

**ΣΧΕΣΙΑΚΗ ΟΜΟΙΟΤΗΤΑ ΜΕΤΑΞΥ ΤΟΥ ΔΙΑΦΟΡΕΤΙΚΟΥ
ΜΕΓΕΘΟΥΣ, ΣΤΥΛ ΚΑΙ ΒΑΡΟΥΣ ΓΡΑΜΜΑΤΟΣΕΙΡΑΣ ΚΕΙΜΕΝΩΝ
ΚΑΙ ΤΗΣ ΕΝΤΑΣΗΣ, ΤΟΥ ΤΟΝΙΚΟΥ ΥΨΟΥΣ ΚΑΙ ΡΥΘΜΟΥ
ΟΜΙΛΙΑΣ ΚΑΤΑ ΤΗΝ ΑΚΟΥΣΤΙΚΗ ΑΠΟΔΟΣΗ ΜΕ ΣΥΝΘΕΤΙΚΗ
ΟΜΙΛΙΑ**

Φίλιππος Κατσούλης* και Γεώργιος Κουρουπέτρογλου**

**Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Μεθοδολογίας, Ιστορίας και Θεωρίας της Επιστήμης, Πρόγραμμα Μεταπτυχιακών Σπουδών στη Βασική και Εφαρμοσμένη Γνωσιακή Επιστήμη, phikats@phs.uoa.gr, Τηλ.: 210 7275507, Fax: 210 7275504*

***Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών, Εργαστήριο Επεξεργασίας Ομιλίας koupe@di.uoa.gr, Τηλ.: 2107275305, Fax:2106018677*

Η έρευνα αυτή εντάσσεται σε μια ευρύτερη προσπάθεια να αποδοθούν τα σχολικά εγχειρίδια σε ακουστική μορφή μέσω συστημάτων μετατροπής-Εγγράφων-σε-Ομιλία (Kouroupetroglou & Tsonos, 2008; Tsonos et al 2007) που αποτελούν την επόμενη γενιά συστημάτων μετατροπής Κειμένου-σε-Ομιλία (Fellbaum & Kouroupetroglou, 2008), με έναν αντιληπτικά αποδεκτό τρόπο, με σκοπό να χρησιμοποιηθούν στην διδασκαλία των τυφλών και γενικότερα εντυποαναπήρων μαθητών (Freitas & Kouroupetroglou, 2008). Στα λογισμικά ανάγνωσης οθόνης (JAWS, HAL, κ.λ.π.) που χρησιμοποιούνται σήμερα είναι ενσωματωμένα συστήματα μετατροπής κειμένου σε ομιλία, τα οποία αποδίδουν το κειμενικό περιεχόμενο του εγγράφου, αλλά παρουσιάζουν προβλήματα κατά την παρουσίαση των οπτικών σημάτων του, όπως το μέγεθος των χαρακτήρων, η έντονη και πλάγια γραφή (bold, italics), οι πίνακες κλπ, Τέτοια σήματα γραφής (text signals) χρησιμοποιούνται από τον συγγραφέα¹ σε κάθε δομημένο κείμενο ως γραπτά τεχνάσματα (writing devices) που έχουν ως σκοπό να τονίσουν το περιεχόμενο ή τη δομή ενός κειμένου, δημιουργώντας μια αντιληπτική αντίθεση, χωρίς να προσθέσουν κάτι στο περιεχόμενό του (Lorch, R. F., 1989). Αυτό έχει ως συνέπεια οι χρήστες λογισμικών ανάγνωσης οθόνης να χάνουν σημαντικές πληροφορίες από ένα δομημένο έγγραφο και να μειονεκτούν σε σχέση με τους βλέποντες στη δημιουργία ενός δομικού σχήματος / structural schema (Kintsch & Yarbrough, 1982; Olhausen & Roller, 1988; Brooks & Dansereau, 1983). Οι βλέποντες, καθώς διαβάζουν

¹ Με τον όρο συγγραφέας εννοούμε όλους τους ειδικούς, οποίοι εμπλέκονται στην οπτική απόδοση ενός δομημένου κειμένου, π.χ. γραφίστας, επιμελητής έκδοσης, κ.λ.π.

δημιουργούν οργανωτικά πλαίσια (organizational framework), τα οποία τους βοηθούν να κωδικοποιήσουν τις πληροφορίες και να τις ανακτήσουν αργότερα από τη μακρόχρονη μνήμη. Το μέγεθος, των χαρακτήρων είναι ένα σημαντικό στοιχείο του κειμένου και ανάλογα με το πώς χρησιμοποιείται (π.χ., σε τίτλους, υποσημειώσεις κλπ) βοηθά τους αναγνώστες να δημιουργούν ένα Οργανωτικό πλαίσιο για την κωδικοποίηση του κειμένου (Spyridakis, 1989a and 1989b), το οποίο τους διευκολύνει να διατηρούν και να ανακτούν τις πληροφορίες. (Spyridakis, 1989a, 1989b; 1991; Hartley et al., 1980; Lorch & Lorch, 1996a; 1996b; Sanchez, Lorch & Lorch, 2001). Το διαφορετικό βάρος και στυλ των χαρακτήρων, εκτός από τους τίτλους διαφορετικών επιπέδων, χρησιμοποιούνται για τη δημιουργία διαφορετικών βαθμών έμφασης, μέσα από την οπτική αντίθεση, τονίζοντας νέους όρους, σημαντικές λέξεις, διαφορετικό ύφος, κ.λ.π.

Οι σχετικές εργασίες που έχουν παρουσιαστεί μέχρι σήμερα έχουν εστιάσει το ενδιαφέρον τους στη χρήση φυσικών ήχων ή αλλιώς ακουστικών εικονιδίων (Gaver, W., 1986; 1988; 1990; Mynatt, E.D. & Edwards, W.K., 1992) και μουσικών τόνων ή Αφηρημένων Ακουστικών Μοτίβων (Blattner, M.M et al, 1990; Brewster, S.A et al., 1993) για να αναπτύξουν μια δομημένη ακουστική απόδοση για τους τυφλούς χρήστες, καθώς και στη χρήση διαφόρων ηχοσειρών για την αναπαράσταση των οπτικών στοιχείων σε έγγραφα ιστοσελίδων μεταβάλλοντας τα χαρακτηριστικά της ομιλίας όπως το τονικό ύψος (ή τονικότητα), η ένταση και η ταχύτητα (Xydias et al., 2004).

Τα Ακουστικά Εικονίδια (ΑΕ) είναι ήχοι από το καθημερινό μας περιβάλλον, οι οποίοι χρησιμοποιούνται για να δημιουργήσουμε μια διαισθητική σύνδεση με το αντικείμενο ή την ενέργεια που αναπαριστούν.

Επειδή, οι σχεδιαστές ακουστικών διεπαφών (auditory interfaces) χρήστη συνήθως δεν είναι εξοικειωμένοι με τα επιστημονικά πεδία της ακουστικής αντίληψης, της προσοχής και της νόησης, σχεδιάζουν διάφορα θέματα βασιζόμενοι κυρίως στη διαίσθησή τους χωρίς να έχουν προηγηθεί οι κατάλληλες αντιληπτικές μελέτες και συνήθως οδηγούνται σε χειρότερα αποτελέσματα από αυτά που αναμένουν.

Προκειμένου να ξεπεραστούν τα προβλήματα που εντοπίστηκαν σε προηγούμενες έρευνες, στις πειραματικές μελέτες που πραγματοποιήσαμε έγινε προσπάθεια εξεύρεσης σχεσιακής ομοιότητας (relational similarity) μεταξύ χαρακτηριστικών της ομιλίας και της τυπογραφίας. Έτσι, βρέθηκε ότι οι συμμετέχοντες συσχέτιζαν με πολύ υψηλά ποσοστά

συνέπειας τα μεγέθη «ένταση» φωνής και «μέγεθος» γραμματοσειράς, όπως επίσης συνδυασμένες τιμές του τονικού ύψους και του ρυθμού ομιλίας με το βάρος ή και το στυλ των χαρακτήρων.

Στην παρούσα έρευνα μελετήσαμε αν η ταυτόχρονη μεταβολή στο μέγεθος, το βάρος ή και το στυλ της γραμματοσειράς μιας λέξης, γίνεται αντιληπτή όταν αυτή αποδίδεται ακουστικά με την κατάλληλη διαφοροποίηση των τιμών της έντασης του ρυθμού και του τονικού ύψους της ομιλίας, σύμφωνα με τα ευρήματα των προηγούμενων ερευνών μας.

Οι πειραματικές διαδικασίες έγιναν εξατομικευμένα με τη χρήση Η/Υ και ακουστικών. Τα ακουστικά ερεθίσματα δημιουργήθηκαν με τη χρήση της ειδικής ερευνητικής πλατφόρμας μετατροπής Εγγράφων-σε-Ακουστική Μορφή (Document-to-Audio) του Εργαστηρίου Επεξεργασίας Ομιλίας του ΕΚΠΑ (Xydas, G. et al, 2005), ενώ τα οπτικά ερεθίσματα παρουσιάστηκαν στην οθόνη του Η/Υ με την εφαρμογή MS Power Point.

Το δείγμα αποτελείτο από 31 εκ γενετής τυφλά παιδιά και 31 βλέποντα. Τα 14 παιδιά από κάθε ομάδα ήταν μαθητές πρωτοβάθμιας εκπαίδευσης και τα 17 μαθητές δευτεροβάθμιας εκπαίδευσης, από τα οποία κανένα δεν αντιμετώπιζε πρόβλημα ακοής και ήταν εξισωμένα ως προς την ηλικία το φύλο και το κοινωνικοοικονομικό επίπεδο.

Παρουσιάστηκαν 36 ερεθίσματα:

- Στα 18 πρώτα ερεθίσματα το υπόδειγμα ήταν ακουστικό, δηλαδή η ακουστική απόδοση μιας λέξης, στην οποία άλλαζε η ένταση της φωνής, το τονικό ύψος και ο ρυθμός, ενώ τα ερεθίσματα επιλογής ήταν κάθε φορά τρεις λέξεις οι οποίες διέφεραν ως προς το μέγεθος το βάρος ή και το στυλ της γραμματοσειράς.
- Στα υπόλοιπα 18 ερεθίσματα το υπόδειγμα ήταν οπτικό, δηλαδή μια λέξη με διαφορετικό κάθε φορά μέγεθος, βάρος ή και στυλ γραμματοσειράς, ενώ τα ερεθίσματα επιλογής ήταν για κάθε οπτικό υπόδειγμα τρεις διαφορετικές ακουστικές αποδόσεις του.

Τα αποτελέσματα έδειξαν σημαντικές διαφορές μεταξύ τυφλών και βλέπόντων [ANOVA μικτού σχεδιασμού $F(1,60) = 8.345$ $p < .05$], με τους βλέποντες να παρουσιάζουν καλύτερες επιδόσεις, όπως φαίνεται στον παρακάτω πίνακα:

Πίνακας 1. Δείκτες κεντρικής τάσης και διασποράς για τη μεταβλητή «κατάσταση όρασης»

ΥΠΟΔΕΙΓΜΑ	ΣΥΜΜΕΤΕΧΟΝΤΕΣ	Mean	%	Std. Deviation	N
ΑΚΟΥΣΤΙΚΟ	ΤΥΦΛΟΙ	12,2903	68,28%	2,91197	31
	ΒΛΕΠΟΝΤΕΣ	13,8065	76,70%	2,56150	31
	Total	13,0484	72,4%	2,82511	62
ΟΠΤΙΚΟ	ΤΥΦΛΟΙ	14,2581	79,2%	3,76800	31
	ΒΛΕΠΟΝΤΕΣ	16,4516	91,39%	1,78585	31
	Total	15,3548	85,3%	3,12629	62

Οι μεταβλητές «φύλο» και «βαθμίδα εκπαίδευσης» δεν φάνηκαν να επηρεάζουν σημαντικά τα ποσοστά των συνεπών απαντήσεων.

Παρά τις σημαντικές διαφορές μεταξύ των τυφλών και των βλέπόντων, οι μέσοι όροι των συνεπών απαντήσεων και των δυο ομάδων ήταν αρκετά υψηλοί (Τυφλοί: 72,4% , Βλέποντες: 85,3%) ώστε να συμπεράνουμε ότι η αρχική μας υπόθεση επιβεβαιώνεται Έτσι, λοιπόν, θεωρούμε ότι κατά την ακουστική απόδοση δομημένων εγγράφων με τη χρήση συνθετικής ομιλίας, η αντιστοίχιση αυτή είναι αντιληπτικά αποδεκτή, δεδομένου ότι οι επιλογές έγιναν χωρίς οι συμμετέχοντες να γνωρίζουν από πριν ποιες απαντήσεις ήταν συνεπείς με την αρχική μας υπόθεση.

References

- Blattner, M.M., D. A. Sumikawa, and R. M. Greenberg. *Earcons and icons: Their structure and common design principles*. E. P Glinert, ed. Visual Programming Environments: Applications and Issues, 582-606, IEEE Computer Society Press, Los Alamitos, CA, 1990.
- Brewster, S.A., P. C. Wright, and A.D.N. Edwards. *An evaluation of earcons for use in auditory human-computer interfaces*. INTERCHI '93 Conference Proceedings, 222-227, Amsterdam, April 1993.
- Brooks, L.W., D.F. Dansereau. Effects of structural schema training and text organization on expository prose processing. *Journal of Educational Psychology*, (75)6: 811-820, 1983.

- Fellbaum, K., and Kouroupetroglou, G. (2008). Principles of Electronic Speech Processing with Applications for People with Disabilities. *Technology and Disability*, 20(2): 55-85.
- Freitas, D., and G. Kouroupetroglou (2008). Speech Technologies for Blind and Low Vision Persons", *Technology and Disability*, 20(2): 135-156.
- Gaver, W. Everyday Listening and Auditory Icons, Unpublished Doctoral dissertation, University of California, 1988.
- Gaver, W.W. Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction*, 2: 167- 177, 1986.
- Gaver, W.W. *The SonicFinder: An interface that uses auditory icons*. E. P. Glinert, ed. Visual Programming Environments: Applications and Issues, 561-581, IEEE Computer Society Press, Los Alamitos, CA, 1990.
- Hartley, J., J. Kenely, G. Owen, M. Trueman. The Effect of Headings on Children's Recall From Prose Text. *British Journal of Educational Psychology*, 40: 304-307, 1980.
- JAWS for Windows (<http://www.freedomscientific.com/>)
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363–394.
- Kouroupetroglou, G & Tsonos, D, (2008). Multimodal Accessibility of Documents. chapter in the book *Advances in Human-Computer Interaction*, I-Tech Education and Publishing, Vienna, 451-470
- Lorch, R. F. (1989), Text-Signaling Devices and Their Effects Reading and Memory Processes. *Educational Psychology Review*, Vol. 1, No. 3
- Lorch, R.F., E.P. Lorch. Effects of Headings on Text Recall and Summarization. *Contemporary Educational Psychology*. 21: 261-278, 1996b.
- Lorch, R.F., E.P. Lorch. Effects of Organizational Signals on Free Recall of Expository Text. *Journal of Educational Psychology*. 88(1): 38-48, 1996a.
- Mynatt, E.D. and W.K. Edwards. (1992). *Mapping GUIs to auditory interfaces*. *Proceedings of the ACM: UIST '92*, 61-70, New York.
- Olhausen, M. M., CM. Roller. The Operation of Text Structure and Content Schemata in Isolation and n Interaction. *Reading Research Quarterly*, (23)1: 70-87, 1988.

- Sanchez, R.P., E.P. Lorch, R.F. Lorch. Effects of Headings on Text Processing Strategies. *Contemporary Educational Psychology*, 26: 418-428, 2001.
- Spyridakis, J.H. Signaling effects: A Review of the Research, part I. *Journal of Technical Writing and Communication*, 19(3): 227-240, 1989a.
- SUPERNOVA Reader Magnifier (<http://www.dolphinuk.co.uk/products/hal.htm>)
- Tsonos, D., G. Xydas, and G. Kouroupetroglou: "Auditory Accessibility of Metadata in Books: A Design for All Approach", *Lecture Notes in Computer Science (LNCS)* Vol. 4556, pp. 436-445, 2007
- Xydas, G., Argyropoulos, V., Karakosta, T., & Kouroupetroglou, G. (2004). *An Open Platform for Conducting Psycho-Acoustic Experiments in the Auditory Representation of Web Documents. Proceedings of the Conference "ACOUSTICS 2004"*, Thessaloniki, 27-28 September, 157-164.
- Xydas, G., Argyropoulos, V., Karakosta, T., & Kouroupetroglou, G. (2005). An Experimental Approach in Recognizing Synthesized Auditory Components in a Non-Visual Interaction with Documents. *Proc. of the 11th Int. Conference on Human-Computer Interaction*, 22-27 July 2005, Las Vegas, Nevada USA, Vol. 3, 411-420.