

Η ανάπτυξη της ορολογίας και η σχετικά πρόσφατη παγίωσή της ως γλωσσικού επιστημονικού κλάδου συνδέεται στενά τόσο με τις μεταφραστικές δραστηριότητες, όσο και με τις εξελίξεις στους χώρους της θεωρίας της μετάφρασης και της μεταφραστικής τεχνολογίας. Χωρίς να ξεχνάμε ότι η ορολογική έρευνα συνδέεται (σε διαχρονικό και επιστημολογικό επίπεδο) και με άλλες επιστήμες, όπως η λογική και η λεξικογραφία, παρατηρούμε ότι τα τελευταία είκοσι χρόνια έχει αναπτυχθεί μία ιδιαίτερα στενή σχέση μεταξύ τεχνικής μετάφρασης και ορολογίας. Οι γενικότερες συνθήκες της παγκοσμιοποίησης της γνώσης και η αμεσότητα των σύγχρονων επικοινωνιών, συμβάλλουν σημαντικά στην παραπάνω εξέλιξη. Έτσι, όταν σήμερα μιλάμε για ορολογία θεωρείται σχεδόν αυτονόητο ότι αναφερόμαστε σε περισσότερες από μία γλώσσες, συνήθως σε δύο ή τρεις. Και δεν είναι τυχαίο το γεγονός ότι οι πρώτες υπηρεσίες ορολογίας συστάθηκαν στα πλαίσια οργανισμών ή σχολών μετάφρασης ή συνδέθηκαν με την ίδρυση μεταφραστικών υπηρεσιών που ενσωμάτωσαν στις δραστηριότητές τους την ορολογική έρευνα¹. Παράλληλα, οι σπουδές ορολογίας αποτελούν ειδικότητα η οποία διδάσκεται

¹ Εκτός από το ευρύτερα γνωστό παράδειγμα της Ευρωπαϊκής Ένωσης, μπορούμε να αναφέρουμε το Μεταφραστικό Τμήμα του ΟΗΕ στη Νέα Υόρκη, όπως και την Καναδική Μεταφραστική Υπηρεσία, που ακολούθησαν το σχήμα αυτό.

κυρίως σε σχολές μετάφρασης και διερμηνείας, ιδιαίτερα σε συνάρτηση με τις σπουδές τεχνικής μετάφρασης και σε εξαιρετικές μόνο περιπτώσεις εντάσσεται στα πλαίσια διδασκαλίας της γλωσσολογίας ή της φιλολογίας (βλ. και Dubuc, 1978).

Ένα τελευταίο, εξίσου σημαντικό σημείο που πρέπει να αναφερθεί είναι ότι οι τεχνολογίες υποστήριξης της μετάφρασης και της ορολογίας αναπτύσσονται ταυτόχρονα, κατά κανόνα σε κοινούς χώρους και από τους ίδιους ανθρώπους. Είναι άλλωστε σαφές ότι αυτό που ονομάζουμε σήμερα «έκρηξη» της ορολογίας συμπίπτει χρονικά με την ανάπτυξη της τεχνικής μετάφρασης και οφείλεται στις σύγχρονες ανάγκες της επικοινωνίας, που αυξάνονται συνεχώς. Στο πλαίσιο αυτό, η μετάφραση αποτελεί πλέον τον καλύτερο «πελάτη» της ορολογίας: έχει γίνει σαφές ότι η ορολογία μπορεί να αποτελέσει ένα θεωρητικό και πρακτικό εργαλείο μεγάλης εμβέλειας για την υποστήριξη της μετάφρασης επιστημονικών και τεχνικών κειμένων, αλλά και για την επιτυχή αντιμετώπιση προβλημάτων όπως η φαινομενική μη-μεταφρασιμότητα όρων και κειμένων.

Με δεδομένη λοιπόν αυτή τη στενή σχέση μεταξύ τεχνικής μετάφρασης και ορολογίας, μπορούμε να προχωρήσουμε σε ορισμένες παρατηρήσεις σχετικά με τη διδασκαλία των δύο αυτών κλάδων, η οποία έχει ακόμα σημαντικά περιθώρια ανάπτυξης στην Ελλάδα. Εκτός από την επιστημονική έρευνα, που αποτελεί απαραίτητη προϋπόθεση για την ανάπτυξη εφαρμογών και εργαλείων διαχείρισης της ορολογίας, η διδασκαλία της είναι σημαντικός παράγοντας για την καλύτερη επαγγελματική αντιμετώπιση των σχετικών δραστηριοτήτων σε όλους τους τομείς: μετάφραση, μετάφραση/προσαρμογή (*localization*), ανάκτηση πληροφοριών, τεκμηρίωση, τυποποίηση της ορολογίας, κλπ. Ο συνδυασμός όλων των παραπάνω στην περίπτωση της διδασκαλίας της ορολογίας, μπορεί να συνοψιστεί ως εξής: πώς ο υποψήφιος μεταφραστής μπορεί αφενός να ενημερώνεται για τους ορολογικούς πόρους που μπορεί να χρησιμοποιήσει και αφετέρου να επιτυγχάνει τη βέλτιστη αξιοποίησή τους σε σχέση με την εργασία του, δηλαδή αφενός με τη διαδικασία της μετάφρασης (ταχύτερη ανεύρεση και ακριβέστερη χρήση των πηγών), και αφετέρου με το αποτέλεσμά της (παραγωγή ποιοτικά καλύτερων προϊόντων μετάφρασης).

Η πρώτη προϋπόθεση για την αποτελεσματική διδασκαλία των παραπάνω είναι η επιλογή κατάλληλου εκπαιδευτικού υλικού, δηλαδή κειμένων προς μετάφραση στην τάξη. Το θέμα των γενικών κριτηρίων επιλογής, ιδιαίτερα του βαθμού δυσκολίας των κειμένων, έχει ήδη συζητηθεί στην ελληνική βιβλιογραφία (Ντάλτας, 1995: *passim*). Το μόνο που χρειάζεται να προσθέσουμε για την τεχνική μετάφραση είναι ότι, ανάλογα με το επίπεδο σπουδών και τον επιθυμητό βαθμό εξειδίκευσης, μπορεί ο διδάσκων να επιλέξει ανάμεσα σε δύο προσεγγίσεις: να δώσει μία μεγάλη ποικιλία διαφορετικών θεμάτων ή να περιοριστεί σε μικρό αριθμό γνωστικών αντικειμένων στα οποία θα εμβαθύνει μελετώντας και μεταφράζοντας περισσότερα κείμενα.

Εξίσου σημαντικός παράγοντας όμως είναι και οι συνθήκες που προκύπτουν από την προηγούμενη εκπαίδευση και κατάρτιση των σπουδαστών η οποία κατά κανόνα δεν είναι επαρκής ως προς τη γνώση της γλώσσας-πηγής (και συχνά ούτε ως προς τον χειρισμό της γλώσσας-στόχου), αλλά και ως προς την απαιτούμενη εξοικείωση με τα επιστημονικά και τεχνικά κείμενα. Αυτό σημαίνει ότι οι σπουδαστές χρειάζονται συνεχή καθοδήγηση, σε όλη τη διάρκεια της διαδικασίας που ακολουθείται μέχρι την ολοκλήρωση μίας μετάφρασης (βλ. και Durieux, 1988: 144-145). Το θέμα είναι εξαιρετικά ευρύ για να παρουσιαστεί εδώ αναλυτικά, για το λόγο αυτό θα περιγράψουμε μόνο το αρχικό στάδιο της διαδικασίας, δηλαδή τον τρόπο με τον οποίο ένα κείμενο προετοιμάζεται για μετάφραση – χωρίς να αναφερθούμε στα προβλήματα σύνταξης του κειμένου στη γλώσσα-στόχο, ούτε στον τελικό έλεγχο του. Ακριβέστερα, θα αναλύσουμε δύο παραμέτρους που, αν και είναι σημαντικές από εκπαιδευτική άποψη, σπανίως εντάσσονται στην διδακτική διαδικασία: πρόκειται για την πρώτη ανάγνωση του προς μετάφραση κειμένου και για την χρήση τεκμηρίωσης στην οποία στηρίζεται η επιλογή των μεταφραστικών αντιστοίχων της ειδικής ορολογίας.

Το πρώτο στάδιο της μεταφραστικής διαδικασίας είναι ασφαλώς η ανάγνωση του κειμένου που πρόκειται να μεταφραστεί, διαδικασία η οποία συνήθως επαφίεται στους σπουδαστές. Είναι γνωστό όμως πόσο εύκολο είναι να οδηγηθεί κανείς σε παρερμηνείες και μεταφραστικά λάθη εξαιτίας μίας βιαστικής και κυ-

ρίως αμεθόδευτης ανάγνωσης. Στην διδακτική προσέγγιση που προτείνεται εδώ, το στάδιο αυτό της μεταφραστικής διαδικασίας αποτελεί αντικείμενο διδασκαλίας, κατά την οποία ο σπουδαστής μαθαίνει να αντιλαμβάνεται το κείμενο ως διαρθρωμένο λόγο, ως ένα συνεκτικό σύνολο του οποίου το νόημα δεν προκύπτει αθροιστικά από τις σημασίες των επιμέρους στοιχείων. Έτσι η ανάγνωση μετατρέπεται σε δυναμική διαδικασία, που δεν περιορίζεται στην αναγνώριση των λεξιλογικών στοιχείων, αλλά στοχεύει στην ουσιαστική κατανόηση του περιεχομένου.

Ένα αντιπροσωπευτικό παράδειγμα της προσέγγισης αυτής δίνει ο Daniel Gile στο βιβλίο του *Basic Concepts and Models for Interpreter and Translator Training* (1995). Το παράδειγμα αντιπροσωπεύεται από ένα ιατρικό κείμενο που περιέχει όρους άγνωστους στον μέσο αναγνώστη και το οποίο εμφανώς δημιουργεί πρόβλημα στον μεταφραστή που δεν έχει ειδικευτεί σε τέτοια θέματα. Η πρώτη εντύπωση είναι ότι το κείμενο είναι ακατανόητο λόγω της πληθώρας των όρων που περιλαμβάνει. Ο Daniel Gile όμως προτείνει μία πρώτη ανάγνωση κατά την οποία οι όροι αυτοί αντικαθίστανται από παραφράσεις. Το αποτέλεσμα της επεξεργασίας αυτής είναι μία μορφή ερμηνείας μέσω ενός νέου κειμένου, το οποίο, αν και αόριστο είναι πιο προσιτό στον αναγνώστη:

Πρωτότυπο κείμενο

Hematogenous tuberculosis

may appear in various morphologic forms:

classical miliary tubercles,

necrotic foci simulating abscesses,

massive tuberculomas simulating neoplasms

and even caseous pneumonia

not at all suggestive of a hematogenous origin²

Ανάγνωση και ερμηνεία

A certain type of disease which may have something to do with tuberculosis

may appear in various forms called:

"miliary tubercles", which are the classical or most widely found form

something that looks like abscesses

something very large ("tuberculomas") that looks like something else ("neoplasms")

and even symptoms that suggest some kind of pneumonia

that apparently would not normally suggest that the origin of the condition is of the something ("hematogenous") type

Ο συγγραφέας του βιβλίου με αυτό το παράδειγμα επιχειρεί να διαλύσει το μύθο περί εγγενούς δυσκολίας των τεχνικών κειμένων, που ενισχύει την αρνητική ψυχολογία των μαθητευόμενων μεταφραστών. Στην εκπαιδευτική διαδικασία, ο τρόπος αυτός προσέγγισης του κειμένου μπορεί να αντιπροσωπεύεται από τον προφορικό σχολιασμό που θα κάνουν οι ίδιοι οι σπουδαστές ή, αν δεν έχουν επαρκή εμπειρία για κάτι τέτοιο, μπορεί να προκύψει από μία σειρά ερωτήσεων και απαντήσεων, όπως:

- Για ποιο πράγμα μιλάει το κείμενο;
 - Για ένα είδος αρρώστιας.
 - Δηλαδή για τι είδος αρρώστιας;
 - Κάτι που μοιάζει με φυματίωση ή είναι είδος φυματίωσης, πάντως λέγεται *hematogenous tuberculosis*.
 - Και τι μαθαίνουμε για την αρρώστια αυτή από το κείμενο;
 - Ότι μπορεί να εμφανιστεί με διάφορες μορφές.
 - Οι μορφές αυτές όμως έχουν συγκεκριμένες ονομασίες. Μπορούμε να κάνουμε μία λίστα με όσες αναφέρονται στο κείμενο;
- κλπ.

Αν και υποτυπώδης, μία τέτοια ανάλυση μας επιτρέπει να προσεγγίσουμε το κείμενο με απλούστερους όρους και κάνει το νόημά του πιο εύληπτο, προετοιμάζοντάς το έτσι για την μετάφραση. Η παραπάνω συζήτηση θα καταλήξει σε ένα συμπέρασμα του τύπου:

«Το κείμενο αναφέρεται σε κάποιο είδος φυματίωσης και στα κλινικά της συμπτώματα που είναι διάφορα είδη οιδημάτων, κάποια από τα οποία είναι παραπλανητικά διότι δεν παραπέμπουν στην πραγματική τους αιτιολογία.»

στο οποίο θα μπορούν να αναφέρονται οι σπουδαστές κατά τη διάρκεια της μεταφραστικής άσκησης.

Παράλληλα, αυτό το είδος προετοιμασίας μας επιτρέπει να εντοπίσουμε τα προβλήματα που θα παρουσιαστούν σε σχέση είτε με κενά στο επίπεδο των ειδικών γνώσεων που απαιτεί το κείμενο για να γίνει κατανοητό, είτε στην ορολογία που περιέχει. Στο συ-

² Rubin Eli H., 1948: *Diseases of the Chest, Philadelphia and London, W.B. Saunders Company: 194.*

γκεκριμένο παράδειγμα, έχουμε και τις δύο περιπτώσεις: ακόμη και αν ο μεταφραστής πληροφορηθεί από ένα λεξικό γενικής γλώσσας ότι *miliary* είναι ένα επίθετο που σημαίνει «με πολλές προεξοχές», θα εξακολουθήσει να αγνοεί τον δόκιμο ιατρικό όρο που πρέπει να χρησιμοποιήσει (και ο οποίος είναι *κεγχροειδής*). Ο ρόλος της διδασκαλίας, του μαθήματος μετάφρασης, σε σχέση με τις δυσκολίες αυτές είναι να προτείνει μία μέθοδο αντιμετώπισής τους, έτσι ώστε η επίλυση του εκάστοτε προβλήματος να μετατρέπεται σε παράδειγμα, σε υπόδειξη για ένα μελλοντικό τρόπο εργασίας (Durieux, 1988: 149). Για παράδειγμα, σε σχέση με το παραπάνω κείμενο, μπορεί να ζητηθεί από μία ομάδα σπουδαστών να κάνει έναν κατάλογο των όρων του κειμένου και να αναζητήσει τη μετάφρασή τους χρησιμοποιώντας λεξικά και από μία δεύτερη ομάδα να βρει πληροφορίες για το θέμα του κειμένου (π.χ. από άρθρα σε εγκυκλοπαίδειες ή σε επιστημονικά περιοδικά), με βάση τις οποίες θα προσπαθήσει να μεταφράσει την ορολογία χωρίς να χρησιμοποιήσει το λεξικό. Η σύγκριση των αποτελεσμάτων στην τάξη είναι ιδιαίτερα χρήσιμη, κυρίως γιατί επιτρέπει στους σπουδαστές να αποκτήσουν μια σφαιρικότερη θεώρηση του ζητήματος. (Gile, 1995: 156-158).

Για να είναι η ανάγνωση «δυναμική» όπως την χαρακτηρίσαμε παραπάνω, θα πρέπει να εντάσσεται σε μία ευρύτερη διαδικασία οικειοποίησης των εννοιών που εμπεριέχονται στο προς μετάφραση κείμενο και των μεταξύ τους σχέσεων. Με την έννοια αυτή, η ανάγνωση δεν περιορίζεται στο κείμενο που πρόκειται να μεταφραστεί, αλλά επεκτείνεται και στην αναζήτηση και μελέτη τεκμηρίωσης που θα επιτρέψει στον μεταφραστή να αποκτήσει ορισμένες γνώσεις σχετικές με το θέμα. Η τεκμηρίωση αποτελείται από το σύνολο των πληροφοριών που μπορούν να χρησιμοποιηθούν ως πηγές για την μετάφραση, άσχετα με το αν η μορφή τους είναι έντυπη, ηλεκτρονική, οπτική, προφορική, ... Πρόκειται για ένα τρόπο εργασίας που συμβάλλει στην εξοικείωση του μεταφραστή με το αντικείμενο και τον μετατρέπει σε επαρκέστερο αναγνώστη³. Αυτό είναι απαραίτητο γιατί, αντίθετα από τον απλό αναγνώστη, ο μεταφραστής έχει ανάγκη να διαμορφώσει σαφή αντίληψη των εννοιών του κειμένου,

ώστε η αποκωδικοποίηση του περιεχομένου να είναι επαρκής και επιτυχημένη – κάτι το οποίο αποτελεί βασική απαίτηση για μια κατανοητή και εύληπτη μετάφραση. Συχνά άλλωστε οι καθηγητές μετάφρασης έχουν παρατηρήσει ότι όταν το κείμενο-πηγή αναφέρεται σε γνωστικούς τομείς των οποίων ο μεταφραστής δεν έχει εμπειρία, ενισχύεται η τάση για κατά λέξη μετάφραση και το μετάφρασμα χάνει κάθε ειρμό (για τα ελληνικά δεδομένα βλ. Ντάλτας, 1995: 115). Το πρόβλημα γίνεται εντονότερο όταν το κείμενο περιέχει και ορολογία η οποία δεν είναι ούτε γνωστή ούτε κατανοητή στο μεταφραστή. Η σωστή επιλογή και χρήση της τεκμηρίωσης, άλλοτε στη γλώσσα-πηγή και άλλοτε στη γλώσσα-στόχο, είναι ένα ιδιαίτερα χρήσιμο εργαλείο για τη λύση τέτοιων προβλημάτων, δεδομένου ότι δίνει τη δυνατότητα στο μεταφραστή να παρατηρήσει τους όρους στο φυσικό τους περιβάλλον και να κατανοήσει τις έννοιες στις οποίες αντιστοιχούν.

Χωρίς να αναφερθούμε εδώ σε θέματα όπως η διδασκαλία για την αναζήτηση και αξιολόγηση της τεκμηρίωσης⁴, τα οποία συνδέονται με την τεχνολογική κατάρτιση των σπουδαστών, θα παραθέσουμε ένα παράδειγμα από την εμπειρία μας κατά τη διδασκαλία της μετάφρασης κειμένων σχετικών με τους θεσμούς και τα όργανα της Ευρωπαϊκής Ένωσης. Στη διάρκεια αυτού του μαθήματος, προτείναμε για μετάφραση στους φοιτητές ένα γαλλικό κείμενο που αναφερόταν στη σύνοδο κορυφής των 15 στη Νίκαια (7-10 Δεκεμβρίου 2000). Μία από τις δυσκολίες που παρουσιάστηκαν σχετικά με την μετάφραση της ορολογίας αφορούσε τη λέξη *repondération*, άγνωστη τόσο στα λεξικά γενικής γλώσσας, όσο και στα ειδικότερα λεξικά όρων της ΕΕ. Το γεγονός δεν είναι τυχαίο, εφόσον πρόκειται για όρο που έχει δημιουργηθεί πρόσφατα στη γαλλική γλώσσα⁵ και χρησιμοποιείται αποκλειστικά σε κείμενα σαν αυτά που θα δούμε παρακάτω. Ο τρόπος με τον οποίο αντιμετωπίστηκε αυτό το πρόβλημα αποτελεί ένα καλό παράδειγμα για την χρήση της τεκμηρίωσης στην τάξη. Στο πρώτο από τα παρακάτω κείμενα, ο όρος *repondération* αντιπροσωπεύει μία άγνωστη λέξη, η οποία όμως μπορεί να γίνει κατανοητή χάρη στη μελέτη του δεύτερου κειμένου:

³ Για την έννοια της επάρκειας, βλ. Ντάλτας, 1995: 97.

⁴ Για το θέμα αυτό, βλ. Κόρδα & Κατσογιάννου, υπό έκδοση.

⁵ Η παλιότερη εμφάνιση του όρου που μπορέσαμε να εντοπίσουμε χρονολογείται το Νοέμβριο του 1999, στα Πρακτικά του Γαλλικού Κοινοβουλίου.

*L'enjeu majeur de cette réunion était de préparer l'élargissement de l'Union en réformant le fonctionnement de ses institutions. Sur les quatre principaux points figurant à l'ordre du jour, à savoir la taille de la Commission européenne, la **repondération** des voix au Conseil des ministres, l'extension du champ des décisions prises à la majorité qualifiée et, enfin, les coopérations renforcées, [...]. Le Monde, le 11-12-00*

*[...] quatre points restés en suspens: le nombre de commissaires, la «**repondération**» des voix (**un meilleur équilibre démographique**), l'extension du vote à la majorité qualifiée et les «coopérations renforcées». [...] Surtout si les Allemands, lors des discussions sur la **repondération**, renonçaient à **peser plus** que les autres grands pays, malgré leur poids démographique. L'Express du 7/12/2000*

Στο δεύτερο κείμενο παρατηρούμε ότι ο όρος, στην πρώτη του εμφάνιση, βρίσκεται μεταξύ εισαγωγικών (κλασικός τρόπος εισαγωγής νεολογισμών στο γραπτό λόγο) και επεξηγείται στην παρένθεση που ακολουθεί⁶: μαθαίνουμε λοιπόν ότι *repondération* των ψήφων είναι η «βελτιωμένη δημογραφική ισορροπία». Στην επόμενη εμφάνιση της λέξης στο ίδιο κείμενο, τα συμφραζόμενα αποτελούν μία επιπλέον πληροφορία: οι συζητήσεις για την *repondération* θα μπορούσαν να καταλήξουν σε μειωμένη εκπροσώπηση της Γερμανίας. Στο στάδιο αυτό, μπορούμε να θεωρήσουμε ότι έχουμε σχηματίσει σαφή ιδέα για τη σημασία της λέξης: αυτό που επρόκειτο να συζητηθεί στη Νίκαια ήταν η ανακατανομή των ψήφων των κρατών-μελών στο Συμβούλιο της ΕΕ.

Όπως όμως και στο παράδειγμα του ιατρικού κειμένου που είδαμε παραπάνω, εξακολουθούμε να μην διαθέτουμε μεταφραστικό αντίστοιχο: βρισκόμαστε στο στάδιο στο οποίο μπορούμε να αποδώσουμε την έννοια χωρίς να ξέρουμε αν υπάρχει στα ελληνικά όρος ειδικά κατασκευασμένος για το συγκεκριμένο θεματολογικό πλαίσιο. Στο σημείο αυτό, η λύση είναι και πάλι να καταφύγουμε στην τεκμηρίωση, αυτή τη φορά στη γλώσσα-στόχο, επιλέγοντας παράλληλα (ή σχεδόν πα-

ράλληλα) κείμενα – που είναι ίσως η καλύτερη πηγή για την εξεύρεση μεταφραστικών αντιστοιχιών. Σε σχέση με το παράδειγμά μας, ήταν εύκολο να φανταστούμε ότι η κατάλληλη τεκμηρίωση θα έπρεπε να αποτελείται από άρθρα ελληνικών εφημερίδων με θέμα τη Σύνοδο της Νίκαιας, στα οποία πράγματι συναντήσαμε τις μεταφράσεις του όρου που μας ενδιέφερε:

*Σε επτά «κύρια σημεία» περιορίζει η Ελλάδα το ενδιαφέρον της στη διάσκεψη της Νίκαιας και αυτές τις ελληνικές επιδιώξεις θα τονίσει στις παρεμβάσεις του ο πρωθυπουργός κ. Σημίτης. Συγκεκριμένα (α) στο μέγεθος και στη σύνθεση της Επιτροπής, (β) στην **στάθμιση** των ψήφων στο Συμβούλιο, [...] ΤΟ ΒΗΜΑ, 03-12-2000*

*Σημασία επίσης έχει ότι η **επαναστάθμιση** των ψήφων κάθε κράτους-μέλους έγινε στη βάση του σεβασμού των ισορροπιών μεταξύ μεγάλων και μικρών κρατών. Έτσι οι ψήφοι της Ελλάδας αυξάνονται σε 12, με την ειδική πλειοψηφία πάντοτε στο 71% των ψήφων. ΤΟ ΒΗΜΑ, 17-12-2000*

Το γεγονός ότι εμφανίζονται δύο μεταφραστικά αντίστοιχα αποτελεί ένα μεταφραστικό πρόβλημα που έχει ιδιαίτερο ενδιαφέρον από διδακτική άποψη, γιατί μπορεί να αποτελέσει αφορμή για περαιτέρω συζήτηση σχετικά με τη δημιουργία των νεωτερισμών στη γλώσσα και με τις πρωτοβουλίες που μπορεί να πάρει ο μεταφραστής σε τέτοιες περιπτώσεις. Με αυτά τα δεδομένα, η διδακτική διαδικασία δεν επικεντρώνεται στο αντικείμενο της επιλογής του μεταφραστή, αλλά στο ίδιο το πρόβλημα της επιλογής που, όπως είπαμε και παραπάνω, είναι συχνό σε τομείς που βρίσκονται σε εξέλιξη και στους οποίους η ορολογία δεν έχει παγιωθεί. Μέσα από τέτοιες ασκήσεις, οι σπουδαστές κατακτούν μία σταθερή μέθοδο εργασίας που μπορεί να εφαρμοστεί σε οποιαδήποτε περιοχή του επιστητού και την οποία αντιλαμβάνονται ως πρακτικό μέσο, ως εργαλείο που τους επιτρέπει να αναλύσουν, να κατανοήσουν και τελικά να μεταφράσουν τον επιστημονικό και τεχνικό λόγο.

Ένα τελευταίο σημείο που θα αναφέρουμε σχετικά με το ρόλο της τεκμηρίωσης συνδέεται με την ανά-

⁶ Παρατηρούμε ότι η ημερομηνία δημοσίευσης του κειμένου συμπίπτει με την έναρξη των εργασιών της Συνόδου. Αυτό σημαίνει ότι ο όρος *repondération* δεν είχε ακόμη γίνει ευρέως γνωστός, εφόσον το θέμα δεν είχε συζητηθεί εκτενώς στον τύπο. Είναι προφανές ότι αυτός είναι και ο λόγος για τον οποίο ο συγγραφέας αισθάνεται την ανάγκη να εξηγήσει την λέξη στους αναγνώστες του.

πτυξη της συγγραφικής δεξιότητας των σπουδαστών στη γλώσσα-στόχο. Όπως είπαμε και παραπάνω, η τάση για μίμηση της συντακτικής διάταξης του πρωτοτύπου ενισχύεται όταν ο σπουδαστής χρειάζεται να αντιμετωπίσει ένα κείμενο που ξεφεύγει θεματολογικά από αυτά στα οποία είναι συνηθισμένος. Άλλωστε, ακόμη και ο πεπειραμένος μεταφραστής μπορεί να παρασυρθεί σε «αδεξιότητες» ακολουθώντας τις γραμματικές δομές ή τις λεξιλογικές επιλογές του πρωτοτύπου, η δυσκολία όμως αυξάνεται όταν η εμπειρία του είναι περιορισμένη. Αντίθετα, αν ο σπουδαστής έχει στη διάθεσή του τα κατάλληλα συγγραφικά πρότυπα, είναι ευκολότερο να ανταποκριθεί στις απαιτήσεις της εργασίας του παράγοντας κείμενα αντίστοιχα με αυτά τα οποία έχει ήδη μελετήσει.

Τελειώνοντας, θα τονίσουμε και πάλι την ιδιαίτερη σημασία που έχει το τρίπτυχο: ανάγνωση – ανάλυση – τεκμηρίωση. Η επιστήμη της μετάφρασης και ειδικότερα η μεθοδολογία για την μετάφραση ειδικών όρων έχει αφήσει στο παρελθόν παρωχημένες αντιλήψεις που θέλουν τη μετάφραση ως στείρα αντικατάσταση λέξεων από μία γλώσσα σε άλλη. Έχουμε απομακρυνθεί από την εικόνα του μεταφραστή με τους ανοικτούς τόμους των λεξικών γύρω του, και έχουμε πλέον αναγνωρίσει την κοινωνική υπόσταση της μετάφρασης, κυρίως χάρη στις έρευνες που ασχολούνται με τις πολιτισμικές διαστάσεις του φαινομένου. Η νέα, σύγχρονη εικόνα του μεταφραστή είναι αυτή του ερευνητή που επιζητεί την λειτουργική απόδοση των σημασιών στο πλαίσιο του κειμένου. Μαζί με την φιλοσοφία για τη μετάφραση, οφείλει να αλλάξει και ο τρόπος διδασκαλίας της. Η σύγχρονη εκπαιδευτική διαδικασία, ξεκινώντας από την ανάγκη του μεταφραστή να αναλύσει και να κατανοήσει ένα κείμενο, τον οδηγεί σε λύσεις αποδεκτές για τον αναγνώστη, δηλαδή σε μεταφράσματα που αποδίδουν πιστά ένα σημασιολογικό περιεχόμενο, χωρίς προσκόλληση στις γλωσσικές μορφές του πρωτοτύπου. Και αυτό υλοποιείται με ευέλικτα μαθήματα, με τεχνικές που φέρνουν τα επιθυμητά αποτελέσματα και που αιτιολογούν συστηματικά τις επιλογές που θα έκανε ένας έμπειρος μεταφραστής. Αυτά τα μέσα χρειάζεται η αίθουσα διδασκαλίας για το μάθημα της μετάφρασης και με αυτά θα προετοιμάσουμε μεταφραστές που θα είναι ανταγωνιστικά στη διεθνή και ευρωπαϊκή αγορά.

Βιβλιογραφία

- Auger Pierre & Boulanger Jean-Claude, 1995 (2e éd.): *Terminologie et Terminographie - Recueil de notes de cours*, Québec, Université Laval.
- Brian Harris, 1997: *Translation and Interpreting Schools*. Amsterdam/Philadelphia, John Benjamins.
- Actes: *Conférence Pour Une Infrastructure Terminologique En Europe*, Paris, le 13-14-15 mars 2000, Maison de l'UNESCO, <http://www.unilat.org/dtil/etis/actasTDCnet>
- Delisle Jean, 1981: *L'enseignement de l'interprétation et de la traduction: de la théorie à la pédagogie*, Ottawa, PU d'Ottawa.
- Durieux, Christine, 1988: *Fondement didactique de la traduction technique*, Paris, Didier Érudition.
- Fry Deborah (ed.), 1996: *POINTER Final Report: Proposals for an operational infrastructure for terminology in Europe*, <http://www.computing.surrey.ac.uk/ai/pointer>.
- Gile Daniel, 1995: *Basic Concepts and Models for Interpreter and Translator Training*, Amsterdam/Philadelphia, John Benjamins.
- Gerding-Salas Constanza, 2000: «Teaching Translation: Problems and Solutions», *Translation Journal*, vol. 4/3, <http://accurapid.com/journal/13educ.htm>.
- Harris Brian, 1997: *Translation and Interpreting Schools. Language International World Directory*, vol. 2, Amsterdam/Philadelphia, John Benjamins.
- Κόρδα Διαμαντούλα & Κατσογιάννου Μαριάννα, υπό έκδοση: «Ο ρόλος της τεκμηρίωσης στην μεταφραστική διαδικασία και την κατάρτιση του μεταφραστή», στο *Μεταφράζοντας στον 21^ο αιώνα: τάσεις και προοπτικές*.
- Κουτσιβίτης Βασίλης, 1994: *Θεωρία της μετάφρασης*, Ελληνικές Πανεπιστημιακές Εκδόσεις.
- Ντάλτας Περικλής, 1995: *Δοκίμια διδακτικής της μετάφρασης - Προς μία κοινωνιογλωσσολογική θεώρηση*, Αθήνα, Επικαιρότητα.

III. Γλωσσάριο Όρων Γλωσσικής Τεχνολογίας και Πληροφορικής

Λάβαμε τις ακόλουθες προτάσεις εκ μέρους των κ.κ. Μ. Δενδρινού, Α. Φωτοπούλου, Γ. Μικρού και Γ. Καραγιάννη στις οποίες εκφράζονται προβληματισμοί για την απόδοση κάποιων ξένων όρων στην ελληνική.

Μάρκος Δενδρινός

- **PMR - Private Mobile Radio:** Ιδιωτικές υπηρεσίες ιδιωτικών συστημάτων κινητής ράδιο-επικοινωνίας
- **PEM - Postal Electronic Mail:** Ταχυδρομική Υπηρεσία Ηλεκτρονικού Ταχυδρομείου
- **data server:** υπηρέτης Η/Υ δεδομένων
- **e-mail server:** υπηρέτης Η/Υ για υπηρεσίες ή ταχυδρομείου
- **fax server:** υπηρέτης Η/Υ για υπηρεσίες τηλεομοιοτυπίας
- **FAQ - frequently asked questions:** συχνά υποβαλλόμενες ερωτήσεις
- **outsourcing:** ανάθεση εργασιών σε τρίτους
- **service level:** επίπεδο παρεχόμενης υπηρεσίας
- **Alternative Dispute Resolution Schemes:** Εναλλακτικοί Μηχανισμοί Επίλυσης Διαφορών
- **portal:** πύλη
- **help desk:** γραφείο υποστήριξης
- **ECDL:** Ευρωπαϊκό Δίπλωμα Κατάρτισης στους Υπολογιστές
- **on-line:** επιγραμμικός
- **MIS:** Πληροφοριακό Σύστημα Διοίκησης (ΠΣΔ)
- **geographic or geospatial information infrastructures:** γεωγραφικές-χωρικές υποδομές πληροφοριών
- **digital divide:** ψηφιακός αποκλεισμός
- **basic ICT skills:** βασικές δεξιότητες ΤΠΕ
- **Vessel Traffic MIS:** ΠΣΔ ελέγχου θαλάσσιας κυκλοφορίας
- **business communities:** επιχειρηματικές κοινότητες
- **comparative shopping:** σύγκριση τιμών προσφερόμενων προϊόντων
- **one-stop shop:** γραφείο συνολικής εξυπηρέτησης
- **call center:** κέντρο κλήσεων
- **infopoint:** σημείο παροχής πληροφοριών
- **venture capital:** επιχειρηματικό κεφάλαιο υψηλού κινδύνου
- **e-business:** ηλεκτρονικό επιχειρείν

- **e-learning:** ηλεκτρονική μάθηση
- **backbone network:** δίκτυο κορμού

Αγγελική Φωτοπούλου

Όροι γλωσσολογίας

- **Phrases figées vs Phrases libres:** Στερεότυπες προτάσεις vs Ελεύθερες - μη στερεότυπες προτάσεις π.χ. Τρώω τα λυσσακά μου vs τρώω τα μπισκότα μου
- **Phrases sémi-figées:** Ημι-στερεότυπες προτάσεις
- **Figement:** Στερεοτυπία (παγίωση)
- **Verbes supports** ex. avoir du courage: Επιβοηθητικά ρήματα π.χ. έχω θάρρος
- **Verbes datifs** ex. donner: Δοτικά ή ρήματα συναλλαγής π.χ. δίνω
- **Verbes de communication** ex. dire: Ρήματα επικοινωνίας π.χ. λέω
- **Contraintes lexicales:** Λεξικοί περιορισμοί
- **Contexte:** Συγκεκριμένο
- **Dativité:** Δοτικότητα
- **Datif lexical vs datif étendu** ex. je lui ai donné un livre vs je lui ai préparé un café: Λεξικοποιημένη ή λεξική δοτική vs Εκτενής δοτική π.χ. Του έδωσα ένα βιβλίο vs του έφτιαξα ένα καφέ
- **Déterminants** ex. le, la, un, chaque: Προσδιοριστές π.χ. ο, η, ένας, κάθε.
- **Les noms abstraits** ex. le subconscient: Τα αόριστα ή μη συγκεκριμένα ουσιαστικά π.χ. το υποσυνείδητο
- **Les noms concrets** ex. le crayon: Τα απτά ή συγκεκριμένα ουσιαστικά π.χ. το μολύβι
- **Les noms inaliénables** ex. la main: Τα αμεταβίβαστα ή αναπαλλοτρίωτα ουσιαστικά π.χ. το χέρι

Γιώργος Μικρός

- **audience design:** Προσαρμογή (ομιλίας) στο ακροατήριο (του ομιλητή). Στη σχετική βιβλιογραφία έχει προταθεί το "Σχεδίαση ακροατηρίου". Αν και είναι ακριβής μετάφραση ο όρος που προκύπτει δεν έχει κάποια συγκεκριμένη εννοιολογική ταυτότητα στη Νέα Ελληνική.
- **balanced bilingualism:** Ισοσκελισμένη διπλογλωσσία
- **bilingualism:** Διπλογλωσσία. Χρησιμοποιείται ευρύτατα ο όρος "δίγλωσσος" για τον ομιλητή που κατέχει σε άριστο επίπεδο δύο γλώσσες. Ωστόσο, ο όρος αυτός συμπίπτει με τον όρο "diglossia" ο

οποίος επίσης μεταφράζεται στη ΝΕ ως διγλωσσία. Προτείνεται η χρήση του όρου διπλογλωσσία για το Bilingualism και διγλωσσία για το Diglossia.

- **diglossia:** Διγλωσσία. Έχει χρησιμοποιηθεί και ο όρος "Κοινωνική Διγλωσσία" για να διακρίνει την περίπτωση της Δημοτικής - Καθαρεύουσας από αυτή της κατοχής δύο γλωσσικών συστημάτων (bilingualism). Ωστόσο, αφού για το "Bilingualism" μπορούμε να κρατήσουμε τον όρο "διπλογλωσσία", για τον όρο "diglossia" μπορεί να διατηρηθεί ο απλός όρος "διγλωσσία", αφού γενικότερα οι μονολεκτικοί όροι θα πρέπει να προτιμούνται έναντι συναφών πολυλεκτικών σχηματισμών.
- **language planning:** Γλωσσικός προγραμματισμός
- **variants:** Ποικίλλοντες τύποι
- **style:** Υφος

Γιώργος Καραγιάννης

- **web page:** Ιστοσελίδα
- **web site:** Ιστότοπος
- **robust:** Έχουν προταθεί κατά καιρούς: Εύρωστος, Ρωμαλαίος, Στιβαρός, Ανθεκτικός, Σθεναρός. Η λέξη "robust" συνοδεύει συνήθως τις εξής λέξεις: αλγόριθμος, λογισμικό, σύστημα, αντίδραση, κ.λ.π. Η λέξη Σθεναρός έχει όλες τις προϋποθέσεις να γίνει δόκιμος όρος γιατί ταυτόχρονα είναι μια σχετικά σπάνια λέξη στην καθημερινή γλώσσα, ενώ αρχίζει να χρησιμοποιείται αρκετά συχνά σε περιβάλλον εντάσεως όρων.
- **agent:** έχουν προταθεί πράκτορας και δράστης. Το "δράστης" είναι ορθότερο γιατί εκφράζει καλύτερα την λειτουργία του συγκεκριμένου λογισμικού.

IV. Τα Νέα του ΙΕΛ

Η Φιλογλωσσία+ στην κορυφαία θέση του εκπαιδευτικού λογισμικού για την Ελληνική Γλώσσα



Την πρώτη θέση μεταξύ των προϊόντων λογισμικού που έχουν ως θέμα τους την ελληνική γλώσσα πήρε το CD-ROM του ΙΕΛ Φιλογλωσσία+ στο σχετικό τεστ που θα δημοσιευτεί στο τεύχος Μαρτίου 2002 του περιοδικού RAM.

Με βαθμολογία 9,6 η Φιλογλωσσία+ πέτυχε την υψηλότερη αξιολόγηση μεταξύ των ελληνικών παραγωγών όλων των κατηγοριών. Στο ίδιο τεστ, άλλα τέσσερα εκπαιδευτικά CD-ROM του ΙΕΛ αξιολογήθηκαν θετικά και εμφανίζονται σε αντίστοιχες θέσεις: Φιλογλωσσία (πολύγλωσση έκδοση), Λογομάθεια+, Ηρόδοτος, Κλεοπάτρα.



Ας σημειωθεί ότι στο πρώτο μέρος του τεστ που θα δημοσιευτεί στο τεύχος Ιανουαρίου του RAM, το εκπαιδευτικό λογισμικό "Ενας Ιππότης στο Κάστρο των Γραμμάτων", για την εκμάθηση της Ελληνικής σε μειονοτικά σχολεία, κέρδισε την πρώτη θέση στην κατηγορία του συγκεντρώνοντας και πάλι την μεγαλύτερη βαθμολογία του τεστ.

Νέο Βιβλίο



"Μύκονος-Δήλος, Συνοδευτικό εγχειρίδιο χρήσης του ομώνυμου DVD-ROM"

Στο πλαίσιο της σειράς "ΣΑΠ-ΦΩ", αναμένεται η κυκλοφορία της έντυπης έκδοσης "Μύκονος-Δήλος, Συνοδευτικό εγχειρίδιο χρήσης του ομώνυμου

DVD-ROM". Με την έντυπη έκδοση συμπληρώνεται η σειρά που αποτελείται επίσης από το αντίστοιχο DVD-ROM και DVD-VIDEO. Στο βιβλίο "Μύκονος-Δήλος" επιχειρείται μία ολοκληρωμένη παρουσίαση των δύο νησιών με πρωτότυπα κείμενα, η σύνθεση των οποίων είναι προϊόν εμπειριστατωμένης έρευνας, και πλούσιο εικονογραφικό υλικό. Αποτελεί ταυτόχρονα έναν χρηστικό οδηγό πλοήγησης για το DVD-ROM "Μύκονος - Δήλος, ένα ταξίδι στο κέντρο των Κυκλάδων".

Στόχος της σειράς "ΣΑΠΦΩ" είναι να προσφέρει ένα δομημένο θησαυρό γνώσεων για τα ελληνικά νησιά που θα αναδείξει την πολιτισμική ιδιαιτερότητα και την ιστορική σημασία του αιγαιακού χώρου.

Νέο Προϊόν



DVD-ROM "NOHMA - Λεξικό της Ελληνικής Νοηματικής Γλώσσας"

Το DVD-ROM NOHMA - Λεξικό της Ελληνικής Νοηματικής Γλώσσας (ΕΝΓ) είναι το πρώτο ηλεκτρονικό λεξικό ελληνικών νοημάτων με μετάφραση στα νέα ελληνικά. Παράλληλα, αποτελεί την πρώτη ελληνική παραγωγή DVD-ROM.

Περιέχει 3.000 βιντεοσκοπημένες λέξεις και απευθύνεται σε γνώστες της νοηματικής, αλλά και σε σπουδαστές που μαθαίνουν την ΕΝΓ ως δεύτερη γλώσσα.

Πρόγραμμα Μεταπτυχιακών Σπουδών στην Γλωσσική Τεχνολογία "Τεχνογλωσσία"

Στο δεύτερο εξάμηνο του δεύτερου κύκλου του, βρίσκεται το Πρόγραμμα Μεταπτυχιακών Σπουδών στην Γλωσσική Τεχνολογία "Τεχνογλωσσία", το οποίο αποτελεί συνεργασία του Γλωσσολογικού Τομέα της Φιλοσοφικής Σχολής του Παν/μίου Αθηνών και του Τομέα Πληροφορικής του ΕΜΠ. Στο εξάμηνο αυτό, οι σπουδαστές θα μελετήσουν θέματα που αφορούν την Επεξεργασία Σωμάτων Κειμένων, τις γλωσσολογικές εφαρμογές της Τυπικής Λογικής, την Ανάλυση και Μοντελοποίηση Φωνής, την Ηλεκτρονική Λεξικογραφία και τον Λογικό Προγραμματισμό.

Μέσα από την εξάσκηση στο εργαστήριο και την εξοικείωση με πραγματικά συστήματα επεξεργασίας φυσικής γλώσσας και ομιλίας, οι συμμετέχοντες στην "Τεχνογλωσσία" θα αποκτήσουν τις απαραίτητες δεξιότητες για την ένταξή τους στον ραγδαία αναπτυσσόμενο τομέα της βιομηχανίας γλώσσας και γνώσης.

EuroMaT

Μετά από σχεδόν οκτώ συναπτά έτη λειτουργίας, το Γραφείο Υπηρεσιών Μεταφραστικής Τεχνολογίας EuroMat σταμάτησε πρόσφατα να λειτουργεί ως κόμβος για την παροχή υπηρεσιών προς το ελληνικό δημόσιο.

Το γραφείο ιδρύθηκε από το Ινστιτούτο Επεξεργασίας του Λόγου τον Μάιο του 1994 για να προωθήσει την ανάπτυξη του γλωσσικού ζεύγους Αγγλικά-Ελληνικά στο πλαίσιο της πλατφόρμας Μηχανικής Μετάφρασης EE-SYSTRAN και να παρέχει δωρεάν υπηρεσίες μηχανικής μετάφρασης προς τον ελληνικό δημόσιο τομέα.

Μεταφράστηκαν περισσότερες από 60.000 σελίδες για λογαριασμό όλων σχεδόν των υπουργείων, με το ΙΕΛ να δρα ως κόμβος για τους τελικούς χρήστες.

Τώρα οι εργαζόμενοι στις δημόσιες υπηρεσίες έχουν τη δυνατότητα να στέλνουν οι ίδιοι, απευθείας χωρίς την μεσολάβηση του γραφείου EuroMat, κείμενα προς μετάφραση στο σύστημα Μηχανικής Μετάφρασης Systran.

Περισσότερες πληροφορίες στη σελίδα:

<http://www.ilsp.gr/euromatfreegr.html>

Οδηγίες υποβολής άρθρων προς δημοσίευση

Τα προς δημοσίευση άρθρα πρέπει να υποβάλλονται σε τρία αντίτυπα στη διεύθυνση:

ΙΕΛ-Ινστιτούτο Επεξεργασίας του Λόγου

Αρτέμιδος 6 & Επιδαύρου,

151 25 Παράδεισος Αμαρουσίου

και σε ηλεκτρονική μορφή στη διεύθυνση: **vpana@ilsp.gr** με την ένδειξη **Λογοπλοήγηση** ως θέμα.

Τα κείμενα πρέπει να κατατίθενται σε μορφή αρχείου **Word** δίσητης σελίδας.

Όλα τα γραφήματα πρέπει να κατατίθενται σε μορφή **.tif** αρχείου ή σε οποιαδήποτε άλλη μορφή υψηλής ανάλυσης (**high resolution 300 dpi**).

Η βιβλιογραφία πρέπει να ακολουθεί το σχήμα που παρατίθεται ως υπόδειγμα:

Daille, B., E. Gaussier & J.M. Lange (1994). Towards automatic extraction of monolingual and bilingual terminology. In *Proceedings of COLING 94*, 515_521.

Jackendoff, R. (1996). *Languages of the Mind*. The MIT Press.

Στο κείμενο οι βιβλιογραφικές αναφορές τοποθετούνται σε παρενθέσεις με το όνομα του συγγραφέα και τη χρονολογία έκδοσης π.χ. (**Jackendoff, 1996**) ή αν παρατίθενται περισσότερες από μία αναφορές (**Jackendoff, 1996; Daille et al, 1994**).

Οι υποσημειώσεις τοποθετούνται στο τέλος κάθε σελίδας.

Πριν από το βασικό κείμενο προηγείται περίληψη 200 λέξεων.

Κάθε άρθρο δεν πρέπει να υπερβαίνει τις 10 σελίδες Α4.

Κάθε άρθρο πρέπει να συνοδεύεται από σημείωμα με το όνομα και τη διεύθυνση, την ηλεκτρονική διεύθυνση, το τηλέφωνο και το fax του βασικού συγγραφέα.

Τα άρθρα που υποβάλλονται προς δημοσίευση στη *Λογοπλοήγηση* δεν πρέπει να έχουν δημοσιευθεί προηγουμένως σε άλλο έντυπο.

Instructions to Authors

Contributions should be submitted in three hard copies to:

ILSP-Institute for Language and Speech Processing

6 Artemidos & Epidaurou str.,

151 25 Paradisos Amarousiou

and by e-mail at: **vpana@ilsp.gr** with the indication **Logonavigation** in the Subject line.

Articles should be submitted as **Word** documents.

All graphs must be submitted in **.tif** format or in any other **high resolution 300 dpi** form.

References should follow the cited pattern:

Daille, B., E. Gaussier & J.M. Lange (1994). Towards automatic extraction of monolingual and bilingual terminology. In *Proceedings of COLING 94*, 515_521.

Jackendoff, R. (1996). *Languages of the Mind*. The MIT Press.

References inside the body text should have the form (**Jackendoff, 1996**) or (**Jackendoff, 1996; Daille et al, 1994**).

Footnotes are placed at the bottom of each page.

An abstract of 200 words, precedes the body text.

Submissions should not exceed 10 pages (A4)

Every submission should be accompanied by a note with the name, address, e-mail address, telephone and fax number of the main author.

Articles submitted to *Logonavigation* should not be published elsewhere before.