



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ**

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

**Μοντελοποίηση της ακουστικής αναπαράστασης της
τυπογραφίας εγγράφων μέσω εκφραστικής συνθετικής
ομιλίας για τυφλούς και βλέποντες**

Δημήτριος Ι. Τσώνος

ΑΘΗΝΑ

Φεβρουάριος 2012

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Μοντελοποίηση της ακουστικής αναπαράστασης της τυπογραφίας εγγράφων μέσω εκφραστικής συνθετικής ομιλίας για τυφλούς και βλέποντες

Δημήτριος Ι. Τσώνος

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ: Γεώργιος Θ. Κουρουπέτρογλου,
Αναπληρωτής Καθηγητής ΕΚΠΑ

ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ ΠΑΡΑΚΟΛΟΥΘΗΣΗΣ:

Γεώργιος Κουρουπέτρογλου, Αναπληρωτής Καθηγητής ΕΚΠΑ
Στέλλα Βοσνιάδου, Καθηγήτρια ΕΚΠΑ
Σέργιος Θεοδωρίδης, Καθηγητής ΕΚΠΑ

ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Γεώργιος Κουρουπέτρογλου,
Αναπληρωτής Καθηγητής ΕΚΠΑ

Στέλλα Βοσνιάδου,
Καθηγήτρια ΕΚΠΑ

Σέργιος Θεοδωρίδης,
Καθηγητής ΕΚΠΑ

Εμμανουήλ Σαγκριώτης,
Αναπληρωτής Καθηγητής ΕΚΠΑ

Κωνσταντίνος Παπαδόπουλος,
Αναπληρωτής Καθηγητής
Πανεπιστήμιο Μακεδονίας

Βασίλειος Αργυρόπουλος,
Επίκουρος Καθηγητής
Πανεπιστήμιο Θεσσαλίας

Βασίλειος Γάτος,
Ερευνητής Β
ΕΚΕΦΕ ΔΗΜΟΚΡΙΤΟΣ

Ημερομηνία εξέτασης 29/2/2012

ΠΕΡΙΛΗΨΗ

Η δυνατότητα πρόσβασης στα έντυπα και ηλεκτρονικά έγγραφα (βιβλία, εφημερίδες, περιοδικά, επιστολές, περιεχόμενο ιστού) από τα άτομα με εντυποαναπηρία, αλλά και από τους μετακινούμενους χρήστες ή τους ηλικιωμένους, βασίζεται στη μετατροπή τους (σε πραγματικό χρόνο) σε ακουστική, κατά κύριο λόγο, ή και απτική μορφή. Ένα έγγραφο, πέραν από το περιεχόμενό του, περιλαμβάνει έναν αριθμό από στοιχεία οπτικής παρουσίασης, όπως: είδος, μέγεθος και χρώμα γραμματοσειράς και τρόπο γραφής (έντονη, πλάγια, υπογραμμισμένη). Παρά τη μεγάλη πρόοδο που επιτεύχθηκε στα συστήματα μετατροπής Κειμένου-σε-Ομιλία, σήμερα αυτά δεν υποστηρίζουν την αποτελεσματική ακουστικοποίηση της σημασιολογίας και των γνωσιακών πτυχών των Στοιχείων Οπτικής Παρουσίασης των Εγγράφων (ΣΟΠΕ). Ουσιαστικά αυτή η επιπλέον πληροφορία «χάνεται» κατά την επεξεργασία του εγγράφου με σκοπό την ακουστική απόδοσή του ή τη μετατροπή του σε απτική μορφή (γραφή Braille).

Η παρούσα διατριβή ασχολείται με την ακουστικοποίηση, για τυφλούς και βλέποντες, της μεταπληροφορίας ΣΟΠΕ κατά τη μετατροπή τους σε ομιλία. Η προσέγγιση του προβλήματος χωρίζεται σε δύο στάδια: α) την αυτόματη εξαγωγή των επαγόμενων συναισθημάτων του αναγνώστη από τις εναλλαγές των ΣΟΠΕ και β) την ακουστική απόδοσή τους χρησιμοποιώντας εκφραστική συναισθηματική συνθετική ομιλία.

Με σκοπό τη δημιουργία ενός συστήματος αυτόματης εξαγωγής της επαγόμενης συναισθηματικής κατάστασης από τα ΣΟΠΕ και τη κατάλληλη επισημείωση των εγγράφων με αυτή την πληροφορία, προτείνεται μία αρχιτεκτονική για την πολυτροπική παραγωγή, παρουσίαση και πλοήγηση σε πραγματικό χρόνο σε καθολικά προσβάσιμα έγγραφα, ανεξάρτητα φυσικής γλώσσας, περιεχομένου και πολιτισμού. Η υλοποίηση του αντίστοιχου πρωτότυπου συστήματος βασίστηκε στην τεχνολογία XML. Στη συνέχεια, αναπτύσσεται ένα μοντέλο για την ποσοτική περιγραφή του τρόπου ακουστικοποίησης των τυπογραφικών εναλλαγών ΣΟΠΕ μέσω: i) της μαθηματικής περιγραφής των μεταβολών της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη που στηρίζεται στη διαστατική φύση των συναισθημάτων στον χώρο «Ευχαρίστηση», «Διέγερση» και «Επικράτηση», και ii) της αντιστοίχισής τους σε μεταβολές προσωδιακών χαρακτηριστικών της εκφραστικής συνθετικής ομιλίας.

Για την αξιολόγηση του προσωδιακού μοντέλου εξετάστηκε, μέσω ψυχοακουστικών πειραμάτων, το κατά πόσο οι ακροατές είναι σε θέση να αναγνωρίσουν τις τυπογραφικές εναλλαγές ΣΟΠΕ με ακουστικό τρόπο. Τα αποτελέσματα ήσαν θετικά ακόμη και για ακροατές χωρίς πρότερη εκπαίδευση. Επίσης, το μοντέλο που αναπτύχθηκε αξιολογήθηκε από βλέποντες και τυφλούς μαθητές δημοτικού σχολείου ότι προάγει την απόδοσή τους κατά τη διδακτική διαδικασία.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Αλληλεπίδραση Ανθρώπου-Υπολογιστή

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Αλληλεπίδραση Ανθρώπου-Υπολογιστή, Καθολική Πρόσβαση, Σχεδίαση για Όλους, Συναισθήματα, Εκφραστική Συνθετική Ομιλία

ABSTRACT

The accessibility to printed and electronic documents (books, newspapers, journals, letters, web content) by the print disabled, as well as the moving users and the elderly, is based on the possibility to convert them (in real time) into, primary, acoustic and or haptic form. Besides its content, a printed or electronic text document contains a number of presentation visual elements that apply design glyphs or typographic elements, such as font (type, size and color) and font style (bold, italics and underline). Regardless the important progress achieved in Text to Speech systems, to date they do not support the efficient sonification of the semantics and cognitive aspects of the Visual Presentation Elements in Documents (VPED). Essentially all this additional metadata information vanishes during the document processing towards its acoustic or haptic (in Braille) rendition.

This dissertation deals with the sonification of the VPED metadata during their transformation to speech. The approach to this problem includes two phases: a) the automatic extraction of the VPED induced emotional states to the reader and b) their acoustic rendition using expressive emotional synthetic speech.

With the scope to develop a system for the automatic extraction of the VPED induced emotional states and the appropriate tagging of the documents with this information, a novel architecture is proposed for the multimodal universal accessibility of documents, regardless of their natural language, content and culture. The realization of the corresponding system is based on the XML technology. Moreover, a quantitative model is developed for the sonification of the VPED typographic alternations by: i) the mathematical formulation of the induced emotional state of the reader, based on the dimensional nature of the emotions in the space "Pleasure", "Arousal" and "Dominance", and ii) their mapping into alternation of the prosodic characteristics of the expressive synthetic speech.

For the evaluation of the prosodic model we have explored, using psychoacoustic experiments, whether the listeners can acoustically recognize the VPED typographic alternations. The results were positive even in the case of listeners without any previous training. Furthermore, the evaluation of the developed model by sighted and blind students of primary education shows enhancement of their performance during the didactic process.

SUBJECT AREA: Human Computer-Interaction

KEYWORDS: Human Computer-Interaction, Universal Accessibility, Design-for-All, Emotions, Expressive Speech Synthesis

Στην Ασημίνα

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω τον επιβλέποντά μου, καθηγητή κ. Γεώργιο Κουρουπέτρογλου, που μου έδωσε την ευκαιρία να εκπονήσω αυτή την διατριβή, καθώς και για την συνεχή συμπαράσταση, καθοδήγηση και κατανόηση του σε αυτή τη δύσκολη αλλά δημιουργική και εποικοδομητική πορεία.

Να ευχαριστήσω την υποψήφια διδάκτορα Καλλιόπη Εικοσπεντάκη για την βοήθεια της και τις συμβουλές καθ' όλη την διάρκεια της διατριβής. Η συμβολή της ήταν πολύτιμη για την κατανόηση του τρόπου προσέγγισης των προβλημάτων που σχετίζονται με την Γνωσιακή Ψυχολογία. Η βοήθεια του Δρ. Γεράσιμου Ξύδα ήταν καθοριστική με τις συμβουλές του στα πρώτα βήματα της διατριβής μου, καθώς και την βοήθεια του για την χρήση του συνθέτη ομιλίας ΔΗΜΟΣΘΕΝΗ και της πλατφόρμας Document-to-Audio. Επίσης, να ευχαριστήσω τον Δρ. Δημήτρη Σπηλιωτόπουλο για την βοήθεια του στα ακουστικά πειράματα, τον υποψήφιο διδάκτορα Φίλιππο Κατσούλη, για την συνεργασία μας και τις συζητήσεις μας ως προς την προσέγγιση του προβλήματος της ακουστικοποίησης των εγγράφων και τη Χερνίσα Κατσόρη για την συνεργασία μας και την ανταλλαγή γνώσεων.

Μέσα από την εκπόνηση της διατριβής και της συμμετοχής μου στις ερευνητικές δραστηριότητες του "Εργαστηρίου Φωνής και Προσβασιμότητας", δημιουργήθηκε μια μακροχρόνια φιλία με τον Δρ. Αθανάσιο Περπερή και τον υποψήφιο διδάκτορα Παναγιώτη Τζεβελέκο.

Θα ήθελα να ευχαριστήσω τους γονείς μου Γιάννη και Κωνσταντία καθώς και την αδερφή μου Αλεξάνδρα, για την στήριξη τους στην επιλογή μου για την εκπόνηση της διδακτορικής διατριβής και καθ' όλη την διάρκεια της.

Τέλος, να ευχαριστήσω την σύζυγο μου Ασημίνα, για την υπομονή και την συνεχή υποστήριξή της στις επιλογές μου.

Η παρούσα διδακτορική διατριβή εκπονήθηκε στο πλαίσιο του Προγράμματος Μεταπτυχιακών Σπουδών του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών (Ε.Κ.Π.Α.). Χρηματοδοτήθηκε από την Γενική Γραμματεία Έρευνας και Τεχνολογίας και την Ευρωπαϊκή Ένωση-Ευρωπαϊκό Κοινωνικό Ταμείο στο πλαίσιο του έργου «ΟΜΗΡΟΣ: Μεθοδολογική προσέγγιση για τη Σχεδίαση Ακουστικής - Απτικής Αλληλεπίδρασης στη Μη-Οπτική Διεπαφή Χρήστη (User Interface) με έμφαση στην Προσβασιμότητα Ατόμων με Απώλεια Όρασης», Πρόγραμμα Ενίσχυσης του Ερευνητικού Δυναμικού (ΠΕΝΕΔ) – 2003.

Δημήτριος Ι. Τσώνος
Φεβρουάριος 2012

ΠΕΡΙΕΧΟΜΕΝΑ

1. ΕΠΙΣΚΟΠΗΣΗ ΤΟΥ ΕΠΙΣΤΗΜΟΝΙΚΟΥ ΠΕΔΙΟΥ	27
1.1 Εισαγωγή.....	27
1.2 Η έννοια του «εγγράφου».....	27
1.3 Η αρχιτεκτονική των κειμενικών εγγράφων.....	29
1.4 Προσβασιμότητα και έγγραφα.....	30
1.4.1 Κατευθυντήριες γραμμές για την Προσβασιμότητα του Ιστού	35
1.4.2 Open Document Format Accessibility	39
1.4.3 Math Markup Language.....	41
1.4.4 Braille Markup Language.....	42
1.4.5 Scalar Vector Graphics.....	42
1.4.6 DAISY/NISO	43
1.5 Συναισθήματα	44
1.5.1 Η έννοια της λέξης «συναίσθημα»	46
1.5.2 Πειραματικές διαδικασίες για τη μελέτη των συναισθημάτων	48
1.5.3 Δοκιμασία αυτοαξιολόγησης με τη χρήση φιγούρων - Self Assessment Manikin Test	50
1.6 Συναισθήματα και έγγραφα.....	53
1.6.1 Επαγόμενα συναισθήματα του αναγνώστη από το περιεχόμενο του εγγράφου	54
1.6.2 Επαγόμενα συναισθήματα του αναγνώστη από τον τρόπο παρουσίασης του εγγράφου	55
1.7 Εκφραστική Συνθετική Ομιλία.....	57
1.7.1 Διαστατική προσέγγιση της εκφραστικής συνθετικής ομιλίας.....	59
1.7.2 Φωνή και συναισθήματα: Οι παράγοντες της γλώσσας και του πολιτισμού	62
1.8 Στόχος Διατριβής.....	63
2. ΠΟΛΥΤΡΟΠΙΚΗ ΠΡΟΣΒΑΣΗ ΣΤΑ ΕΓΓΡΑΦΑ	67
2.1 Εισαγωγή.....	67
2.2 Η προτεινόμενη αρχιτεκτονική για πολυτροπική πρόσβαση στα έγγραφα.....	69
2.3 Παραγωγή πολυτροπικά προσβάσιμων εγγράφων.....	70
2.4 Προ-επεξεργασία εγγράφων.....	70
2.5 Ανάλυση των εγγράφων	71
2.5.1 Σημασιολογικό Επίπεδο	71
2.5.2 Συναισθηματικό Επίπεδο.....	72

2.6	Πολυτροπική πλοήγηση στα προσβάσιμα έγγραφα.....	72
2.7	Πολυτροπική παρουσίαση των προσβάσιμων εγγράφων	74
2.8	Ακουστική παρουσίαση των πολυτροπικά προσβάσιμων εγγράφων	75
2.9	Εφαρμογή της προτεινόμενης αρχιτεκτονικής κατά την διδακτική διαδικασία	80
3.	ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΕΠΑΓΩΜΕΝΗΣ ΣΥΝΑΙΣΘΗΜΑΤΙΚΗΣ ΚΑΤΑΣΤΑΣΗΣ ΤΟΥ	
	ΑΝΑΓΝΩΣΤΗ	83
3.1	Εισαγωγή.....	83
3.2	Πιλοτικό πείραμα.....	83
3.2.1	Μεθοδολογία.....	84
3.2.2	Πειραματική διαδικασία.....	86
3.2.3	Αποτελέσματα.....	87
3.2.4	Συμπεράσματα.....	91
3.3	Πειραματική διαδικασία.....	93
3.3.1	Ερευνητικό Μοντέλο - Υπόθεση.....	93
3.3.2	Μεθοδολογία.....	98
3.3.3	Πειραματική διαδικασία.....	100
3.3.4	Αποτελέσματα.....	101
3.3.5	Μαθηματική μοντελοποίηση των αποτελεσμάτων	112
3.3.6	Συμπεράσματα.....	120
3.4	Το σύστημα DocEmoX	122
3.4.1	Η προτεινόμενη αρχιτεκτονική	122
3.4.2	Προ-επεξεργασία των εγγράφων	122
3.4.3	Εξαγωγή της επαγόμενης συναισθηματικής κατάσταση του αναγνώστη και επισημείωση των εγγράφων με την αντίστοιχη πληροφορία.....	125
3.4.4	Υλοποίηση και αξιολόγηση του συστήματος	128
4.	ΑΚΟΥΣΤΙΚΟΠΟΙΗΣΗ ΤΩΝ ΤΥΠΟΓΡΑΦΙΚΩΝ ΣΤΟΙΧΕΙΩΝ ΜΕ ΤΗΝ ΧΡΗΣΗ	
	ΕΚΦΡΑΣΤΙΚΗΣ ΣΥΝΘΕΤΙΚΗΣ ΟΜΙΛΙΑΣ.....	131
4.1	Εισαγωγή.....	131
4.2	Το προσωδιακό μοντέλο.....	131
4.2.1	Ελάχιστες προσωδιακές μεταβολές της ομιλίας που είναι αντιληπτές από τον ακροατή.....	136
4.2.2	Γραμμική κβάντιση των προσωδιακών μεταβολών	137
4.2.3	Προσωδιακές μεταβολές μετά τη γραμμική κβάντιση.....	139

4.3	Επιλογή βέλτιστης ακουστικής απόδοσης των τυπογραφικών μεταβολών.....	142
4.3.1	Η πειραματική διαδικασία	143
4.3.2	Αποτελέσματα.....	144
4.3.3	Σύνοψη των αποτελεσμάτων	144
4.4	Συμπεράσματα	146
5.	ΑΞΙΟΛΟΓΗΣΗ ΤΟΥ ΑΚΟΥΣΤΙΚΟΥ ΜΟΝΤΕΛΟΥ	149
5.1	Εισαγωγή.....	149
5.2	Ακουστική αναγνώριση των τυπογραφικών στοιχείων	149
5.2.1	Η μεθοδολογία	149
5.2.2	Η πειραματική διαδικασία	150
5.2.3	Τα αποτελέσματα	150
5.3	Εφαρμογή του ακουστικού μοντέλου στη διδακτική διαδικασία για τυφλούς και βλέποντες μαθητές.....	154
5.3.1	Μεθοδολογία.....	154
5.3.2	Πειραματική διαδικασία.....	155
5.3.3	Αποτελέσματα.....	156
5.4	Συμπεράσματα	156
6.	ΣΥΜΠΕΡΑΣΜΑΤΑ – ΜΕΛΛΟΝΤΙΚΗ ΕΡΕΥΝΑ	159
6.1	Συμπεράσματα	159
6.2	Περιορισμοί.....	162
6.3	Μελλοντική έρευνα	163
	ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΑΣ	165
	ΣΥΝΤΜΗΣΕΙΣ – ΑΡΚΤΙΚΟΛΕΞΑ – ΑΚΡΩΝΥΜΙΑ	167
	ΑΝΑΦΟΡΕΣ	169

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

Σχήμα 1: Το βασικό μοντέλο συσχέτισης των συστατικών, των τεχνικών χαρακτηριστικών και των οδηγιών WAI της W3C για προσβασιμότητα στο περιεχόμενο του Ιστού.....	36
Σχήμα 2: Οι φιγούρες (manikins) της κλίμακας εννέα σημείων της δοκιμασίας SAM, όπως παρουσιάζονται κατά την πειραματική διαδικασία. Οι εκφράσεις «Ευχαρίστηση», «Διέγερση» και «Επικράτηση» δεν εμφανίζονται κατά το πείραμα.	52
Σχήμα 3: Η λεκτική – σημασιολογική έκφραση των συναισθημάτων τοποθετημένη στο πλέγμα «Ευχαρίστηση» - «Διέγερση».	52
Σχήμα 4: Δύο στιγμιότυπα από α) Geneva Emotion Wheel (GEW) και β) FeelTrace....	53
Σχήμα 5: Το διάγραμμα διαδικασίας μετατροπής Κειμένου-σε-Ομιλία.....	58
Σχήμα 6: Γενικό μοντέλο «αρχιτεκτονικής» κειμενικού εγγράφου.	67
Σχήμα 7: Το προτεινόμενο σύστημα για την πολυτροπική πρόσβαση στα έγγραφα.	69
Σχήμα 8: Το άρθρωμα Προ-Επεξεργασίας των Εγγράφων.	70
Σχήμα 9: Το άρθρωμα της Ανάλυσης των Εγγράφων: η προσέγγιση της εξαγωγής Σημασιολογικής Πληροφορίας.	72
Σχήμα 10: Το άρθρωμα της Ανάλυσης των Εγγράφων: η προσέγγιση της εξαγωγής Συναισθηματικής Πληροφορίας	73
Σχήμα 11: Το διάγραμμα της Πολυτροπικής Απεικόνισης των εγγράφων, με έμφαση στον ακουστικό τρόπο	76
Σχήμα 12: Η προσωδιακή και τυπογραφική όψη.....	78
Σχήμα 13: Η τροποποιημένη αρχιτεκτονική του προτεινόμενου συστήματος για πολυτροπική προσβασιμότητα. Η λειτουργικότητα του έχει προσαρμοστεί για τη χρήση του στη διδακτική διαδικασία.	81
Σχήμα 14: Ο πολυτροπικός προσαρμογέας τροποποιημένος έτσι ώστε να εξυπηρετεί και την ακουστική απόδοση των μαθηματικών εκφράσεων.	82
Σχήμα 15: Τα πειραματικά αποτελέσματα της συναισθηματικής κατάστασης α) «Ευχαρίστηση» β), «Διέγερση» (γ), «Επικράτηση» ως προς το μέγεθος της γραμματοσειράς.....	88

Σχήμα 16: Τα πειραματικά αποτελέσματα των διαστάσεων «Ευχαρίστησης» και «Διέγερσης» με α) τις τιμές κάθε ερεθίσματος και β) τις μέσες τιμές των ερεθισμάτων για κάθε χρωματικό συνδυασμό.	90
Σχήμα 17: Δύο δείγματα ερεθισμάτων όπως παρουσιάστηκαν κατά την πειραματική διαδικασία. (α) Ελληνικό κείμενο σε έντονη γραφή, «Times New Roman» με μέγεθος γραμματοσειράς 16px και (β) ελληνικό απλό κείμενο, «Times New Roman» με μέγεθος γραμματοσειράς 26px.	100
Σχήμα 18: Οι μέσες τιμές των απαντήσεων και τα αντίστοιχα τυπικά σφάλματα για όλα τα ερεθίσματα στους χρωματικούς συνδυασμούς για τον άξονα α) «Ευχαρίστηση» β) «Διέγερση» και γ) «Επικράτηση».....	106
Σχήμα 19: Το προτεινόμενο μοντέλο συσχετισμού των ανεξάρτητων μεταβλητών με τις τρεις διαστάσεις «Ευχαρίστηση», «Διέγερση» και «Επικράτηση».	112
Σχήμα 20: Οι μέσες τιμές των τριών συναισθηματικών διαστάσεων «Ευχαρίστηση», «Διέγερση» και «Επικράτηση» για κάθε χρωματικό συνδυασμό. Η κωδικοποίηση του άξονα X είναι, 1. Κίτρινο σε Μπλε, 2. Άσπρο σε Μαύρο, 3. Άσπρο σε Μπλε, 4. Κόκκινο σε Πράσινο, 5 Μαύρο σε Γκρι, 6. Μαύρο σε Άσπρο και 7. Πράσινο σε Κίτρινο.	113
Σχήμα 21: Η γραφική παράσταση της διάστασης (α) «Ευχαρίστηση», (β) «Διέγερση» και (γ) «Επικράτηση» ως προς το μέγεθος της γραμματοσειράς.....	116
Σχήμα 22: Το άρθρωμα της προ-επεξεργασίας εγγράφου του DocEmoX.	124
Σχήμα 23: Το δομικό στοιχείο «Heading», όπως περιγράφεται στο αρχείο styles.xml	125
Σχήμα 24: Το άρθρωμα του DocEmoX για την εξαγωγή και επισημείωση της συναισθηματικής κατάστασης.	125
Σχήμα 25: Η περιγραφή της συναισθηματικής κατάστασης σύμφωνα με την EmotionML χρησιμοποιώντας: (α) διπολικές τιμές από το -1 έως το 1, και (β) διπολικές διακριτές τιμές.	126
Σχήμα 26: Ένα παράδειγμα της δομής του αρχείου baseline.xml με τα βασικά τυπογραφικά στοιχεία	127
Σχήμα 27: Παράδειγμα XSL υλοποίησης ενός συναισθηματικού κανόνα	128
Σχήμα 28: Τα συναισθηματικά επισημειωμένα δομικά στοιχεία στο αρχείο styles.xml	128
Σχήμα 29: Παράδειγμα ερεθίσματος, όπως παρουσιαζόταν οπτικά κατά την πειραματική διαδικασία	150

Σχήμα 30: Βαθμός αναγνώρισης για την εναλλαγή του τρόπου γραφής.	153
Σχήμα 31: Βαθμός αναγνώρισης για τη μείωση/αύξηση του μεγέθους γραμματοσειράς.	153

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 1: Οι επτά χρωματικοί συνδυασμοί που χρησιμοποιήθηκαν	85
Πίνακας 2: Βαθμολογία αξιοπιστίας κάθε συμμετέχοντα για κάθε μία συναισθηματική διάσταση.....	92
Πίνακας 3: Οι χρωματικοί συνδυασμοί γραμματοσειράς / υποβάθρου που χρησιμοποιούνται στην παρούσα εργασία: (α) τα χρώματα σύμφωνα με τον χρωματικό χώρο αντίληψης του Munsell (perceptual color space) και (β) στον χρωματικό χώρο RGB μαζί με τις αντίστοιχες τιμές διαφοράς χρωματικής φωτεινότητας του ζεύγους χρωμάτων / χρωματικής αντίθεσης.....	95
Πίνακας 4: Τα αποτελέσματα της three-way r-m ANOVA για κάθε μία εξαρτημένη μεταβλητή (συναισθηματικές διαστάσεις) ως προς τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου, τύπου γραμματοσειράς και τρόπου γραφής μαζί με τις αλληλεπιδράσεις.....	107
Πίνακας 5: Τα αποτελέσματα της two-way r-m ANOVA για κάθε μία εξαρτημένη μεταβλητή (συναισθηματικές διαστάσεις) ως προς τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου και τον τρόπο γραφής, την αλληλεπίδραση τους για «Times New Roman» και «Arial».....	108
Πίνακας 6: Οι μέσες τιμές των τριών συναισθηματικών διαστάσεων ως προς το μέγεθος της γραμματοσειράς (χρησιμοποιώντας μαύρο κείμενο, «Times New Roman» σε άσπρο υπόβαθρο) μαζί με το τυπικό σφάλμα (SE).....	110
Πίνακας 7: Τα αποτελέσματα της two-way r-m ANOVA για κάθε μία εξαρτημένη μεταβλητή (συναισθηματικές διαστάσεις) ως προς τις εναλλαγές του τύπου γραμματοσειράς και του τρόπου γραφής καθώς και η αλληλεπίδραση αυτών των μεταβολών.....	111
Πίνακας 8: Οι μέσες τιμές της «Ευχαρίστησης» και οι ποσοστιαίες μεταβολές από το απλό κείμενο, για τον τρόπο γραφής χρησιμοποιώντας α) «Arial» και β) «Times New Roman».....	111
Πίνακας 9: Οι πολυωνυμικοί συντελεστές της μαθηματικής περιγραφής του συσχετισμού των συναισθηματικών διαστάσεων με το μέγεθος γραμματοσειράς και το τυπικό σφάλμα (SE).....	116
Πίνακας 10: Ο τρόπος επισημείωσης του μεγέθους γραμματοσειράς, του τύπου γραμματοσειράς και του τρόπου γραφής κατά ODF.....	124

Πίνακας 11: Μεγέθη γραμματοσειράς σε σημεία και η μετατροπή τους σε εικονοστοιχεία.	134
Πίνακας 12: Οι μεταβολές των συναισθηματικών διαστάσεων από τη βάση αναφοράς σε ποσοστιαία κλίμακα.	134
Πίνακας 13: Οι προσωδιακές μεταβολές για α) εναλλαγές του μεγέθους γραμματοσειράς για τρεις διαφορετικές βάσεις αναφοράς, 10pt, 12pt και 14 pt και β) εναλλαγές τρόπου γραφής από το απλό κείμενο, για κάθε τύπο γραμματοσειράς.	135
Πίνακας 14: Τα έξι επίπεδα γραμμικής κβάντισης των προσωδιακών μεταβολών για: α) το μέγεθος και β) τον τύπο γραμματοσειράς – τρόπο γραφής.	138
Πίνακας 15: Τα επίπεδα γραμμικής κβάντισης των προσωδιακών εναλλαγών όπως προκύπτουν για: α) το μέγεθος και β) τον τύπο γραμματοσειράς – τρόπο γραφής. ...	140
Πίνακας 16: Ελάχιστες και μέγιστες τιμές των προσωδιακών μεταβολών.....	141
Πίνακας 17: Οι προσωδιακές μεταβολές μετά την κβάντιση (ποσοστιαίες μεταβολές) και οι αντίστοιχες πραγματικές τιμές για την εναλλαγή του μεγέθους γραμματοσειράς	141
Πίνακας 18: Οι προσωδιακές μεταβολές μετά τη κβάντιση (ποσοστιαίες μεταβολές) και οι αντίστοιχες πραγματικές τιμές για την εναλλαγή του τρόπου γραφής.	142
Πίνακας 19: Τα στατιστικά αποτελέσματα της ανάλυσης των αποτελεσμάτων από την πειραματική διαδικασία. Στις παρενθέσεις αναφέρεται ο τύπος γραμματοσειράς (TNR – Times New Roman και Arial).	145
Πίνακας 20: Τα στατιστικώς σημαντικά αποτελέσματα μαζί με τα επίπεδα αύξησης/μείωσης, με τις αντίστοιχες ποσοστιαίες μεταβολές και τις απόλυτες τιμές των προσωδιακών χαρακτηριστικών	146
Πίνακας 21: Αποτελέσματα για την ακουστικοποίηση της εναλλαγής του τρόπου γραφής.....	152
Πίνακας 22: Αποτελέσματα για την ακουστικοποίηση της εναλλαγής του μεγέθους της γραμματοσειράς.....	152

ΠΡΟΛΟΓΟΣ

Τα κείμενα με τα οποία ερχόμαστε σε καθημερινή επαφή είναι σε έντυπη μορφή (π.χ. εφημερίδες, σχολικά εγχειρίδια) ή σε ηλεκτρονική μορφή (π.χ. doc, HTML, XML) τα οποία προβάλλονται σε μια οθόνη τερματικής συσκευής ή σε οθόνη προβολής (π.χ. σε μία διάλεξη ή σχολική τάξη). Τα περισσότερα από αυτά έχουν παραχθεί χωρίς καμιά πρόβλεψη για την απόδοσή τους σε ακουστική μορφή από συστήματα κειμένου σε ομιλία. Το γεγονός αυτό έχει ως αποτέλεσμα η οπτική πληροφορία (όπως για παράδειγμα έντονη ή πλάγια γραφή, μέγεθος γραμμάτων, είδος γραμματοσειράς, κλπ) που σχετίζεται με τον τρόπο παρουσίασης των εγγράφων, να αγνοείται από π.χ. τις εφαρμογές ανάγνωσης οθονών (screen readers) ή τα συστήματα μετατροπής Κειμένου-σε-Ομιλία ή τις οθόνες Braille για άτομα με απώλεια όρασης.

Τα οπτικά στοιχεία ενός εγγράφου μπορούν να «κρύβουν» σημασιολογικές - εννοιολογικές πληροφορίες για το κείμενο (π.χ. να υποδηλώνεται έμφαση ή σε άλλες περιπτώσεις μειωμένη σημαντικότητα) τα οποία ένας συνθέτης ομιλίας δεν τα λαμβάνει υπόψη του. Μέχρι σήμερα δεν υπάρχει ένας δομημένος τρόπος αντιμετώπισης του προβλήματος της ακουστικής απόδοσης της οπτικής πληροφορίας των εγγράφων. Η σχεδίαση της ακουστικής αλληλεπίδρασης γίνεται από τους σχεδιαστές της διεπαφής χρήστη με έναν ad hoc και προσωποπαγή τρόπο, με συχνό αποτέλεσμα μια αναποτελεσματική χρήση του ήχου - φωνής. Βασικός στόχος της παρούσας διατριβής είναι η μελέτη του τρόπου ακουστικοποίησης των τυπογραφικών στοιχείων των εγγράφων, με τη χρήση συνθετικής ομιλίας, με σκοπό τη συνεισφορά για τη βελτίωση της προσβασιμότητας των εγγράφων από τυφλούς και βλέποντες.

Η εργασία δομείται σε 6 κεφάλαια ως εξής:

Κεφάλαιο 1^ο

Γίνεται παρουσίαση μιας εκτενούς ανασκόπησης του επιστημονικού πεδίου που πραγματεύεται η διατριβή. Αναλύονται έννοιες και ορισμοί, τα οποία είναι χρήσιμα για την κατανόηση του προβλήματος, του τρόπου προσέγγισής του και την παρουσίαση της προτεινόμενης λύσης. Μέσω της ανάλυσης των παραπάνω γίνεται η επιλογή των βέλτιστων μεθοδολογιών. Τέλος, παρουσιάζονται αναλυτικά οι στόχοι και τα κίνητρα της διατριβής.

Κεφάλαιο 2^ο

Σε αυτό το κεφάλαιο εισάγεται ένα μοντέλο αρχιτεκτονικής εγγράφων και γίνεται η παρουσίαση ενός προτεινόμενου συστήματος για πολυτροπική πρόσβαση στα έγγραφα με έμφαση, τον ακουστικό τρόπο. Μέσω της θεωρίας των συναισθημάτων και της μαθηματικής μοντελοποίησής τους, προτείνεται μια ενιαία μαθηματική περιγραφή της αντιστοίχισης των τυπογραφικών εναλλαγών σε προσωδιακές μεταβολές.

Κεφάλαιο 3^ο

Γενικά μέχρι σήμερα, τα επαγόμενα συναισθήματα του αναγνώστη που πηγάζουν από τον τρόπο παρουσίασης του εγγράφου, έχουν μελετηθεί με τη χρήση ποιοτικών και όχι ποσοτικών χαρακτηριστικών. Σε αυτό το κεφάλαιο παρέχεται ο τρόπος με τον οποίο μπορεί να γίνει αυτή η ποσοτική μελέτη. Παρουσιάζεται ένα πιλοτικό πείραμα για την εξαγωγή των κανόνων αντιστοίχισης των τυπογραφικών στοιχείων στη συναισθηματική κατάσταση του αναγνώστη. Αυτό το πείραμα παρείχε τις κατευθυντήριες γραμμές για τη διεξαγωγή μίας εκτεταμένης πειραματικής διαδικασίας, ώστε να είναι στατιστικώς έγκυρη η μοντελοποίηση των επαγόμενων συναισθημάτων από τις τυπογραφικές μεταβολές. Στο τέλος του κεφαλαίου παρουσιάζεται ένα σύστημα με τον οποίο μπορεί να γίνει η υλοποίηση των κανόνων που προέκυψαν, με σκοπό την αυτοματοποιημένη διαδικασία της εξαγωγής της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη εγγράφων (σύστημα DocEmoX).

Κεφάλαιο 4^ο

Έχοντας πλέον την ποσοτικοποίηση της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη, που παρουσιάστηκε στο προηγούμενο κεφάλαιο, σε αυτό προτείνεται μια μεθοδολογία για τον τρόπο με τον οποίο μπορούν να ακουσικοποιηθούν οι τυπογραφικές εναλλαγές. Αυτό γίνεται με τη χρήση της συνθετικής συναισθηματικής ομιλίας. Παρουσιάζεται η διαδικασία ακουσικοποίησης των τυπογραφικών εναλλαγών και τα προβλήματα που ενέκυψαν, καθώς και ο τρόπος επίλυσής τους. Από αυτή, προκύπτει το μοντέλο αντιστοίχισης των εναλλαγών των τυπογραφικών στοιχείων σε μεταβολές των προσωδιακών χαρακτηριστικών (τονικό ύψος, ρυθμός και ένταση).

Κεφάλαιο 5^ο

Το προτεινόμενο ακουστικό μοντέλο αντιστοίχισης αξιολογείται μέσω δύο πειραματικών διαδικασιών. Η πρώτη οδήγησε σε συμπεράσματα για το κατά πόσο ο χρήστης μπορεί να διακρίνει τις εναλλαγές των τυπογραφικών στοιχείων με ακουστικό τρόπο. Το τελικό μοντέλο που προέκυψε, δοκιμάστηκε σε μία εκπαιδευτική διαδικασία, με τυφλούς και βλέποντες μαθητές, για την μελέτη του κατά πόσο βελτίωσε την απόδοσή τους.

Κεφάλαιο 6^ο

Στο τελευταίο κεφάλαιο παρουσιάζονται τα συμπεράσματα της παρούσας μελέτης και η συνεισφορά της. Επίσης, τίθενται νέα ερευνητικά ερωτήματα, τα οποία αποτελούν κίνητρο για μελλοντική μελέτη και βελτιστοποίηση ανοικτών θεμάτων που ανακύπτουν από την παρούσα εργασία.

Δημήτριος Ι. Τσώνος

Φεβρουάριος 2012

1. ΕΠΙΣΚΟΠΗΣΗ ΤΟΥ ΕΠΙΣΤΗΜΟΝΙΚΟΥ ΠΕΔΙΟΥ

1.1 Εισαγωγή

Σε αυτό το κεφάλαιο παρουσιάζονται βασικές έννοιες, οι οποίες είναι χρήσιμες για την κατανόηση του προβλήματος που πραγματεύεται η παρούσα διατριβή, καθώς και για τον τρόπο με τον οποίο προσεγγίζεται το πρόβλημα της ακουστικοποίησης των τυπογραφικών στοιχείων εγγράφων με τη χρήση εκφραστικής συνθετικής ομιλίας. Αναλύονται έννοιες όπως «έγγραφο – αρχιτεκτονική εγγράφου», «προσβασιμότητα - πολυτροπική πρόσβαση», «συναίσθημα – συναισθηματική κατάσταση», «εκφραστική συνθετική ομιλία». Περιγράφονται οι βασικές μεθοδολογίες μελέτης των συναισθημάτων και ο συσχετισμός των τελευταίων με τα έγγραφα και την ομιλία. Στο τέλος του κεφαλαίου προσδιορίζεται ο στόχος της παρούσας διατριβής.

1.2 Η έννοια του «εγγράφου»

Παραδοσιακά, ως «έγγραφο» (document) θεωρείται ένα αρχείο - σύνολο κειμένων. Ο Schamber [1] όρισε ως «έγγραφο» την οντότητα *«η οποία περιέχει δυναμικό, ευέλικτο, μη-γραμμικό περιεχόμενο, που παρουσιάζεται ως ένα σύνολο από συνδεδεμένα στοιχεία πληροφοριών, αποθηκευμένο σε ένα ή περισσότερα μέσα ή δικτυωμένες θέσεις, τα οποία έχουν δημιουργηθεί ή χρησιμοποιούνται από έναν ή περισσότερα ανεξάρτητα άτομα με σκοπό την εξυπηρέτηση κάποιας διεργασίας ή έργου»*. Ο Buckland [2] προσπαθεί να απαντήσει την ερώτηση *«τι είναι ένα έγγραφο;»* μέσα από μια συζήτηση για το που μπορεί να φτάσει η έννοια του «εγγράφου» και ποια είναι τα όρια της έννοιας *«τεκμηρίωση» (documentation)*.

Η πρόοδος της επιστήμης των υπολογιστών και της τεχνολογίας της πληροφορίας δημιούργησε καινούριες προοπτικές για την έννοια του «εγγράφου». Ο όρος *«ηλεκτρονικό έγγραφο»* μπορεί να διαφοροποιηθεί από εκείνη του *«έντυπου εγγράφου»* αποδίδοντας συγκεκριμένα χαρακτηριστικά στα έγγραφα [1]: ευκολία στο χειρισμό τους, διασυνδέσιμα εσωτερικά ή εξωτερικά, εύκολα μετατρέψιμα, εγγενώς εξερευνήσιμα, άμεσα μεταφερόμενα και απείρως αντιγράψιμα.

Ο όρος «έγγραφο» που χρησιμοποιείται σε αυτή τη διατριβή, αναφέρεται σε όλα τα είδη τυπωμένα ή ηλεκτρονικά, των οποίων το περιεχόμενο, στην πλειονότητά του είναι κείμενο, όπως οι εφημερίδες, τα βιβλία, τα περιοδικά, το εκπαιδευτικό υλικό, οι επιστολές, τα φυλλάδια, κλπ.

Ένα έγγραφο περιέχει στοιχεία που τακτοποιούν και δομούν το περιεχόμενο στη σελίδα ή ακόμα και σε όλο το έγγραφο. Για παράδειγμα, ο τίτλος ενός κεφαλαίου μπορεί να αναγνωριστεί τοποθετώντας τον στην κορυφή της σελίδας με μεγαλύτερου μεγέθους γράμματα σε σχέση με το κύριο μέρος του κειμένου (σώμα κειμένου). Επίσης, το ευρετήριο σελίδων στο τέλος ενός εγγράφου, συνδέει τον αναγνώστη σε συγκεκριμένα κομμάτια του.

Η *λειτουργικότητα ενός εγγράφου* μπορεί να διακριθεί στις εξής ενέργειες: *ξεφύλλισμα, αναζήτηση, πλοήγηση και ανάγνωση* [3]. Άλλες έννοιες συνδεδεμένες με τη λειτουργικότητα είναι η αναγνωσιμότητα, πόσο ευανάγνωστο είναι και η αισθητική του.

Ο όρος «ευανάγνωστο» (legibility) ορίζεται ως *το μέτρο του πόσο εύκολα μπορεί ο αναγνώστης να διακρίνει τα γράμματα σε ένα κείμενο με συγκεκριμένο τύπο γραμματοσειράς και τρόπο γραφής* και ο όρος «αναγνωσιμότητα» (readability) είναι *το μέτρο του πόσο εύκολα οι λέξεις, φράσεις ή κομμάτια του κειμένου μπορούν να αναγνωστούν*. Αυτές οι δύο μετρικές εισήχθησαν ως συγκριτικά μεγέθη μεταξύ έντυπων και ηλεκτρονικών εγγράφων κειμένου που προβάλλονται στην οθόνη του υπολογιστή, για την περιγραφή της τυπογραφικής σαφήνειας (typographic clarity) [4]. Και οι δύο έννοιες είναι στενά συνδεδεμένες με το είδος γραμματοσειράς και με το συνδυασμό του χρώματος γραμματοσειράς/υποβάθρου [5] [6]. Η αναγνωσιμότητα είναι περισσότερο συνδεδεμένη με τη συνολική δομή ή διάταξη της σελίδας, όπως μελετήθηκε για την αγγλική γλώσσα [7] [8], αλλά και για τα Δανέζικα, Σουηδικά, Νορβηγικά και Φιλανδικά [9] [10] [11]. Η επίδραση της δομής του εγγράφου στην κατανόηση κατά την ανάγνωση και το ξεφύλλισμα (reading comprehension και browsing) μελετήθηκε από τους Calisir et al. [12] για την Τουρκική γλώσσα. Οι Hill και Scharf [13] συμπέραναν ότι δεν υπάρχει χρωματικός συνδυασμός γραμματοσειράς/υποβάθρου, γραμματοσειράς ή τρόπος γραφής που οδηγεί σε μεγαλύτερο βαθμό αναγνωσιμότητας. Επίσης, η αισθητική παρουσίαση ενός εγγράφου παίζει σημαντικό ρόλο κατά τη διάρκεια της ανάγνωσης του, αλλά και στην προτίμηση του αναγνώστη [14] [15] [16].

Ο όρος «τρόπος γραφής» έχει διπλή σημασία. Μπορεί να αναφέρεται είτε στο περιεχόμενο του κειμένου είτε στον τρόπο παρουσίασης του. Για παράδειγμα, όταν αναφερόμαστε στο περιεχόμενο εννοούμε το λογοτεχνικό ύφος του συγγραφέα, ενώ για στην παρουσίαση του κειμένου, εννοούμε το στυλ της γραμματοσειράς, δηλαδή, την έντονη, πλάγια, έντονη-πλάγια και υπογραμμισμένη γραφή. Στην παρούσα διατριβή ο όρος «τρόπος γραφής» αναφέρεται **μόνο** στον τρόπο παρουσίασης του κειμένου.

Επίσης, να σημειωθεί ότι ο όρος «απλό κείμενο» χρησιμοποιείται ως *το κείμενο στο οποίο δεν έχει εφαρμοστεί κανένα τυπογραφικό στοιχείο*.

«Δυναμικές» έννοιες όπως η «πλοήγηση» και το «ξεφύλλισμα» είναι συνδεδεμένες με τη διαδικασία αλληλεπίδρασης του αναγνώστη με το έγγραφο. Η πλοήγηση στα ηλεκτρονικά έγγραφα είναι ένα σύνολο από οδηγίες-εντολές με στόχο τη δημιουργία της κατάλληλης ροής πληροφοριών ενός εγγράφου για κάθε τύπο συσκευής [17]. Ο παραδοσιακός τρόπος πλοήγησης είναι απλά «η μετακίνηση ή κύλιση» σε άλλη θέση στο παράθυρο (scrolling the window) όπου παρουσιάζεται στον αναγνώστη το έγγραφο - αντίστοιχα, όπως το ξεφύλλισμα στα έντυπα έγγραφα. Η ψηφιοποίηση και η μετατροπή σε ηλεκτρονική μορφή κάθε έντυπου εγγράφου, μας έχει οδηγήσει στην εξεύρεση και χρήση νέων τεχνικών για αποτελεσματικότερη διαδικασία πλοήγησης, συνδυάζοντάς την με το ξεφύλλισμα [18] [19] [20]. Έτσι η λειτουργικότητα ενός εγγράφου μπορεί να διακριθεί σε δύο διεργασίες [21]:

- **Διεργασία Παρουσίασης (Presentation task)**, ως έξοδος του εγγράφου, για παράδειγμα: η παρουσίαση του περιεχομένου και των πληροφοριών του εγγράφου στον αναγνώστη.
- **Διεργασία Πλοήγησης (Navigation task)**, ως είσοδος στο έγγραφο, για παράδειγμα: οι ενέργειες που εκτελεί ο αναγνώστης όπως, η αναζήτηση ή το ξεφύλλισμα για την ανεύρεση συγκεκριμένης πληροφορίας στο έγγραφο.

1.3 Η αρχιτεκτονική των κειμενικών εγγράφων

Σύμφωνα με τον Marshall McLuhan, ένα έγγραφο είναι το μέσο δια του οποίου μεταδίδεται-επικοινωνείται ένα μήνυμα-πληροφορία [22]. Πέρα από το περιεχόμενο, ένα έγγραφο συμπεριλαμβάνει στοιχεία παρουσίασης τα οποία εφαρμόζονται με τη χρήση της τυπογραφίας (για παράδειγμα, η οπτική παρουσίαση ενός γράμματος ή χαρακτήρα με συγκεκριμένο είδος γραμματοσειράς και είδος γραφής) και τη διάθρωση-διάταξη του περιεχομένου στη σελίδα ή ακόμα και σε όλο το έγγραφο. Ο όρος «σήμα» (signal) εισάγεται ως «ο μηχανισμός γραφής που δίνει έμφαση σε πτυχές του περιεχομένου ή τη δομή του κειμένου χωρίς την προσθήκη στο περιεχόμενο του κειμένου» [23]. Στοχεύει στο να προαναγγείλει ή να τονίσει το περιεχόμενο ή/και να αποκαλύψει τις σχέσεις του περιεχομένου [24] [25]. Σύμφωνα με αυτή τη θεώρηση, για παράδειγμα ο τίτλος-υπότιτλος και τα τυπογραφικά στοιχεία νοούνται ως σήματα. Επίσης, η «επαύξηση εισόδου» (input enhancement) είναι η λειτουργία όπου η προβολή των γλωσσικών χαρακτηριστικών ενισχύεται μέσω της π.χ. οπτικής επαύξησης (π.χ. έντονη γραφή) και

των φωνολογικών χειρισμών για τον ακουστικό τρόπο (όπως είναι η προφορική επανάληψη) [26]. Τα στοιχεία παρουσίασης ενός εγγράφου (ή σήματα) μπορούν να κατηγοριοποιηθούν σε τρία επίπεδα (views ή layers) σύμφωνα με το γενικό «αρχιτεκτονικό» μοντέλο του εγγράφου που προτείνουμε στο Κεφάλαιο 2 [27]: Λογικό Επίπεδο (Logical Layer), Επίπεδο Διάταξης (Layout Layer), Τυπογραφικό Επίπεδο (Typography Layer).

Ως μεταδεδομένα (metadata) γενικά ορίζονται: οι δομημένες πληροφορίες που περιγράφουν, εξηγούν, προσδιορίζουν ή διευκολύνουν στην ανάκληση, χρήση ή διαχείριση ενός πόρου πληροφοριών. Συχνά τα μεταδεδομένα χαρακτηρίζονται ως «δεδομένα για τα δεδομένα» ή «πληροφορίες για τις πληροφορίες» [28]. Έτσι, οι πληροφορίες που περιλαμβάνονται στα τρία επίπεδα του «αρχιτεκτονικού» μοντέλου ενός εγγράφου χαρακτηρίζονται ως μεταδεδομένα, που αφορούν την οπτική παρουσίαση του εγγράφου. Στην παρούσα εργασία στο εξής θα αναφέρονται ως «μεταδεδομένα» του εγγράφου.

Τα μεταδεδομένα, ανάλογα τον τύπο του ηλεκτρονικού εγγράφου (το πρότυπο που ακολουθεί), κωδικοποιούνται με διαφορετικό τρόπο. Για παράδειγμα, μία φράση που είναι σε έντονη γραφή, σύμφωνα με το πρότυπο της HTML, συμβολίζεται με το στοιχείο . Κατά το πρότυπο ODF συμβολίζεται με <fo:font-weight="bold">. Στο εδάφιο 3.4.2 παρουσιάζεται αναλυτικά η κωδικοποίηση των τυπογραφικών στοιχείων που χρησιμοποιούνται στην παρούσα εργασία, σύμφωνα με το πρότυπο ODF (Πίνακας 10).

1.4 Προσβασιμότητα και έγγραφα.

Η εντυποαναπηρία εμποδίζει τους ανθρώπους από την ανάγνωση έντυπων εγγράφων και όχι μόνο. Αυτού του είδους η αναπηρία μπορεί να οφείλεται σε οπτική αναπηρία, αναπηρία αντίληψης ή φυσική αναπηρία λόγω οπτικής ανικανότητας, όπως τύφλωση, χαμηλή όραση ή δυσχρωματοψία, μαθησιακή αναπηρία – συμπεριλαμβανομένης και της δυσλεξίας ή αναπηρία η οποία εμποδίζει τη χρήση του έντυπου βιβλίου (όπως η απώλεια επιδεξιότητας). Δημογραφικές μελέτες που έχουν πραγματοποιηθεί για τους εντυποανάπηρους δείχνουν ότι ο αριθμός τους ποικίλλει, από το 10% του γενικού πληθυσμού στον Καναδά [29] έως το 17.5% στην Αυστραλία [30].

Οι απαιτήσεις των ανθρώπων με εντυποαναπηρία είναι η πρόσβαση στα έντυπα ή ηλεκτρονικά έγγραφα με ένα εναλλακτικό τρόπο, όπως π.χ. απτικά με χρήση της γραφής Braille, ακουστικά με ομιλία και μέσω μεγάλης-μεγενθυμένης εκτύπωσης ή προβολής τους σε οθόνη. Σε πολλές περιπτώσεις απαιτείται η χρήση υποστηρικτικών

τεχνολογιών για την εκπλήρωση της ανάγκης για πρόσβαση. Οι οθόνες Braille παρέχουν σε πραγματικό χρόνο τις πληροφορίες του κειμένου με απτικό τρόπο. Οι μεγεθυντές οθόνης (screen magnifiers) είναι εφαρμογές λογισμικού που βοηθούν τους ανθρώπους με χαμηλή όραση ή δυσχρωματοψία στην ανάγνωση ενός εγγράφου.

Τα συστήματα μετατροπής Κείμενο-σε-Ομιλία (Text-to-Speech, TtS) είναι εφαρμογές λογισμικού που μετατρέπουν σε πραγματικό χρόνο οποιοδήποτε ηλεκτρονικό κείμενο σε ομιλία [31]. Στις περισσότερες περιπτώσεις στη γραφική διεπαφή χρήστη εντοπίζεται το κείμενο από ένα λογισμικό που ονομάζεται «αναγνώστης οθόνης», που τροφοδοτεί το σύστημα μετατροπής Κειμένου-σε-Ομιλία. Τα συστήματα μετατροπής Κειμένου-σε-Ομιλία βρίσκουν εφαρμογή στους προσωπικούς υπολογιστές, αλλά και έξυπνες φορητές συσκευές και τηλέφωνα (όπως Smart Mobile Phones και Personal Digital Assistants - PDAs). Οι αυτόματες συσκευές ανάγνωσης (Automated Reading Devices - ARDs) είναι αυτόνομες συσκευές που μπορούν να μετατρέψουν έντυπα ή ηλεκτρονικά κείμενα σε φωνή, χρησιμοποιώντας συστήματα μετατροπής Κειμένου-σε-Ομιλία. Δεν προαπαιτούν τη σύνδεσή τους σε κάποια άλλη συσκευή, όπως για παράδειγμα έναν υπολογιστή. Τα ARDs έχουν σχεδιαστεί για τη χρήση από άτομα με εντυποαναπηρία. Ανάλογα με το είδος του πηγαίου υλικού (είσοδος συσκευής) υπάρχουν δύο κατηγορίες ARDs [32]:

- αυτόματες συσκευές ανάγνωσης για έντυπα κείμενα - Printed-Text ARDs,
- αυτόματες συσκευές ανάγνωσης για ηλεκτρονικά κείμενα - Electronic-Text ARDs.

Στα περισσότερα συστήματα μετατροπής Κειμένου-σε-Ομιλία, δε λαμβάνεται υπόψη οι σημασιολογικές και γνωσιακές πτυχές (semantics - cognitive aspects) των οπτικών και μη-οπτικών στοιχείων ενός εγγράφου (όπως το είδος της γραμματσειράς και ο τρόπος γραφής). Υπάρχουν κάποιες προσπάθειες για τη δημιουργία συστημάτων μετατροπής Εγγράφου-σε-Ακουστική μορφή (Document-to-Audio - DtA), τα οποία αποτελούν την επόμενη γενιά των συστημάτων μετατροπής Κειμένου-σε-Ομιλία, υποστηρίζοντας την εξαγωγή των σημασιών από τα μεταδεδομένα των εγγράφων [33] και την επαρκή ακουστική παρουσίαση-απόδοση της μορφής του κειμένου [34] [35] [36] και των πινάκων [37] [38] μοντελοποιώντας τις παραμέτρους της συνθετικής ομιλίας.

Σύμφωνα με την [39], «η **προσβασιμότητα** αφορά την πρόβλεψη και την υποστήριξη της πρόσβασης, από αναπήρους και ηλικιωμένα άτομα, στην κωδικοποιημένη

πληροφορία και στη διαπροσωπική επικοινωνία, μέσω κατάλληλων εφαρμογών βασισμένων στον ηλεκτρονικό υπολογιστή και στις τηλεματικές υπηρεσίες».

Επίσης, «*η καθολική πρόσβαση αφορά το δικαίωμα όλων των πολιτών για την απόκτηση και τη διατήρηση της πρόσβασης σε όλο το εύρος των πληροφοριακών πόρων της κοινωνίας και των εγκαταστάσεων για διαπροσωπική επικοινωνία, λαμβάνοντας υπόψη την ποικιλία των πλαισίων*» [40].

Παραδείγματα της ποικιλίας των πλαισίων αποτελούν περιπτώσεις όπου οι χρήστες:

- Ίσως δεν μπορούν εύκολα ή και καθόλου να δουν, να ακούσουν, να χειριστούν ή να επεξεργαστούν κάποιες μορφές πληροφορίας,
- Ίσως έχουν δυσκολία να διαβάσουν ή να κατανοήσουν κείμενα,
- Ίσως δεν μπορούν να χειριστούν το πληκτρολόγιο ή το ποντίκι
- Ίσως διαθέτουν μικρή οθόνη, ή οθόνη χωρίς γραφικά ή χαμηλής ταχύτητας σύνδεση στο διαδίκτυο
- Ίσως δεν μιλούν ή δεν καταλαβαίνουν με ευχέρεια τη γλώσσα στην οποία είναι γραμμένη μια πληροφορία
- Ίσως βρίσκονται σε μια κατάσταση όπου τα μάτια τους ή και τα χέρια τους ή και τα αυτιά τους είναι απασχολημένα σε άλλες κύριες δραστηριότητες (π.χ. ενώ οδηγούν, όταν εργάζονται σε ένα θορυβώδες περιβάλλον)
- Ίσως διαθέτουν μια παλαιότερη έκδοση ενός φυλλομετρητή, ή έναν εντελώς διαφορετικό φυλλομετρητή, ή έναν φωνητικό φυλλομετρητή ή ένα διαφορετικό λειτουργικό σύστημα.

Το δικαίωμα όλων των πολιτών για καθολική πρόσβαση έχει κατοχυρωθεί νομοθετικά σε διάφορα επίπεδα:

- Εθνικό, π.χ. Σύνταγμα της Ελλάδος, άρθρο 5Α(2) : «*Καθένας έχει δικαίωμα συμμετοχής στην Κοινωνία της Πληροφορίας. Η διευκόλυνση της πρόσβασης στις πληροφορίες που διακινούνται ηλεκτρονικά, καθώς και της παραγωγής, ανταλλαγής και διάδοσής τους αποτελεί υποχρέωση του Κράτους*».
- Ευρωπαϊκό, π.χ. Η Διακήρυξη της Ρήγα (11/6/2006) που υπεγράφη από τους ηγέτες των 25 κρατών μελών της Ευρωπαϊκής Ένωσης προωθεί μια Ευρωπαϊκή Κοινωνία της Πληροφορίας χωρίς κοινωνικό αποκλεισμό.

- Παγκόσμιο, π.χ. Συνθήκη των Ηνωμένων Εθνών για τα Δικαιώματα των ΑμεΑ (2008), Άρθρο 21: Ελευθερία έκφρασης και γνώμης, και πρόσβαση στην πληροφόρηση: *«Τα Κράτη Μέλη λαμβάνουν όλα τα κατάλληλα μέτρα για να ενθαρρύνουν τα ΜΜΕ, συμπεριλαμβανομένων των παρόχων υπηρεσιών μέσω του διαδικτύου, για να κάνουν τις υπηρεσίες τους προσβάσιμες στα ΑμεΑ».*

Τις τελευταίες δύο δεκαετίες, γίνονται πολλές προσπάθειες στην επιστημονική περιοχή της προσβασιμότητας των εγγράφων. Μερικές από αυτές εξετάζουν την προσβασιμότητα στο περιεχόμενο του παγκόσμιου ιστού για χρήστες με οπτική αναπηρία [41] [42] [43] [44]. Οι Bigham et al. [45] εξετάζουν πώς οι εικόνες μπορούν να είναι προσβάσιμες και οι Saito et al. [46] προτείνουν μια μέθοδο για τη μετατροπή του ήδη υπάρχοντος περιεχομένου υπό μορφή Flash, σε δομές τύπου XML. Οι Spiliotopoulos et al. [38, 47] στοχεύουν στην προσβασιμότητα των πινάκων στα έγγραφα με ακουστικό τρόπο. Οι Edwards et al. [48] προσπαθούν να «κάνουν» τα μαθηματικά προσβάσιμα σε τυφλούς μαθητές και οι Francioni & Smith [49] προτείνουν ένα πλαίσιο εργασίας για την προσβασιμότητα σε μαθήματα της επιστήμης των υπολογιστών για τυφλούς μαθητές. Στις εργασίες που δημοσιεύτηκαν στο πλαίσιο της διατριβής [50] [51] παρουσιάζεται μια προσέγγιση με γνώμονα τη Σχεδίαση-για-Όλους (Design-for-All) για προσβάσιμα έγγραφα στον πίνακα κατά την παρουσίαση στην τάξη (παράγραφος 2.9) και στηρίζεται στην αρχιτεκτονική που παρουσιάζεται στο Κεφάλαιο 2.

Πολυτροπική αλληλεπίδραση με τα έγγραφα, θεωρείται: *η εκτέλεση διεργασιών παρουσίασης και πλοήγησης σύμφωνα με τις προτιμήσεις του αναγνώστη σε έναν από τους τρεις τρόπους αλληλεπίδρασης, οπτικός, ακουστικός και απτικός ή οποιοσδήποτε προτεινόμενος συνδυασμός τους.* Οι Guillon et al. [52] προτείνουν μια ολοκληρωμένη διαδικασία έκδοσης (integrated publishing procedure) για πολυτροπική πρόσβαση στα έγγραφα βασισμένη στο πρότυπο DAISY 3.0 [53].

Η κοινοπραξία του παγκόσμιου ιστού (World Wide Web Consortium - W3C) [54] προσδιορίζει την πρόσβαση στον παγκόσμιο ιστό ως *μέσο για τους ανθρώπους με αναπηρίες ώστε να μπορούν να αντιληφθούν κατανοήσουν, πλοηγηθούν και να αλληλεπιδράσουν με τον Παγκόσμιο Ιστό και το περιεχόμενό του* [55]. Σύμφωνα με την W3C, πρέπει να υποστηρίζεται η πρόσβαση στα έγγραφα από οποιαδήποτε συσκευή:

- thin clients: συσκευές με μικρή επεξεργαστική ισχύ ή δυνατότητες που μπορούν να χρησιμοποιηθούν για τη λήψη της εισόδου του χρήστη – μικρόφωνα, οθόνες αφής,

stylus κ.α. – όπως και είσοδο πέρα από το χρήστη, όπως είναι από μια συσκευή GPS

- thick clients: συσκευές όπως PDA ή notebook
- medium clients: συσκευές με δυνατότητα ενός βαθμού ερμηνείας (interpretation) [56].

Επίσης, η έννοια **Σχεδιασμός-για-Όλους** (Design-for-all, ισοδύναμοι όροι: Καθολική Σχεδίαση, Σχεδίαση Ενσωμάτωσης) *δε συνεπάγεται μια ενιαία σχεδιαστική λύση, κατάλληλη για όλους τους χρήστες*. Αλλά, πρέπει να ερμηνευθεί ως «η ενσυνείδητη και συστηματική προσπάθεια της εκ των προτέρων εφαρμογής αρχών, μεθόδων και εργαλείων με σκοπό την ανάπτυξη προϊόντων και υπηρεσιών Πληροφορικής και Τηλεπικοινωνιών που είναι προσβάσιμα και χρηστικά από όλους τους πολίτες και επομένως η αποφυγή της ανάγκης εκ των υστέρων προσαρμογών ή εξειδικευμένης σχεδίασης» [40]. Έτσι, χρησιμοποιώντας αυτή την έννοια στην προσβασιμότητα των εγγράφων, στοχεύουμε σε ένα μεγαλύτερο εύρος χρηστών, πέρα από τα άτομα με εντυποαναπηρία, που προτιμούν την ακουστική πρόσβαση, όπως για παράδειγμα, οι μετακινούμενοι χρήστες (που μπορεί να μην έχουν άμεσα οπτική πρόσβαση στο έγγραφο), οι ηλικιωμένοι και γενικότερα κάποιος που θέλει-έχει ανάγκη για ακουστική πρόσβαση στο περιεχόμενο των εγγράφων.

Οι αρχές της Σχεδίασης για Όλους περιλαμβάνουν:

1. Δικαίωμα στη χρήση.
2. Ευελιξία στη χρήση.
3. Απλότητα και διαισθητικότητα.
4. Αντιληπτικότητα πληροφοριών.
5. Ανοχή στα σφάλματα.
6. Χαμηλή φυσική προσπάθεια.
7. Κατάλληλο μέγεθος και χώρος για προσέγγιση και χρήση.

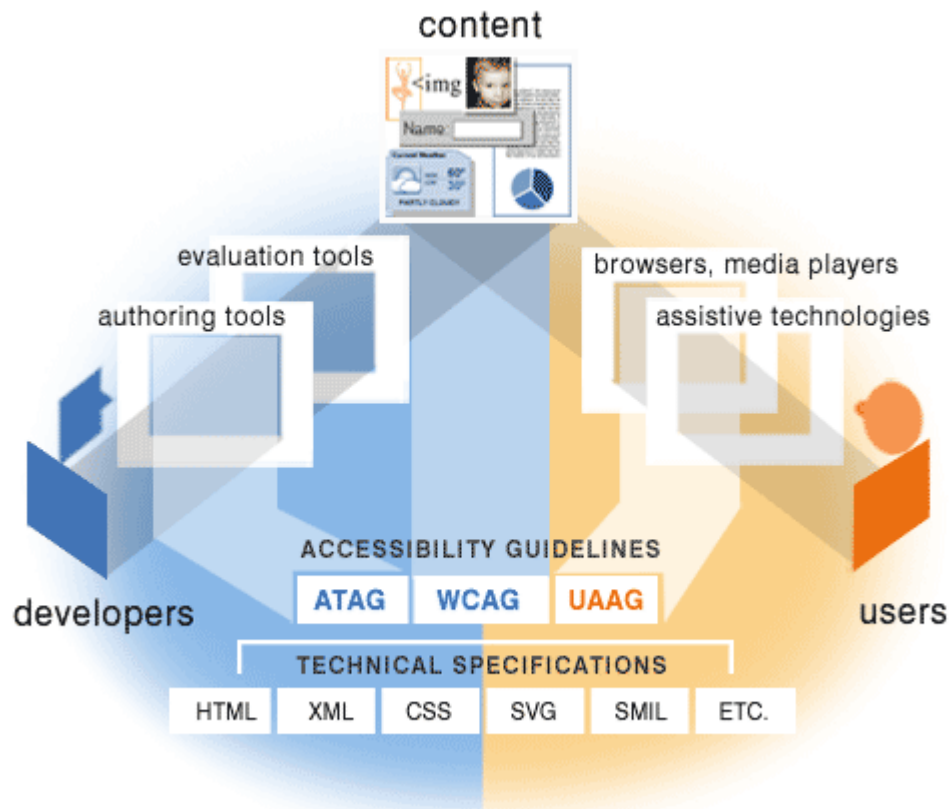
Πέρα από τις παραπάνω προσπάθειες, δημιουργείται ένα σημαντικό ερώτημα για τις διεργασίες παρουσίασης και πλοήγησης: *Είναι τα έγγραφα προσβάσιμα από όλους;* Στις παρακάτω παραγράφους παρουσιάζονται μερικά σημαντικά πρότυπα που συνεισφέρουν προς την κατεύθυνση της προσβασιμότητας των εγγράφων από όλους (Universal Accessibility – Καθολική Πρόσβαση).

1.4.1 Κατευθυντήριες γραμμές για την Προσβασιμότητα του Ιστού

Οι κατευθυντήριες γραμμές για την προσβασιμότητα του περιεχομένου στον Ιστό (Web Content Accessibility Guidelines - WCAG) [57] είναι μέρος των οδηγιών που παρέχει η πρωτοβουλία Web Accessibility Initiative – WAI [55] της W3C. Ο σκοπός αυτών των οδηγιών είναι η δημιουργία περιεχομένου του Παγκόσμιου Ιστού σε προσβάσιμη μορφή για τα άτομα με αναπηρία. Τα ουσιαστικά συστατικά της προσβασιμότητας του ιστού (Σχήμα 1) παρουσιάζουν τον τρόπο με τον οποίο η προσβασιμότητα εξαρτάται από τα διάφορα συστατικά της ανάπτυξης εφαρμογών Ιστού και την αλληλεπίδραση τους, καθώς και τη συσχέτιση αυτών με τις οδηγίες του WAI [55].

Οι οδηγίες αφορούν τους προγραμματιστές-δημιουργούς περιεχομένου για τον παγκόσμιο ιστό, καθώς και εκείνους που δημιουργούν εργαλεία (authoring – evaluation). Ακολουθώντας αυτές τις οδηγίες το περιεχόμενο γίνεται ανεξάρτητο από το είδος-τύπο συσκευής και εφαρμογής, κάνοντάς το προσβάσιμο και με τους τρεις τρόπους αλληλεπίδρασης. Οι οδηγίες απευθύνονται όχι μόνο σε ανθρώπους με μόνιμη ανικανότητα, αλλά και σε εκείνους με παροδική ή προσωρινή αναπηρία, για παράδειγμα ένα άτομο το οποίο χρησιμοποιεί μια κινητή συσκευή σε θορυβώδες περιβάλλον.

Οι υπεύθυνοι ανάπτυξης ενθαρρύνονται για τη χρήση περιεχομένου πολυμέσων (όπως εικόνες, βίντεο, ήχους κ.α.), αλλά με τέτοιο τρόπο ώστε να είναι σωστά και κατάλληλα προσβάσιμο.



Σχήμα 1: Το βασικό μοντέλο συσχέτισης των συστατικών, των τεχνικών χαρακτηριστικών και των οδηγιών WAI της W3C για προσβασιμότητα στο περιεχόμενο του Ιστού

Το περιεχόμενο μιας ιστοσελίδας μπορεί να παρουσιάζεται με ένα ισοδύναμο κείμενο ή χωρίς κείμενο (non-textual). Αυτοί οι δύο όροι είναι κατανοητοί χρησιμοποιώντας τα παρακάτω παραδείγματα:

- Ας υποθέσουμε ότι υπάρχει μια εικόνα σε μια ιστοσελίδα. Ένας μη-βλέπων χρήστης είναι αδύνατο να δει την εικόνα. Μια περιγραφή της εικόνας (ισοδύναμο κείμενο) μπορεί να καταστήσει την εικόνα προσβάσιμη. Αυτή μπορεί να είναι, σύμφωνα με τις προθέσεις του συγγραφέα, απλή περιγραφή της εικόνας, που είναι συμπληρωματική στο κυρίως περιεχόμενο της ιστοσελίδας ή να καθοδηγήσει το χρήστη να επιλέξει την εικόνα με τον κέρσορα για λόγους πλοήγησης. Επομένως, δύο χρήστες με διαφορετικές ανάγκες μπορούν να έχουν πρόσβαση στην εικόνα χρησιμοποιώντας το ίδιο είδος φυλλομετρητή: ο βλέπων χρήστης μπορεί να δει την εικόνα και ο μη-βλέπων να ακούσει την περιγραφή χρησιμοποιώντας συνθετική ομιλία με τη χρήση συστημάτων μετατροπής Κειμένου-σε-Ομιλία ή να «διαβάσει» την περιγραφή με τη χρήση μιας οθόνης Braille.

- Άλλα πολυμέσα που χρησιμοποιούνται σε μια ιστοσελίδα είναι η προηγρογραφημένη φωνή (για παράδειγμα ένα μήνυμα καλωσορίσματος στην ιστοθέση). Ένας κωφός χρήστης είναι αδύνατον να ακούσει το καλωσόρισμα. Μπορεί όμως η ύπαρξη ενός κειμένου να καταστήσει το ακουστικό μήνυμα προσβάσιμο στο χρήστη.
- Στο προηγούμενο παράδειγμα, ένας κωφός χρήστης μπορεί εναλλακτικά να «ακούσει» την περιγραφή του μηνύματος χρησιμοποιώντας μια ροή βίντεο ή ένα εικονικό πράκτορα ο οποίος μεταφράζει την περιγραφή σε νοηματική γλώσσα.

Οι οδηγίες WAI θέτουν δύο βασικούς στόχους: να βεβαιώνεται η *ορθή μετατροπή του περιεχομένου* και να *καθίσταται κατανοητό και πλοηγήσιμο*. Οι οδηγίες Web Content Accessibility Guidelines Ver. 1.0 [57] παροτρύνουν ώστε:

- Να παρέχονται τα ισοδύναμα εναλλακτικά για κάθε αισθητηριακό τρόπο (οπτικό, ακουστικό και απτικό).
- Να επιβεβαιώνεται ότι το κείμενο και τα γραφικά είναι καταληπτά ακόμα και όταν παρουσιάζονται χωρίς χρώμα.
- Να χρησιμοποιείται η επισημείωση (markup) και τα φύλλα μορφοποίησης (style sheets) με όσο πιο ορθό τρόπο γίνεται. Να γίνεται η επισημείωση των εγγράφων με τα κατάλληλα δομικά στοιχεία.
- Να γίνεται ορθή χρήση και αποσαφήνιση της φυσικής γλώσσας. Δηλαδή, να χρησιμοποιείται η επισημείωση που διευκολύνει την προφορά ή την ερμηνεία των συντομογραφιών ή του ξένου κειμένου.
- Να δημιουργούνται πινάκες που μετασχηματίζονται ομαλά. Να βεβαιώνεται ότι οι πίνακες έχουν την απαραίτητη επισημείωση, ώστε να μετατρέπονται χρησιμοποιούνται από προσβάσιμους φυλλομετρητές και άλλα προγράμματα (user agents).
- Να βεβαιώνεται ότι οι ιστοσελίδες που χρησιμοποιούν νέες τεχνολογίες είναι εύκολα και ομαλά μετατρέψιμες. Οι σελίδες πρέπει να είναι προσβάσιμες ακόμα και όταν οι νεώτερες τεχνολογίες δεν υποστηρίζονται ή είναι απενεργοποιημένες.
- Να διασφαλίζεται ο έλεγχος από τον χρήστη των εναλλαγών του χρονικά εξαρτώμενου περιεχομένου. Να βεβαιώνεται ότι κινούμενα, αναβοσβήνοντα, κυλιόμενα ή αυτόματα ενημερωμένα αντικείμενα ή σελίδες μπορούν να σταματήσουν ή να διακοπεί η λειτουργία τους (pause).

- Να διασφαλίζεται η άμεση προσβασιμότητα όλων των ενσωματωμένων διεπαφών χρήστη. Το περιβάλλον εργασίας χρήστη πρέπει να ακολουθεί τις αρχές του προσβάσιμου σχεδιασμού: ανεξαρτήτου συσκευής πρόσβαση στη λειτουργικότητα, λειτουργικότητα του πληκτρολογίου κτλ.
- Να γίνεται σχεδιασμός ανεξάρτητα συσκευής. Να χρησιμοποιούνται χαρακτηριστικά που ενεργοποιούν τα στοιχεία της σελίδας μέσω μιας ποικιλίας συσκευών εισόδου.
- Να χρησιμοποιούνται ενδιάμεσες λύσεις προσβασιμότητας, ώστε οι υποστηρικτικές τεχνολογίες και οι παλαιότεροι φυλλομετρητές να λειτουργούν σωστά.
- Να γίνεται χρήση των τεχνολογιών και οδηγιών της W3C.
- Να παρέχονται πληροφορίες περιεχομένου και προσανατολισμού, με σκοπό τη βοήθεια των χρηστών στην κατανόηση σύνθετων σελίδων και στοιχείων .
- Να παρέχονται σαφείς και συνεπείς μηχανισμοί πλοήγησης, όπως πληροφορίες προσανατολισμού, γραμμές πλοήγησης, κλπ., ώστε να αυξάνεται η πιθανότητα εύρεσης των ζητούμενων πληροφοριών από ένα άτομο σε μια ιστοθέση.
- Να βεβαιώνεται ότι τα έγγραφα είναι σαφή και απλά, ώστε να είναι πιο ευνόητα.

Οι οδηγίες Web Content Accessibility Guidelines Ver. 2.0 [57] περιλαμβάνει συνοπτικά τα εξής:

1: Αντιληπτό περιεχόμενο

- 1.1 Παρέχετε εναλλακτικά κείμενα για κάθε περιεχόμενο που δεν διατίθεται σε μορφή κειμένου, ώστε να μπορεί να αποδοθεί σε άλλες μορφές που χρειάζονται οι χρήστες, όπως για παράδειγμα με μεγάλη γραμματοσειρά, σε Braille, σε ομιλία, με χρήση συμβόλων, ή σε πιο απλή γλώσσα.
- 1.2 Παρέχετε εναλλακτικές λύσεις για πολυμέσα που εξαρτώνται από το χρόνο.
- 1.3 Δημιουργείτε περιεχόμενο που μπορεί να αποδοθεί με διαφορετικούς τρόπους (για παράδειγμα με απλούστερη διάταξη), χωρίς απώλεια πληροφοριών ή δομής.
- 1.4 Διευκολύνετε την οπτική και ηχητική αντίληψη του περιεχομένου από τους χρήστες, συμπεριλαμβανομένης της διάκρισης των πληροφοριών προσκηνίου από το παρασκήνιο.

2: Λειτουργικό περιεχόμενο

- 2.1 Καταστήστε δυνατή τη χρήση όλων των λειτουργιών μέσω πληκτρολογίου.
- 2.2 Παρέχετε στους χρήστες επαρκή χρόνο για την ανάγνωση και χρήση του περιεχομένου.
- 2.3 Μην σχεδιάζετε περιεχόμενο με τρόπο που είναι γνωστό ότι προκαλεί επιληπτικές κρίσεις λόγω φωτοευαισθησίας.
- 2.4 Παρέχετε μηχανισμούς που βοηθούν τους χρήστες να εντοπίζουν περιεχόμενο, να προσανατολίζονται και να περιηγούνται σε αυτό.

3: Κατανοητό περιεχόμενο

- 3.1 Σε μορφή κειμένου περιεχόμενο πρέπει να είναι αναγνώσιμο και κατανοητό.
- 3.2 Δημιουργείτε ιστοσελίδες με προβλέψιμη διάταξη και λειτουργία.
- 3.3 Βοηθάτε τους χρήστες να αποφεύγουν και να διορθώνουν τυχόν λάθη τους.

4: Εύρωστο περιεχόμενο (περιεχόμενο χωρίς σφάλματα)

- 4.1 Ενισχύστε τη συμβατότητα με τρέχοντες και μελλοντικούς πράκτορες χρήστη, συμπεριλαμβανομένων των υποστηρικτικών τεχνολογιών.

Οι WCAG έχουν υιοθετηθεί στο νομικό πλαίσιο αρκετών χωρών. Για παράδειγμα, στην Ευρωπαϊκή Ένωση η Digital Agenda for Europe (26/8/2010) αναφέρει στο άρθρο 2.6.2: *«Υπάρχει ανάγκη για συντονισμένες δράσεις ώστε να εξασφαλιστεί ότι το νέο ηλεκτρονικό περιεχόμενο είναι επίσης πλήρως στη διάθεση των ατόμων με αναπηρίες. Ειδικότερα, οι δημόσιοι δικτυακοί τόποι και οι υπηρεσίες της ΕΕ που είναι απαραίτητες για πλήρη συμμετοχή στη δημόσια ζωή πρέπει να ευθυγραμμιστούν με τα διεθνή πρότυπα πρόσβασης στο διαδίκτυο (WCAG 2.0). Εξάλλου, η Σύμβαση των Ηνωμένων Εθνών για τα δικαιώματα των ατόμων με αναπηρία περιέχει υποχρεώσεις σχετικά με την προσβασιμότητα».*

1.4.2 Open Document Format Accessibility

Το OpenDocument Format (ODF) [58] είναι ένα τύπος αρχείου, ανοιχτού κώδικα βασισμένο στην XML [59], για τη χρήση του σε εφαρμογές γραφείου, για έγγραφα που περιέχουν κείμενο, ιστογράμματα-διαγράμματα και στοιχεία γραφικών. Αυτό το είδος αρχείου καθιστά τη μετατροπή σε άλλους τύπους αρχείων απλή επαναχρησιμοποιώντας και ενισχύοντας ήδη υπάρχοντα πρότυπα όπου είναι δυνατόν [60]. Η δημιουργία και υποστήριξη αυτού του πρότυπου, θέτει τη δυνατότητα για την

ανάπτυξη νέων εφαρμογών και τη συμβατότητα με παλαιότερες - παραδοσιακές εφαρμογές (backward compatibility). Το σχήμα του ODF (ODF schema) παρέχει υψηλού επιπέδου πληροφορίες κατάλληλες για την τροποποίηση-επεξεργασία των εγγράφων. Προσδιορίζει τις κατάλληλες δομές XML για τα έγγραφα γραφείου και είναι φιλικό στις μετατροπές χρησιμοποιώντας XSLT [61] ή παρόμοια εργαλεία βασισμένα στην XML.

Υπό τις οδηγίες και την υποστήριξη του ODF, έχουν δημιουργηθεί προδιαγραφές με σκοπό την προσβασιμότητα. Στόχος των προδιαγραφών είναι να μελετηθούν και να βελτιωθούν θέματα που αφορούν την προσβασιμότητα καθώς και την ενίσχυση των διαδικασιών δημιουργίας, ανάγνωσης και τροποποίησης των εγγράφων για άτομα με αναπηρία [62]. Το Open Document Format περιλαμβάνει τόσο δομικές όσο και σημασιολογικές πληροφορίες που χρειάζονται για την ορθή προσπέλαση και πρόσβαση στις πληροφορίες από άτομα με αναπηρίες. Η Open Document Accessibility Subcommittee κατηγοριοποιεί την προσβασιμότητα σε τρεις τύπους προσπέλασης: *άμεση, με διαμεσολάβηση και έμμεση* [62].

Συνδυάζοντας τη χρήση των υποστηρικτικών τεχνολογιών και των προδιαγραφών Open Document Accessibility, ένας χρήστης με αναπηρία μπορεί να έχει άμεση πρόσβαση στο περιεχόμενο ενός εγγράφου. Τα είδη των ανικανότητων που υποστηρίζονται από το ODF – Accessibility είναι:

- Ελαφρές ανικανότητες όρασης.
- Σημαντικές ανικανότητες όρασης.
- Μερική ή ολική τύφλωση.
- Ελαφρές σωματικές ανικανότητες.
- Σημαντικές σωματικές ανικανότητες χωρίς αναγνώριση ομιλίας.
- Σημαντικές σωματικές ανικανότητες με αναγνώριση ομιλίας.
- Ανικανότητες ακοής.
- Γνωσιακές Ανικανότητες.

Οι εφαρμογές που υποστηρίζουν το ODF πρέπει να λαμβάνουν υπόψη τις επιλογές του χρήστη για την επιφάνεια εργασίας του Η/Υ, όπως επιλογή γραμματοσειράς, χρώματος και αντίθεσης. Για παράδειγμα, οι χρήστες με ανικανότητες όρασης (όπως αχρωματοψία) χρησιμοποιούν διαφορετικές ρυθμίσεις από τις προκαθορισμένες του

συστήματος, ώστε να μπορούν να δουν τις διεπαφές στην επιφάνεια εργασίας. Η ODF εφαρμογή, πρέπει να χρησιμοποιεί αυτές τις διαφορετικές ρυθμίσεις, π.χ. κείμενα σε παράθυρα διαλόγου και μενού. Ιδανικά, η εφαρμογή θα πρέπει να τροποποιεί την παρουσίαση της χωρίς την παρέμβαση του χρήστη [62].

Σύμφωνα με τις οδηγίες, η εφαρμογή ODF είτε χρησιμοποιεί ένα υπάρχον πλαίσιο εργασίας προσβασιμότητας, είτε το δημιουργεί η ίδια, είναι σημαντικό ότι θα παρέχει ακριβώς όλες τις πληροφορίες για τα στοιχεία διεπαφής χρήστη που μπορεί να μεταδώσει το πλαίσιο προσβασιμότητας. Δηλαδή, η μετάδοση της πληροφορίας (συμπεριλαμβανομένων και των τυπογραφικών στοιχείων και γενικά του τρόπου παρουσίασης) δεν έγκειται στις οδηγίες του ODF αλλά στο εκάστοτε πλαίσιο εργασίας που επικοινωνεί με την εφαρμογή.

1.4.3 Math Markup Language

Η MathML [63] παρέχει οδηγίες για την παραγωγή και παρουσίαση των μαθηματικών με την χρήση της XML. Γενικά, η παρουσίαση των μαθηματικών και επιστημονικών εκφράσεων στα ηλεκτρονικά έγγραφα γινόταν με τη χρήση εικόνων ή φωτογραφιών και τη λεζάντα ως την περιγραφή τους. Αλλά, αυτές οι εικόνες δεν είναι προσβάσιμες π.χ. από ένα τυφλό φοιτητή. Αυτό οφείλεται στο ότι ο αναγνώστης οθόνης δεν μπορεί να «διαβάσει» την εικόνα, αλλά μόνο την περιγραφή που παρέχεται από τη λεζάντα ή τη μεταπληροφορία που συνοδεύει την εικόνα.

Η MathML έχει σχεδιαστεί με γνώμονα:

- Την κωδικοποίηση του μαθηματικού υλικού κατάλληλο για τη διδασκαλία και την επιστημονική επικοινωνία σε όλες τις βαθμίδες εκπαίδευσης.
- Την κωδικοποίηση μαθηματικών εκφράσεων και εννοιών.
- Τη διευκόλυνση της μετατροπής από και προς άλλους τύπους μαθηματικών, παρουσίασης ή σημασίας (presentational - semantic). Η έξοδος πρέπει να υποστηρίζει: γραφικές απεικονίσεις, TtS, άλλες γλώσσες μορφοποίησης μαθηματικών, όπως είναι η TEX π.χ. μέσα εκτύπωσης (print media), συμπεριλαμβανομένης και της γραφής Braille.

Η MathML παρέχει χαμηλού επιπέδου προδιαγραφές για την περιγραφή των μαθηματικών ως βάση για την επικοινωνία μεταξύ των μηχανών-υπολογιστών. Παρέχει τις βάσεις ώστε να περιλαμβάνονται οι μαθηματικές περιγραφές στις ιστοσελίδες [64].

Δεν υποστηρίζει μόνο την οπτική παρουσίαση, αλλά και τις έννοιες-σημασίες των μαθηματικών εκφράσεων (όπως «δια», «επί», «ρίζα του»). Για παράδειγμα, οι ενσωματωμένες μαθηματικές εκφράσεις χρησιμοποιώντας MathML στις ιστοσελίδες, μπορούν να φαίνονται ως κανονική ιστοσελίδα (οπτική παρουσίαση). Οι τυφλοί μπορούν να χρησιμοποιούν έναν αναγνώστη οθόνης μαζί με ένα σύστημα μετατροπής Κειμένου-σε-Ομιλία για να ακούσουν την περιγραφή της μαθηματικής έκφρασης χρησιμοποιώντας MathML (σημασιολογική παρουσίαση).

Μια δωρεάν εφαρμογή (plug-in) προγράμματος για τους φυλλομετρητές [65] [66], μας δίνει τη δυνατότητα της ακουστικής απόδοσης των μαθηματικών εκφράσεων χρησιμοποιώντας MathML.

1.4.4 Braille Markup Language

Ένα πρότυπο για την απτική παρουσίαση του περιεχομένου των εγγράφων σε Braille αποτελεί η Braille Markup Language (BrailleML). Η BrailleML αποτελεί μια προσπάθεια για την προτυποποίηση των εγγράφων σε μορφή Braille. Οι Masanori et al., [67] προτείνουν την αυτόματη μετατροπή των ODF εγγράφων σε έγγραφα Braille χρησιμοποιώντας την BrailleML.

Γενικά, τα τυπογραφικά στοιχεία ενός εγγράφου (όπως η έντονη και η πλάγια γραφή) μπορούν να παρέχονται στις οθόνες Braille π.χ. με τη χρήση συμβόλων πριν από τη λέξη-φράση, που υποδηλώνει το τυπογραφικό στοιχείο που εφαρμόζεται σε αυτή [68].

1.4.5 Scalar Vector Graphics

Scalar Vector Graphics (SVGs) είναι μια γλώσσα για την περιγραφή δισδιάστατων γραφικών και γραφικών εφαρμογών σε XML [69]. Το πρότυπο SVG παρέχει ένα αριθμό στοιχείων-χαρακτηριστικών για να γίνουν τα γραφικά του παγκόσμιου ιστού πιο προσβάσιμα σε ένα μεγαλύτερο εύρος ανθρώπων. Οι χρήστες που μπορούν να βοηθηθούν από αυτό το πρότυπο είναι: οι έχοντες χαμηλή όραση, αχρωματοψία-δυσχρωματοψία ή τυφλοί χρήστες και γενικά εκείνοι που χρησιμοποιούν υποστηρικτικές τεχνολογίες. Ένας αριθμός από αυτά τα χαρακτηριστικά μπορεί να αυξήσει τη χρησιμότητα του περιεχομένου για πολλούς χρήστες χωρίς αναπηρίες, όπως χρήστες φορητών συσκευών (PDA, κινητά τηλέφωνα και άλλες μη-παραδοσιακές συσκευές για πρόσβαση στον παγκόσμιο ιστό) [70].

Στους χρήστες με αναπηρία παρέχονται πολλές δυνατότητες προσβασιμότητας από τις προδιαγραφές του προτύπου. Οι εικόνες σε μορφή SVG είναι βαθμωτές (scalable) –

μπορούν να μεγεθυνθούν και να μεταβληθεί το μέγεθός τους σύμφωνα με τις ανάγκες του χρήστη. Η δυνατότητα της κλιμάκωσης μπορεί να βοηθήσει τους χρήστες με χαμηλή όραση και τους χρήστες ορισμένων υποστηρικτικών τεχνολογιών (π.χ. απτικές γραφικές συσκευές, που τυπικά έχουν χαμηλή ανάλυση).

1.4.6 DAISY/NISO

Η κοινοπραξία DAISY (Digital Accessible Information SYstem) [53] είναι ένας μη-κερδοσκοπικός οργανισμός που υποστηρίζει το πρότυπο DAISY/NISO (επίσημα γνωστό και ως ANSI/NISO Z39.86 [71]). Το πρότυπο προσδιορίζεται από τη μορφή και το περιεχόμενο του συνόλου των ηλεκτρονικών αρχείων, το οποίο συνθέτει ένα ψηφιακό ομιλόν βιβλίο – ΨΟΒ (Digital Talking Book - DTB) και καθορίζει ένα συγκεκριμένο σύνολο απαιτήσεων για τις συσκευές αναπαραγωγής ψηφιακών ομιλούντων βιβλίων (DTB playback devices):

- Την απόδοση του συστήματος αναπαραγωγής που σχετίζεται με τις απαιτήσεις του αρχείου.
- Τη συμπεριφορά του συστήματος αναπαραγωγής σε περιοχές που ορίζονται στις απαιτήσεις των χρηστών.

Το πρότυπο προσδιορίζει τις οδηγίες για την παραγωγή και την παρουσίαση των ΨΟΒ για τους εντυποανάπηρους αναγνώστες (τυφλούς, οπτικά ανάπηρους, σωματικά ανάπηρους και εκείνους με μαθησιακά προβλήματα).

Ένα ΨΟΒ είναι: μια συλλογή από ηλεκτρονικά αρχεία οργανωμένα με τέτοιο τρόπο, ώστε να παρουσιάζει τις πληροφορίες σε συγκεκριμένο μέρος του πληθυσμού δια εναλλακτικών μέσων, δηλαδή, συνθετική ή πραγματική ομιλία, ανανεώσιμες οθόνες Braille ή οθόνες π.χ. μεγάλης εκτύπωσης-παρουσίασης [71].

Τα αρχεία που συνθέτουν ένα ΨΟΒ μπορούν να χωριστούν σε 10 κατηγορίες, όπως αρχείο με το περιεχόμενο του κειμένου, αρχεία ήχου, εικόνων, συγχρονισμού, τρόπου προβολής κ.α. Ο συνδυασμός και η ενοποίηση των λειτουργιών αυτών των αρχείων και η δημιουργία ενός ΨΟΒ σύμφωνα με το πρότυπο, ενσωματώνει πολλά χαρακτηριστικά είτε για διεργασίες Πλοήγησης είτε Παρουσίασης κατά τη διαδικασία ανάγνωσης. Μερικά χαρακτηριστικά είναι:

- Κατά τη διεργασία πλοήγησης: γρήγορη και ευέλικτη πλοήγηση, χρήση σελιδοδείκτη και επισημείωση σημαντικών σημείων, αναζήτηση λέξεων-κλειδιών.

- Κατά τη διεργασία παρουσίασης: έλεγχο του τρόπου παρουσίασης επιλεγμένων στοιχείων, π.χ. υποσημείωση, αριθμοί σελίδας κτλ.

Μερικά από τα χαρακτηριστικά πλοήγησης που παρέχονται από τα ΨΟΒ είναι: γρήγορη μετάβαση εμπρός ή πίσω, ανάγνωση σε μεταβλητή ταχύτητα, σημειώσεις - επισημειώσεις, αναζήτηση, λίστες, μαθηματικές εξισώσεις-επιστημονικά σύμβολα κ.α. Αυτά τα χαρακτηριστικά δίνουν τη δυνατότητα διεργασιών παρουσίασης ή/και πλοήγησης με οπτικό, ακουστικό ή απτικό τρόπο. Σύμφωνα με τις ανάγκες του χρήστη ή/και την αναπηρία του, το έγγραφο μπορεί να παρουσιαστεί ή προσπελαστεί με ένα ή με συνδυασμό των παραπάνω τρόπων.

Εάν κάποιος ακολουθήσει το πρότυπο μπορεί να υλοποιήσει ένα σύστημα αναπαραγωγής ΨΟΒ (DTB player) με ποικιλία δυνατοτήτων και λειτουργιών. Ένα σύστημα αναπαραγωγής ΨΟΒ μπορεί να είναι μια φορητή συσκευή, η οποία απλά «διαβάζει» στο χρήστη το περιεχόμενο του βιβλίου χρησιμοποιώντας συνθετική ομιλία, υποστηρίζοντας βασικές λειτουργίες πλοήγησης, όπως γρήγορη μετάβαση εμπρός ή πίσω, μεταβλητή ταχύτητα ανάγνωσης. Εναλλακτικά, ένα πιο αποδοτικό σύστημα αναπαραγωγής βασισμένο στον υπολογιστή, μπορεί να υποστηρίζει κάθε τρόπο πρόσβασης και όλα τα χαρακτηριστικά πλοήγησης που παρέχει το πρότυπο. Οι προγραμματιστές, κατά τη δημιουργία των ΨΟΒ και την ανάπτυξη συστημάτων αναπαραγωγής τους, παροτρύνονται να δημιουργούν τις κατάλληλες συνθήκες για να ορίζει ο χρήστης τον τρόπο με τον οποίο παρουσιάζεται το περιεχόμενο, είτε σε οπτικό, ακουστικό ή απτικό (και πώς να παρουσιάζεται σε κάθε αισθητηριακό τρόπο).

Στα ΨΟΒ μπορεί να περιλαμβάνεται η οπτική πληροφορία για τη διάταξη, δομή και παρουσίαση του, σε περίπτωση που χρησιμοποιείται από εφαρμογές-συσκευές οι οποίες υποστηρίζουν αυτή την πληροφορία. Αυτό μπορεί να επιτυγχάνεται με τη χρήση προτύπων, όπως το CSS ή XSL ώστε να διαχωρίζεται το περιεχόμενο από την παρουσίαση. Για παράδειγμα, εάν χρησιμοποιείται οθόνη Braille, τότε το Braille Style Sheet μπορεί να ελέγχει τη μορφοποίηση ώστε το έγγραφο να είναι εύκολα πλοηγήσιμο. Σε περίπτωση που χρησιμοποιείται για την απόδοση με ακουστικό τρόπο του περιεχομένου, μπορεί να χρησιμοποιηθεί το Aural Style Sheets (ASS) [72] για την ισοδύναμη ακουστική απόδοση των οπτικών χαρακτηριστικών των εγγράφων.

1.5 Συναισθήματα

Τα συναισθήματα αποτελούν αναπόσπαστο κομμάτι της ανθρώπινης ύπαρξης και κατ' επέκταση της καθημερινής μας δραστηριότητας. Σε αυτό το μέρος του κεφαλαίου,

περιγράφονται γενικά οι θεωρίες-προσεγγίσεις που αφορούν τα συναισθήματα. Η θεωρητική ανασκόπηση μας βοηθάει στην κατανόηση του τρόπου με τον οποίο επηρεάζονται τα συναισθήματα και η συναισθηματική κατάσταση του αναγνώστη εγγράφων από τα τυπογραφικά στοιχεία, στη μέτρηση και τη βέλτιστη πειραματική διαδικασία για την ορθότερη μοντελοποίηση τους.

Στόχος είναι να γίνει κατανοητή η σύνδεση των συναισθημάτων του αναγνώστη με τα έγγραφα (είτε σε επίπεδο παρουσίασης, είτε σε επίπεδο περιεχομένου). Δε θα αναλυθούν λεπτομερώς οι θεωρίες περί συναισθημάτων, γιατί είναι πέρα από τον σκοπό της παρούσας εργασίας. Η διατριβή της Μαλατέστα [73] αποτελεί μια πολύ καλή αναφορά, όπου αναλύονται διεξοδικά οι επικρατούσες θεωρίες των συναισθημάτων και ο τρόπος χρήσης τους στην επιστήμη των υπολογιστών.

Οι θεωρήσεις για τα συναισθήματα, σύμφωνα με τον Cornelius [74], κατηγοριοποιούνται ως εξής:

- Δαρβινική Θεωρία (Darwinian Theory),
- Θεωρία του James (Jamesian Theory),
- Γνωσιακή Θεώρηση – Γνωστική Θεωρία (Cognitive Theory),
- Θεωρία Κοινωνικού Δομισμού ή Δομητισμού (Social Constructivist Theory).

Η Δαρβινική θεωρία υποθέτει ότι τα συναισθήματα είναι [74]:

«δυναμικά-εξελιξίμα φαινόμενα με σημαντικές λειτουργίες επιβίωσης, και η επιλογή τους οφείλεται στο ότι έχουν λύσει συγκεκριμένα προβλήματα που έχουμε αντιμετωπίσει ως είδος.»

Όλοι οι οργανισμοί που μοιράζονται ένα εξελικτικό παρελθόν πρέπει να μοιράζονται και τα ίδια συναισθήματα. Τα συναισθήματα πρέπει να αναλύονται υπό όρους λειτουργικότητας και αξίας επιβίωσης [74] [75].

Σύμφωνα με τη θεώρηση του James [76] [74]:

«οι σωματικές αλλαγές ακολουθούν άμεσα την αντίληψη του ερεθίσματος και η αίσθησή μας για τις ίδιες αλλαγές όσο συμβαίνουν, είναι το συναίσθημα»

Η πεποίθηση του James είναι ότι βιώνουμε τα συναισθήματα διότι το σώμα μας έχει εξελιχθεί ώστε να αντιδρά αυτόματα και να προσαρμόζεται σε χαρακτηριστικά του περιβάλλοντος, που σχετίζονται σημαντικά με την επιβίωση μας. Πρώτα αντιδρά-ανταποκρίνεται το σώμα μας και η εμπειρία πάνω σε αυτές τις αλλαγές συνθέτουν αυτό

που ονομάζουμε συναίσθημα [74]. Σε αυτή την προσέγγιση, εάν θεωρήσουμε ένα σύστημα σαν «μαύρο κουτί», το ερέθισμα (π.χ. έγγραφο, ομιλία) θεωρείται ως η είσοδος του και οι σωματικές αλλαγές η έξοδος του. Έτσι μπορούμε να αλλάξουμε τα συναισθήματα κάποιου αλλάζοντας το ερέθισμα [75].

Η Γνωσιακή θεώρηση (Cognitive Theory) προτείνει ότι [74]:

«κάθε συναίσθημα είναι συνδεδεμένο με ένα συγκεκριμένο πρότυπο αξιολόγησης, εάν μεταβληθεί η αξιολόγηση, πρέπει να μεταβληθεί και το συναίσθημα αυτό.»

Η σκέψη και το συναίσθημα δεν μπορούν να διαχωριστούν. Το τι αισθάνεται κάποιος εξαρτάται από τον τρόπο που καταλαβαίνει και κρίνει τα γεγονότα. Τα συναισθήματα είναι αποτέλεσμα της νοητικής δραστηριότητας [74] [75].

Τέλος, σύμφωνα με τη θεώρηση Κοινωνικού Δομισμού, η *κουλτούρα και οι κοινωνικοί κανόνες παίζουν σημαντικό ρόλο στην οργάνωση των συναισθημάτων σε διάφορα επίπεδα*. Οι δύο τελευταίες θεωρήσεις είναι στενά συνδεδεμένες και η ανάπτυξη των σκέψεων (development of thoughts) είναι στενά συνδεδεμένες με την πολιτισμική και την κοινωνική κατάσταση [75].

1.5.1 Η έννοια της λέξης «συναίσθημα»

Έχοντας παραθέσει τις βασικές θεωρίες περί συναισθημάτων μπορούμε πλέον να προσεγγίσουμε τον ορισμό της έννοιας «συναίσθημα». Υπάρχει μια πληθώρα ερμηνειών-ορισμών. Σύμφωνα με τον Scherer [77] ορίζεται ως:

«Ένα επεισόδιο από διασυνδεδεμένες και συγχρονισμένες αλλαγές στις καταστάσεις όλων ή των περισσότερων από τα πέντε οργανικά υποσυστήματα, ως απόκριση στην αξιολόγηση ενός εξωτερικού ή εσωτερικού ερεθίσματος σχετικό με τις βασικές μέριμνες του οργανισμού.»

Παρόμοιο, αλλά πιο απλό ορισμό, δίνουν οι Brave και Nass [78]:

«Το συναίσθημα είναι η αντίδραση σε γεγονότα που σχετίζεται με τις ανάγκες, τους στόχους ή τις ανησυχίες ενός ατόμου και περικλείει τα συστατικά φυσιολογίας, επιρροής, συμπεριφοράς και γνωσιακής.»

Βέβαια, θα ήταν χρήσιμο σε αυτό το σημείο να ξεχωρίσουμε την έννοια του «συναίσθηματος» από τις έννοιες, «διάθεση» (mood) και «αίσθημα» (sentiment). Αυτές οι τρεις έννοιες συγχέονται πολύ συχνά, κυρίως κατά τη χρήση τους στην αγγλική γλώσσα [78].

Διάθεση (Mood)

Η διάθεση ενός ατόμου δεν είναι η αντίδραση του σε κάποιο π.χ. γεγονός ή αντικείμενο (συναίσθημα), αλλά [77]:

«Οι διαθέσεις θεωρούνται ως οι διάχυτες καταστάσεις επιρροής, που χαρακτηρίζονται από μία σχετική διαρκή επικράτηση συγκεκριμένων τύπων υποκειμενικών αισθητηριακών ερεθισμάτων που επηρεάζουν την πείρα-εμπειρία και τη συμπεριφορά ενός ατόμου.»

Η διάθεση σε σχέση με το συναίσθημα είναι μακράς διάρκειας. Μπορεί να διατηρηθεί για ώρες ακόμα και για μέρες. Θεωρείται ότι υπάρχει συσχετισμός-αλληλεπίδραση μεταξύ αυτών των δύο εννοιών. Η διάθεση «λειτουργεί» σαν φίλτρο για τα συναισθήματα μειώνοντας ή αυξάνοντας την έντασή τους (σχετιζόμενων με τη διάθεση) [78]. Ουσιαστικά, δρα σαν θόρυβος κατά τη μελέτη των συναισθημάτων. Αλλά και η μεταβολή των συναισθημάτων μπορεί να επηρεάσει άμεσα τη διάθεση του ατόμου.

Αίσθημα (Sentiment)

Σε αντίθεση με το «συναίσθημα» και τη «διάθεση» [78]:

«Τα αισθήματα δεν είναι καταστάσεις ενός ατόμου αλλά ιδιότητες που ανατίθενται σε ένα αντικείμενο.»

Το παράδειγμα που αναφέρεται στην ίδια εργασία και θα το αναπαράγουμε σε αυτό το σημείο, είναι κατάλληλο για να γίνει κατανοητή η έννοια του «αισθήματος». Όταν κάποιος/α λέει ότι του/της «αρέσει» μια διεπαφή ή τη βρίσκει «ενοχλητική», αντιστοιχεί μια ιδιότητα στη διεπαφή (δηλαδή στο αντικείμενο, όπως αναφέρεται στον ορισμό), αναμένοντας ότι η αλληλεπίδραση με τη διεπαφή θα τον/την οδηγήσει σε θετικά ή αρνητικά συναισθήματα.

Παρατηρούμε ότι τα συναισθήματα είναι άρρηκτα συνδεδεμένα με τις έννοιες της «διάθεσης» και του «αισθήματος». Μπορούν να επηρεάσουν τα τελευταία είτε θετικά είτε αρνητικά. Επίσης, μπορούν να επηρεάσουν και άλλες ανθρώπινες λειτουργίες, όπως [78]:

- την προσοχή
- την κατανόηση
- τη μνήμη
- την απόδοση

- την αξιολόγηση

Έτσι τα συναισθήματα πρέπει να λαμβάνονται υπόψη, π.χ. κατά το σχεδιασμό της εμφάνισης ενός κειμένου-εγγράφου, ώστε να αυξάνουν την ευχαρίστηση κατά την ανάγνωσή του, καθώς και τη συγκέντρωση του αναγνώστη για τη βέλτιστη αποστήθιση εννοιών από αυτό.

Πέρα από τα έγγραφα, έχει αναπτυχθεί ένας ολόκληρος επιστημονικός τομέας για τη μελέτη χρήσης των συναισθημάτων κατά την αλληλεπίδραση του ανθρώπου με τον υπολογιστή. Ονομάζεται «Συναισθηματική Υπολογιστική» (Affective Computing), δηλαδή η υπολογιστική που σχετίζεται, πηγάζει ή επηρεάζει τα συναισθήματα [79] [80] [81], αξιοποιώντας στοιχεία από τις γενικές θεωρίες συναισθημάτων αλλά και από τη Θεωρία Αποτίμησης (Appraisal Theory) [82] [83] [84] [85].

1.5.2 Πειραματικές διαδικασίες για τη μελέτη των συναισθημάτων

Η μελέτη των συναισθημάτων μπορεί να διακριθεί σε:

1. αναπαράσταση σε κατηγορίες (διακριτά συναισθήματα-κατηγορική περιγραφή). Για παράδειγμα, ο Ekman [86] προτείνει τη χρήση έξι βασικών συναισθημάτων: *θυμός, αηδία, φόβος, χαρά, λύπη* και *έκπληξη*. Βέβαια, αυτός ο αριθμός μπορεί να αυξηθεί και να φτάσει μέχρι τα δεκατέσσερα [77].
2. διαστατική προσέγγιση (dimensional approach). Σύμφωνα με τις [87] και [88] ο χώρος των συναισθημάτων δομείται από τρεις διαστάσεις: *ευχαρίστησης* (pleasure), *διέγερσης* (arousal) και *επικράτησης* (dominance) (αναλύονται λεπτομερώς στο εδάφιο 1.5.3). Οι διαστάσεις της ευχαρίστησης και της διέγερσης, θεωρούνται ως οι πιο επικρατέστερες [89]. Βέβαια σε πολλές περιπτώσεις δεν αρκούν μόνο οι δύο διαστάσεις για την ορθή μελέτη του εκάστοτε προβλήματος [90].

Η παρατήρηση των συναισθημάτων ενός ανθρώπου μπορεί να γίνει με πολλούς τρόπους, αξιοποιώντας τις διάφορες θεωρήσεις και ιδιότητες αυτών [73]. Για παράδειγμα, η αξιολόγηση της συναισθηματικής κατάστασης ενός ατόμου μπορεί να πραγματοποιηθεί:

- με παρατήρηση σε πραγματικό χρόνο της ψυχοσωματικής (psychophysical) και ψυχοφυσιολογικής (psychophysiological) αντίδρασης στο ερέθισμα (π.χ. έκφραση προσώπου, εφίδρωση, χτύποι καρδιάς, εγκεφαλικά κύματα κ.α.)

- με ερώτηση του χρήστη και περιγραφή του συναισθήματος που βιώνει μέσα από ειδικά διαμορφωμένα ερωτηματολόγια, αφού έχει πρώτα εκτεθεί στο ερέθισμα (αυτοαξιολόγηση).

Επικεντρώνοντας την προσοχή μας στη μεθοδολογία της αυτοαξιολόγησης, οι ψυχολόγοι έχουν αναπτύξει διάφορα εργαλεία και πειραματικές διαδικασίες που μπορούν να μετρήσουν τα συναισθήματα. Σύμφωνα με τον Scherer [77] μπορούν να κατηγοριοποιηθούν σε *Μέτρηση Ελεύθερης Απόκρισης - ΜΕΛΑ* (Free Response Measurement – FRM) και *Μέτρηση Εξαναγκασμένης Επιλογής Απόκρισης – ΜΕΞΑ* (Forced Choice Response Measurement – FCRM).

Χρησιμοποιώντας τη ΜΕΛΑ, οι συμμετέχοντες στην πειραματική διαδικασία καλούνται να απαντήσουν ελεύθερα, δηλαδή να χρησιμοποιήσουν σύμβολα ή μικρές φράσεις για να χαρακτηρίσουν τη φύση των συναισθημάτων που βιώνουν. Χρησιμοποιώντας αυτή τη διαδικασία, οι ερευνητές αντιμετωπίζουν προβλήματα, όπως στην περίπτωση που οι συμμετέχοντες δεν μπορούν να εκφράσουν τα συναισθήματά τους είτε λόγω της χρήσης ακατάλληλων συμβόλων, λέξεων ή φράσεων είτε λόγω του περιορισμένου εύρους του λεξιλογίου τους που περιορίζει τις απαντήσεις τους. Επίσης, είναι αρκετά δύσκολη η στατιστική ανάλυση των ελεύθερων απαντήσεων [77]. Οι ερευνητές ταξινομούν πολυάριθμες ελεύθερες απαντήσεις σε μικρό αριθμό από κατηγορίες συναισθημάτων, χρησιμοποιώντας συνώνυμα ή παρόμοιες λέξεις-εκφράσεις. Δεν είναι τυποποιημένη (standard) πειραματική διαδικασία, αλλά γίνονται προσπάθειες προς την προτυποποίηση των συναισθηματικών κατηγοριών-ετικετών (labels).

Οι ΜΕΞΑ μπορούν να κατηγοριοποιηθούν σε [77]: α) Προσέγγιση Διακριτών Συναισθημάτων – ΠΔΣ (Discrete Emotion – DE) και β) Διαστατική Προσέγγιση των Συναισθημάτων – ΔΠΣ (Emotion Dimensions – ED).

Κατά την ΠΔΣ ζητείται από τους συμμετέχοντες στην πειραματική διαδικασία να αξιολογήσουν τα συναισθήματα τους με τη χρήση λεκτικών εκφράσεων (verbal expressions) που περιγράφουν καλύτερα τα συναισθήματα τους ή με την ανάδραση σε μια κλίμακα επιδεικνύοντας εάν το συναίσθημα που βίωσαν ήταν ασθενές ή δυνατό ή τη χρήση αναλογικής κλίμακας για να δείξουν την ένταση ενός συναισθήματος που βίωσαν. Η συγκεκριμένη πειραματική διαδικασία είναι τυποποιημένη, αλλά πολλοί ερευνητές προτιμούν να αναπτύσσουν τις δικές τους συναισθηματικές κατηγορίες. Τέτοιες προσεγγίσεις οδηγούν σε αναντιστοιχία των κατηγοριών και μπορεί να παρουσιάζονται προβλήματα κατά τη σύγκριση των αποτελεσμάτων [77]. Επίσης,

υπάρχουν προβλήματα στη συγκριτική στατιστική ανάλυση μεταξύ των μελετών (cross-study), λόγω της πληθώρας των δεδομένων που μπορεί να λείπουν.

Οι συμμετέχοντες στην πειραματική διαδικασία, χρησιμοποιώντας τη ΔΠΣ, καλούνται π.χ. να δείξουν πόσο θετικά (ευχάριστα) ή αρνητικά (δυσάρεστα) και πόσο έντονα (ενθουσιώδεις) ή ήρεμα νιώθουν. Τα συναισθήματα μπορούν να αντιστοιχηθούν χρησιμοποιώντας το δυσδιάστατο χώρο «Ευχάριστο» – «Έντονο». Αυτή η μέθοδος είναι απλή και αξιόπιστη [77]. Επίσης, μπορεί να πραγματοποιηθεί απλή (single study) ή προχωρημένη (cross-study) στατιστική ανάλυση σε αντίθεση με τις προσεγγίσεις ΜΕΛΑ και ΠΔΣ. Σε αυτή την προσέγγιση είναι δύσκολο να διαφοροποιηθούν η ένταση του αισθήματος από τη σωματική διέγερση [77].

1.5.3 Δοκιμασία αυτοαξιολόγησης με τη χρήση φιγούρων - Self Assessment Manikin Test

Το 1985, ο P.J. Lang παρουσίασε την πειραματική δοκιμασία που ονομάζεται Self Assessment Manikin Test (SAM) - Δοκιμασία Αυτοαξιολόγησης με τη Χρήση Φιγούρων [91], με την οποία δίνεται η δυνατότητα να αποφύγουμε τη λεκτική έκφραση των συναισθημάτων κατά την αξιολόγηση, εισάγοντας με αυτό τον τρόπο μια γρήγορη και εύκολη πειραματική διαδικασία. Επίσης, χρησιμοποιώντας εικόνες για την αξιολόγηση αντί για λέξεις, η δοκιμασία SAM καθίσταται διαπολιτισμική και ανεξάρτητη γλώσσας [92] [93].

Με αυτή τη δοκιμασία καταγράφεται η συναισθηματική κατάσταση (emotional state) των συμμετεχόντων μετά την έκθεση τους σε διάφορα ερεθίσματα και απεικόνισή της σε τρεις άξονες. Αυτοί είναι:

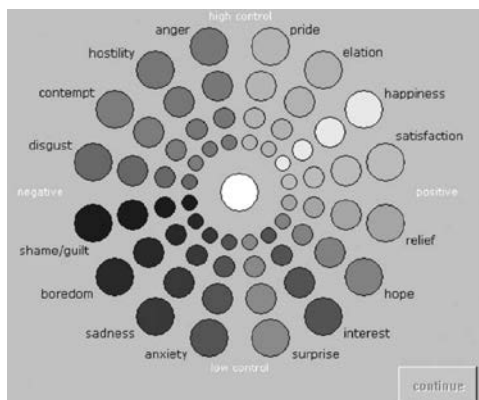
- «Ευχαρίστηση» (Pleasure, άλλοι όροι που χρησιμοποιούνται evaluation, valence). Θετικά ή αρνητικά συναισθήματα.
- «Διέγερση» (Arousal, άλλοι όροι που χρησιμοποιούνται activation, activity). Δυναμικότητα συναισθηματικής κατάστασης.
- «Επικράτηση» (Dominance, άλλοι όροι που χρησιμοποιούνται power, potency, controlled). Έλεγχος της συναισθηματικής κατάστασης.

Επίσης, η πειραματική διαδικασία είναι γνωστή και ως «PAD test», από τα αρχικά κάθε διάστασης. Κατά το σχεδιασμό και τη διεξαγωγή της πειραματικής διαδικασίας μπορεί να χρησιμοποιηθούν και οι εναλλακτικές-συνώνυμα κάθε διάστασης.

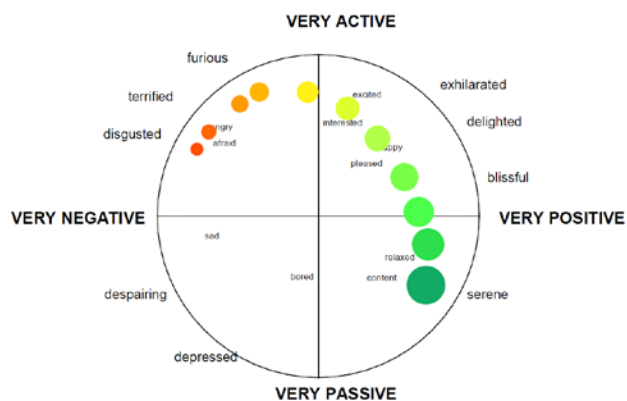
Οι συμμετέχοντες στο πείραμα μπορούν να επιλέξουν μία εκ των πέντε (τουλάχιστον) φιγούρων (manikins). Όπως αναφέρθηκε και παραπάνω δεν υπάρχουν λεκτικές εκφράσεις των συναισθημάτων για την αξιολόγηση των συμμετεχόντων. Σε μερικές περιπτώσεις χρησιμοποιείται μεγαλύτερη κλίμακα, 9 έως 25 σημείων [92] [94]. Για παράδειγμα, στην κλίμακα 9 σημείων (Σχήμα 2) υπάρχουν πέντε φιγούρες επιλογής και τέσσερα σημεία ενδιάμεσων τιμών ανάμεσα σε αυτές. Για την αξιολόγηση της διάστασης «Ευχαρίστηση», οι ακραίες τιμές αποτελούνται από μια χαμογελαστή φιγούρα και μια λυπημένη (συνοφρυωμένη). Αντίστοιχα, για τη διάσταση της «Διέγερσης», τα άκρα αναπαρίσταται από μία φιγούρα υψηλής ενεργητικότητας και μία ήρεμη, με τα μάτια κλειστά. Όμοια για την «Επικράτηση» το άκρο ελεγχόμενος από το συναίσθημα αναπαρίσταται από μία μικρή φιγούρα ενώ ο μη-ελεγχόμενος από μία μεγάλη. Οι αξιολογήσεις-απαντήσεις των συμμετεχόντων μπορούν πολύ εύκολα να μετατραπούν από την κλίμακα σημείων σε σύνολο τιμών με πεδία $[-1, 1]$ ή $[-100\%, 100\%]$. Χρησιμοποιώντας την εκατοστιαία κλίμακα, είναι ευδιάκριτη η μεταβολή της συναισθηματικής κατάστασης από την ουδέτερη (η τιμή «μηδέν» αντιπροσωπεύει την ουδέτερη κατάσταση σε κάθε διάσταση).

Χρησιμοποιώντας τον τρισδιάστατο χώρο των συναισθηματικών καταστάσεων, μας δίνεται η δυνατότητα να τις αντιστοιχήσουμε σε συγκεκριμένα συναισθήματα. Ένα γνωστό παράδειγμα είναι το κυκλικό σύμπλεγμα – διάταξη του Russell [77] [95]. Οι δύο διαστάσεις της «Ευχαρίστησης» και της «Διέγερσης» απεικονίζονται στο πλέγμα X-Y, αντίστοιχα στο οποίο τοποθετήθηκαν οι λεκτικές εκφράσεις των συναισθημάτων. Το Σχήμα 3 παρουσιάζει τον κύκλο με τη λεκτική - σημασιολογική αντιστοίχιση των συναισθημάτων.

Άλλη μια έκδοση του κύκλου είναι το Geneva Emotion Wheel (GEW) [96] [77] και το FeelTrace [97] που εισήχθησαν ως πειραματικά εργαλεία. Δύο στιγμιότυπα παρουσιάσής τους στην οθόνη του υπολογιστή φαίνονται στο Σχήμα 4.



(α)



(β)

Σχήμα 4: Δύο στιγμιότυπα από α) Geneva Emotion Wheel (GEW) και β) FeelTrace.

1.6 Συναισθήματα και έγγραφα

Τα συναισθήματα του αναγνώστη επηρεάζονται από τα έγγραφα με δύο τρόπους: α) μέσω του περιεχομένου, χρησιμοποιώντας συγκεκριμένες λέξεις ή φράσεις για την επίτευξη συγκεκριμένου συναισθήματος, β) μέσω του τρόπου παρουσίασης, χρησιμοποιώντας π.χ. διάφορα τυπογραφικά στοιχεία ή χρώματα υποβάθρου. Να σημειώσουμε σε αυτό το σημείο ότι η επιρροή της συναισθηματικής κατάστασης του αναγνώστη μπορεί να επιτευχθεί και με το συνδυασμό των δύο παραπάνω τρόπων. Η έκφραση των συναισθημάτων μέσω του περιεχομένου είναι κυρίως ευθύνη του συγγραφέα του εγγράφου, σε αντίθεση με τον τρόπο παρουσίασης του εγγράφου που είναι κυρίως ευθύνη του σχεδιαστή σελίδας (page designer) κατά τη διαδικασία έκδοσης. Σύμφωνα με τη θεωρία του James και το παράδειγμα που παρουσιάστηκε στην παράγραφο 1.5, θεωρώντας τον αναγνώστη σαν «μαύρο κουτί», τα συναισθήματα του (έξοδος) μπορούν να μεταβληθούν και εν γένει να επηρεαστούν από τα οπτικά χαρακτηριστικά του εγγράφου (είσοδος).

Σε αυτό το μέρος του κεφαλαίου θα παρουσιαστούν διάφορες μελέτες με σκοπό την παρατήρηση του τρόπου επιρροής των συναισθημάτων του αναγνώστη είτε μέσω του περιεχομένου είτε μέσω του τρόπου παρουσίασης του εγγράφου. Μέσω αυτής της βιβλιογραφικής αναδρομής θα μπορέσουμε να δούμε τα ερευνητικά κενά και εμπόδια που υπάρχουν και πώς η παρούσα διατριβή συνεισφέρει στην επίλυση των συγκεκριμένων προβλημάτων (παράγραφος 1.8).

1.6.1 Επαγόμενα συναισθήματα του αναγνώστη από το περιεχόμενο του εγγράφου

Αρκετές μελέτες εστιάζονται στη σημασιολογική εξαγωγή και τη μοντελοποίηση επαγόμενων συναισθημάτων από το περιεχόμενο των εγγράφων. Στην εργασία [98] παρουσιάζεται μια εφαρμογή πραγματικού χρόνου, η οποία ενισχύει τη διαπροσωπική επικοινωνία μέσω της αναγνώρισης συγκεκριμένων συναισθημάτων (ευτυχία, λύπη, θυμός, φόβος, αηδία και έκπληξη) κατά τη δακτυλογράφηση του κειμένου. Η παρουσίαση του κειμένου στην οθόνη του αποδέκτη γίνεται με την κατάλληλη εικόνα προσώπου-φιγούρας. Μία εφαρμογή για την αυτόματη εξαγωγή συναισθημάτων (Text-to-Emotion engine) βασισμένη στην επισημείωση των λέξεων και των προτάσεων του κειμένου, έχει αναπτυχθεί από τους Zhe et al. [99]. Οι Liu et al. [100] μελέτησαν την επίδραση του κειμένου στα συναισθήματα του αναγνώστη και οι Wu et al. [101] παρουσιάζουν μία προσέγγιση ανεξαρτήτου τομέα για την αυτόματη αναγνώριση συναισθημάτων (τριών καταστάσεων: χαρούμενος, λυπημένος και ουδέτερος) από το κείμενο για την Κινεζική γλώσσα. Οι Min και Park [102] προτείνουν ένα αριθμό από διακριτά χαρακτηριστικά που πηγάζουν από το κείμενο (για τα Κορεάτικα) και την κατάλληλη απεικόνιση τους χρησιμοποιώντας ετικέτες, τη διάρκεια και την ένταση. Βασισμένοι σε εικονικούς ή πραγματικούς χαρακτήρες (avatars), οι Zhixiao et al. [103] εισήγαγαν την έννοια της ενσωμάτωσης μηνυμάτων κειμένου για την κινεζική γλώσσα (simplified Chinese), με σκοπό την αποδοτική έκφραση μη-λεκτικών πληροφοριών, όπως τα ανθρώπινα συναισθήματα που ενσωματώνονται στο κείμενο. Οι Gill et al. [104] μελέτησαν την ικανότητα αξιολογητών να ανιχνεύουν μία από τις 8 συναισθηματικές κατηγορίες σε κείμενο 50 και 200 λέξεων από δείγματα κειμένου (weblog text). Οι Hancock et al. [105] μελέτησαν τον τρόπο που οι άνθρωποι εκφράζουν και ανιχνεύουν τα συναισθήματα κατά την επικοινωνία με κείμενο, όταν εξαλείφονται τα μη-λεκτικά ερεθίσματα (cues) που τυπικά συνδέονται με τα συναισθήματα (όπως χειρονομίες, εκφράσεις προσώπου κ.α.). Με τη χρήση Supervised Machine Learning για την πρόβλεψη συναισθημάτων από το κείμενο, γίνεται μία προσπάθεια στην εργασία [106] να ταξινομηθεί η συναισθηματική συγγένεια των προτάσεων (emotional affinity of sentences) σε αφηγηματικού τύπου κείμενα (παραμύθια για παιδιά), με σκοπό την ενσωμάτωση των αποτελεσμάτων στην εκφραστική συναισθηματική ομιλία. Στη μελέτη τους χρησιμοποιούν συγκεκριμένα θετικά και αρνητικά συναισθήματα όπως: θυμωμένος, αηδιασμένος, έντρομος, χαρούμενος και λυπημένος.

Ο Read [107] προτείνει μια καινοτόμα προσέγγιση, η οποία βασίζεται στη χρήση της γλώσσας σε συνδυασμό με τα συναισθήματα, η οποία μπορεί να εκπαιδεύσει ένα ταξινομητή και να λειτουργεί ανεξάρτητα από τον τομέα, το κεφάλαιο και το χρόνο. Οι Cowie et al. [108] περιγράφουν μια εμπειρική προσέγγιση για τον προσδιορισμό του είδους της διεργασίας για την ορθή απόδοση ενός συστήματος αναγνώρισης συναισθημάτων. Τρία επίπεδα πληροφορίας προκύπτουν: ένα βασικό λεξικό συναισθημάτων, μία βασική απεικόνιση στο χώρο «Ευχαρίστηση» - «Διέγερση» της έννοιας κάθε λέξης και ένα πιο πλούσιο σχήμα απεικόνισης. Οι Owsley et al. [109] προτείνουν ένα σύνολο τεχνικών που μπορούν να χρησιμοποιηθούν για την κατηγοριοποίηση ιστολογίων (weblogs) σύμφωνα με το συναισθηματικό περιεχόμενο. Αντί να χρησιμοποιηθεί μια γενικού σκοπού στρατηγική για τη συναισθηματική κατηγοριοποίηση, η προσέγγισή τους στοχεύει στην δημιουργία ταξινομητών αισθημάτων για συγκεκριμένους τομείς. Οι Lin et al. [110] κατηγοριοποίησαν άρθρα ειδήσεων για την Κινεζική γλώσσα, σε οχτώ συναισθήματα: χαρούμενος, θυμωμένος, λυπημένος, έκπληκτος, ευχάριστο (heart warming), φοβερός (awesome), βαρετός και χρήσιμος.

1.6.2 Επαγόμενα συναισθήματα του αναγνώστη από τον τρόπο παρουσίασης του εγγράφου

Ο χρωματικός συνδυασμός του κειμένου και του υποβάθρου σε μια ιστοσελίδα επιδρά στην αναγνωσιμότητα και στην αισθητική του εγγράφου [5] [13] [14]. Έτσι, ένα σωστά και κατάλληλα σχεδιασμένο έγγραφο του παγκόσμιου ιστού μπορεί να γίνει φιλικό στον αναγνώστη [111]. Οι Hall και Hanna [5] μελέτησαν την επίδραση των χρωματικών συνδυασμών του κειμένου και του υποβάθρου των ιστοσελίδων στα παρακάτω: αναγνωσιμότητα, μνημονικό (retention), αισθητική και πρόθεση συμπεριφοράς (behavioral intention). Ένα σημαντικό εύρημα είναι ότι οι προτιμώμενοι χρωματικοί συνδυασμοί οδηγούν σε υψηλή αξιολόγηση της αισθητικής. Οι Éthier et al. [112] μελετούν την επίδραση των χαρακτηριστικών της διεπαφής από τέσσερις ιστοσελίδες στη γνωσιακή διαδικασία που ενεργοποιούνται τα συναισθήματα των καταναλωτών μέσω διαδικτύου. Η συγκεκριμένη εργασία επικεντρώνεται στη μελέτη έξι συγκεκριμένων συναισθημάτων που βιώνονται κατά την αλληλεπίδραση με μια ιστοσελίδα.

Ο Laarni [16] μελέτησε την επίδραση του χρώματος, του είδους και του τύπου της γραμματοσειράς στις προτιμήσεις των χρηστών. Συμπέρανε ότι 1) οι πιο ευανάγνωστοι συνδυασμοί είναι 1α) απλό κείμενο, «Times New Roman», με μαύρη γραμματοσειρά σε

άσπρο υπόβαθρο, Ιβ) πλάγια γραφή, «Arial» με λευκή γραμματοσειρά σε μπλε υπόβαθρο και γ) απλό κείμενο, «Arial» με άσπρη γραμματοσειρά σε μαύρο υπόβαθρο, και (II) οι λιγότερο ευανάγνωστοι συνδυασμοί είναι κόκκινη γραμματοσειρά σε πράσινο υπόβαθρο. Επίσης, μελέτησε τη επίδραση του χρώματος στην αισθητική του εγγράφου, για παράδειγμα συνδυάζοντας κόκκινη γραμματοσειρά σε πράσινο υπόβαθρο αξιολογήθηκε ως ο πιο δυσάρεστος συνδυασμός και μαύρο σε άσπρο θεωρήθηκε ως ο λιγότερο έντονος. Εξάλλου, άλλες μελέτες έδειξαν πώς πολλαπλοί συνδυασμοί χρωμάτων [113], μεγέθους, είδους και τύπου γραμματοσειράς σε ένα έγγραφο επηρεάζουν τη συναισθηματική κατάσταση του αναγνώστη [114] [115] [116] και την αναγνωσιμότητα του εγγράφου, όχι μόνο στην έντυπη μορφή αλλά και στην ηλεκτρονική [117].

Το TextTone [118] είναι ένα σύστημα που προσθέτει συναισθήματα στην απευθείας επικοινωνία με την χρήση κειμένου (online textual) αλλάζοντας τα γνωρίσματα (attributes) των τυπογραφικών στοιχείων. Υποστηρίζονται συναισθήματα (που αναφέρονται ως συναισθηματικοί τόνοι - emotional tones) όπως: χαρούμενος, πολύ χαρούμενος, παραγμένος, απογοητευμένος, θυμωμένος, πολύ θυμωμένος, σοκαρισμένος, μπερδεμένος κ.α. Το TextTone χρησιμοποιεί τα εξής τυπογραφικά στοιχεία: μέγεθος, χρώμα, είδος και τύπος γραμματοσειράς, έντονη, πλάγια γραφή, υπογραμμισμένη και διαγραφή.

Οι Ohene-Djan et al. [119, 120] παρουσιάζουν τον τρόπο με τον οποίο οι εναλλαγές των τυπογραφικών στοιχείων μπορούν να χρησιμοποιηθούν για τη μετάδοση συναισθημάτων στους υπότιπλους, επικεντρώνοντας στους κωφούς και τα άτομα με προβλήματα ακοής. Στη μελέτη τους χρησιμοποίησαν το χρώμα και το μέγεθος της γραμματοσειράς για να μεταδώσουν συναισθήματα όπως, χαρά, λύπη, σαρκασμός κ.α. Η Yannicopoulou [121] μελέτησε πώς τα παιδιά προσχολικής ηλικίας αναγνωρίζουν συμβάσεις αναλογίας, για παράδειγμα η μεταβολή της έντασης της φωνής οπτικοποιείται με το μέγεθος της γραμματοσειράς, ή συμβατικές οπτικές αλληγορίες (conventional visual metaphors) αντιστοιχούν σε διάφορα συναισθήματα. Επίσης, ο Rosenberger [122, 123] εισήγαγε τον όρο «prosodic font», ο οποίος παίρνει τη χρονική του μορφή (temporal form) από συγκεκριμένες προσωδιακές παραμέτρους της φωνής. Συγκεκριμένα, το «prosodic font» εξομοιώνει την τονική και ρυθμική κίνηση της φωνής [123] και είναι συνδεδεμένο με τη δυναμική και κινητική τυπογραφία (dynamic and kinetic typography).

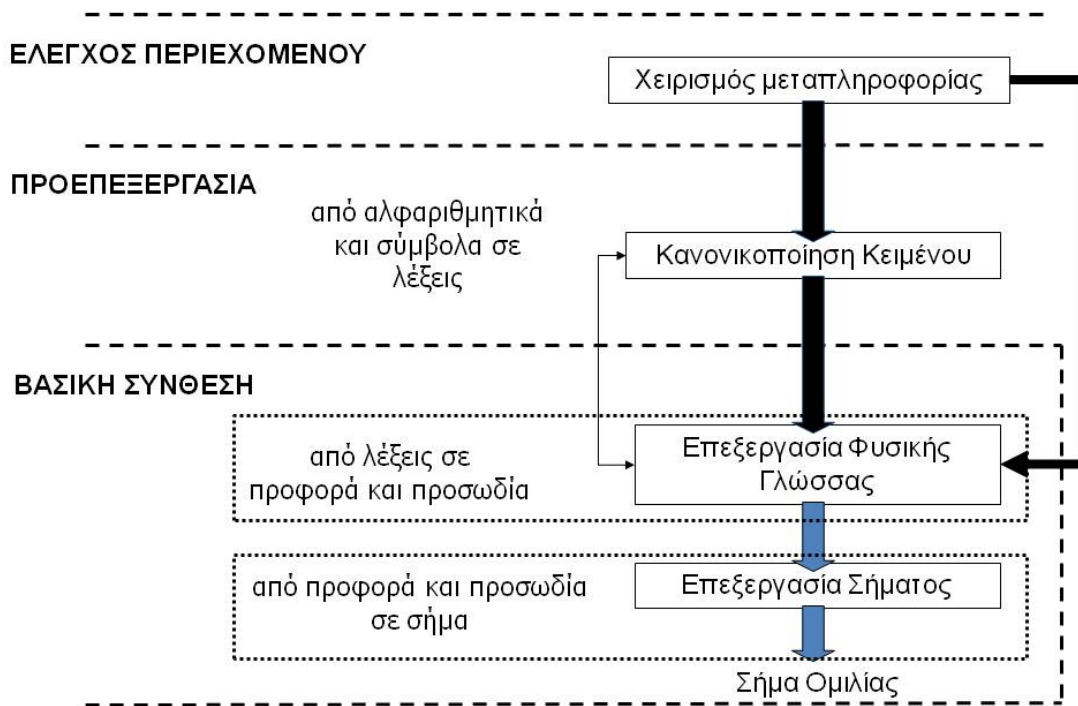
Όλες αυτές οι μελέτες, προτείνουν κυρίως τον τρόπο με τον οποίο μπορούν να χρησιμοποιηθούν τα τυπογραφικά στοιχεία. Έτσι, μπορούν να χαρακτηριστούν περισσότερο ως ποιοτικές παρά ως ποσοτικές. Στην εργασία [124] (που δημοσιεύτηκε στο πλαίσιο της παρούσας διατριβής) προτείνεται μια μαθηματική περιγραφή του τρόπου με τον οποίο μεταβάλλεται η συναισθηματική κατάσταση σύμφωνα με τις τυπογραφικές μεταβολές, με σκοπό τη δημιουργία ενός αυτόματου συστήματος εξαγωγής της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη και στη συνέχεια την κατάλληλη επισημείωση των εγγράφων [125] (όλα όσα αναφέρονται παρουσιάζονται αναλυτικά στο Κεφάλαιο 3). Για παράδειγμα, περιγράφεται η επίδραση των εναλλαγών του μεγέθους της γραμματοσειράς (συνεχείς τιμές) στις διαστάσεις της επαγόμενης συναισθηματικής κατάστασης «Ευχαρίστηση», «Διέγερση» και «Επικράτηση», χρησιμοποιώντας δευτέρου και τρίτου βαθμού εξισώσεις. Για την περίπτωση, του τρόπου γραφής (διακριτές τιμές), δημιουργείται ένα σύνολο κανόνων για την αντιστοίχιση των εναλλαγών των τυπογραφικών στοιχείων από το απλό κείμενο.

1.7 Εκφραστική Συνθετική Ομιλία

«Σύνθεση ομιλίας» είναι η δημιουργία τεχνητών (συνθετικών) εκφωνημάτων μέσω κατάλληλων αλγορίθμων, οι οποίοι υλοποιούνται με κάποια υπολογιστική διάταξη (π.χ. προσωπικό ηλεκτρονικό υπολογιστή) [126]. Υπάρχουν τρεις τυπικές μέθοδοι που μπορούν να παράγουν διαφορετικών βαθμών αποδεκτή συνθετική ομιλία: η αρθρωτική σύνθεση (articulatory synthesis), η σύνθεση φωνοσυντονισμών (formant synthesis) και η σύνθεση μέσω συρραφής (concatenative synthesis) [126] [127].

Η μετατροπή Κειμένου-σε-Ομιλία είναι μία σύνθετη διαδικασία η οποία πραγματοποιείται σε διάφορα στάδια. Στο Σχήμα 5 παρουσιάζονται τα τρία βασικά βήματα:

1. Έλεγχος του περιεχόμενου.
2. Προεπεξεργασία κειμένου.
3. Βασική σύνθεση



Σχήμα 5: Το διάγραμμα διαδικασίας μετατροπής Κειμένου-σε-Ομιλία.

Στα δύο πρώτα στάδια γίνεται η επεξεργασία του κειμένου ώστε να εξαχθούν πληροφορίες που μπορούν να επηρεάσουν την επεξεργασία της φυσικής γλώσσας του περιεχομένου και κατ' επέκταση την αντιστοίχιση των λέξεων-φράσεων-προτάσεων στα προσωδιακά χαρακτηριστικά. Τέλος, στο κομμάτι της βασικής σύνθεσης πραγματοποιείται η προσωδιακή περιγραφή (με τη χρήση προσωδιακών κανόνων - μοντέλων) του κειμένου και η μετατροπή του σε σήμα ομιλίας.

Το πρόβλημα με τα περισσότερα υπάρχοντα συστήματα μετατροπής Κειμένου-σε-Ομιλία, είναι ότι δε λαμβάνουν υπόψη την επιπλέον πληροφορία (μεταπληροφορία – μεταδεδομένα) που αφορά την παρουσίαση του κειμένου. Δηλαδή, το πρώτο στάδιο του Σχήματος 5 («Χειρισμός μεταπληροφορίας») ουσιαστικά παρακάμπτεται, «απογυμνώνοντας» το έγγραφο από κάθε σχετική πληροφορία. Στις [128] [129] παρουσιάζεται μία προσπάθεια για να λαμβάνεται υπόψη αυτή η μεταπληροφορία, αλλά χρησιμοποιούν συγκεκριμένους κανόνες αντιστοίχισης των οπτικών χαρακτηριστικών (περιορισμένου αριθμού) σε προσωδιακά χαρακτηριστικά.

Η «Εκφραστική Ομιλία» είναι η ομιλία που μας δίνει πληροφορίες, εκτός από το απλό μήνυμα, και για τον ομιλητή και επιπλέον δημιουργεί-διεγείρει μία απόκριση στον ακροατή [130]. Υπάρχει μια πληθώρα μελετών για τη δημιουργία-υλοποίηση

«εκφραστικής συνθετικής ομιλίας» (*expressive speech synthesis*) με σκοπό τη δημιουργία φυσικού αποτελέσματος κατά την αλληλεπίδραση ανθρώπου-μηχανής. Το συναίσθημα νοείται ως ένας τρόπος έκφρασης, και σύμφωνα με τον Campbell [131] ορίζεται ως μία μέθοδος για τη μετάδοση συναισθημάτων (και άλλων παραγλωσσολογικών πληροφοριών) μέσω της ομιλίας, χρησιμοποιώντας μεταβολές και διαφοροποιήσεις των χαρακτηριστικών της φωνής. Ο ίδιος γενικά αναφέρει ότι [132]: η εκφραστική συνθετική ομιλία αφορά την προσθήκη πρόσθετων χαρακτηριστικών που σχετίζονται με τη συναισθηματική κατάσταση του (συνθετικού) ομιλητή όπως επίσης και τα συγκεκριμένα χαρακτηριστικά που αναγνωρίζουν τον ομιλητή στην ουδέτερη συνθετική ομιλία. Έτσι, πτυχές όπως τα συναισθήματα και η ποιότητα ομιλίας πρέπει να λαμβάνονται υπόψη κατά τη σύνθεση εκφραστικής ομιλίας.

Πολλές μελέτες εστιάζονται στην έκφραση συγκεκριμένων συναισθημάτων που μπορούν να εξαχθούν από τα χαρακτηριστικά της ομιλίας [133] [134] [101]. «Ποικιλία τρόπων έκφρασης» είναι ένας όρος που χρησιμοποιείται για τη μελέτη που εξαρτάται από το θεματικό πεδίο. Για παράδειγμα, χρησιμοποιώντας εκφραστική συνθετική ομιλία μπορούμε να μεταδώσουμε μηνύματα, όπως «καλά-άσχημα νέα», «ερωτήσεις ναι-όχι» [135], εξιστόρησης (storytelling) [136] και εκφράσεις στρατιωτικού τρόπου [137].

1.7.1 Διαστατική προσέγγιση της εκφραστικής συνθετικής ομιλίας

Οι θεωρητικές μελέτες των Schröder [138] [139] και των Tatham και Morton [130] παρουσιάζουν μια διεξοδική ανάλυση των διάφορων τεχνικών και μεθοδολογιών για την εκφραστική συνθετική ομιλία. Για αυτό το λόγο, θα αναφερθούν περιληπτικά οι προσεγγίσεις με σκοπό να παρουσιαστούν οι διαφορές, τα πλεονεκτήματα και τα μειονεκτήματα της κάθε μεθοδολογίας.

Η Εκφραστική Συνθετική Ομιλία μπορεί να υλοποιηθεί χρησιμοποιώντας συνθέτες φωνοσυντονισμών, τη σύνθεση μέσω συρραφών, την αρθρωτική σύνθεση, με την επιλογή φωνητικών μονάδων ή με σαφή έλεγχο της προσωδίας [138] [139] [130]. Στη σύνθεση φωνοσυντονισμών το αποτέλεσμα που προκύπτει είναι αφύσικο συγκρινόμενο με τη σύνθεση μέσω συρραφών [138]. Στα περισσότερα συστήματα σύνθεσης διφώνων, μπορούν να ελεγχθούν μόνο η βασική συχνότητα (F0) και η διάρκεια (και πιθανόν η ένταση). Είναι συνήθως αδύνατο να ελεγχθεί η ποιότητα της φωνής [138] που καθορίζεται από την απόδοση (rendition) του ομιλητή κατά την ηχογράφηση των διφώνων [139]. Η πιο φυσική τεχνική σύνθεσης θεωρείται η επιλογή φωνητικών μονάδων, ή η σύνθεση με τη χρήση μεγάλης βάσης δεδομένων. Παρόλο που αυτός ο

τρόπος σύνθεσης δίνει πολύ φυσικά αποτελέσματα, μπορεί να προκύπτουν και πολύ άσχημα όταν δε βρίσκονται οι κατάλληλες φωνητικές μονάδες. Η μέθοδος συρραφής κυματομορφής έρχεται δεύτερη - έπεται της σύνθεσης φωνοσυντονισμών, όταν ασχολούμαστε με το χειρισμό του σήματος για εκφραστικό περιεχόμενο [130]. Επίσης, έχουν προταθεί μεθοδολογίες [139] με σκοπό τη βελτιστοποίηση των ήδη υπαρχόντων, καθώς και νέες προσεγγίσεις, όπως επιλογή μονάδων με βάση την εκφραστικότητα (expressivity-based selection of units), επιλογή μονάδων και τροποποίηση του σήματος και στατιστικώς παραμετρική σύνθεση βασισμένη σε μοντέλα Hidden Markov [140].

Μερικές μελέτες για τη μοντελοποίηση της εκφραστικής συνθετικής ομιλίας χρησιμοποιούν τη διαστατική προσέγγιση των συναισθημάτων [141] [142]. Το πλεονέκτημα χρήσης αυτής της μεθοδολογίας είναι ότι οι τρεις διαστάσεις παίρνουν συνεχείς τιμές. Επίσης, οι τιμές αυτές μπορούν να αντιστοιχηθούν σε συγκεκριμένα συναισθήματα (ή μεταβολές τους). Για παράδειγμα, το συναίσθημα «χαρούμενος» μπορεί να έχει μεταβολές όπως «λίγο χαρούμενος», «πολύ χαρούμενος», «λιγότερο χαρούμενος».

Ο Schröder [142] παρουσιάζει ένα μοντέλο για τη δημιουργία εκφραστικής συνθετικής ομιλίας χρησιμοποιώντας τη διαστατική προσέγγιση των συναισθημάτων. Το μοντέλο υλοποιήθηκε και δοκιμάστηκε χρησιμοποιώντας το σύστημα MARY TTS [143]. Ο ίδιος προτείνει ότι η φωνητική έκφραση των συναισθημάτων είναι παρόμοια μεταξύ των γλωσσών, όπως παρουσιάζεται διεξοδικά στην επόμενη παράγραφο (1.7.2). Έτσι, ακολουθώντας τη θεωρητική μελέτη μπορούμε να χρησιμοποιήσουμε το προτεινόμενο μοντέλο και στην ελληνική γλώσσα.

Με τη χρήση διαφόρων εξισώσεων περιγράφεται ο τρόπος μεταβολής των προσωδιακών χαρακτηριστικών, καθώς μεταβάλλονται οι συναισθηματικές καταστάσεις.

Οι παράμετροι της ομιλίας διακρίνονται ως:

- Πρότυπες καθολικές (Standard global) παράμετροι: τονικό ύψος, εύρος, ρυθμός και ένταση.
- Μη-κλασικές (Non-standard) καθολικές παράμετροι: δυναμική τονικού ύψους (pitch-dynamics) και δυναμική εύρους (range-dynamics) .
- Συγκεκριμένες οντότητες (entities): όπως τονισμός (accents) και όρια GToBI (boundaries).

Και οι δύο εξαρτημένες (προσωδικά στοιχεία) και ανεξάρτητες (συναισθηματικές διαστάσεις) μεταβλητές είναι συνεχείς και συνοψίζονται χρησιμοποιώντας την παρακάτω σχέση:

$$\Delta S = A * \Delta E + I \quad (1)$$

όπου:

$$\Delta S = \begin{bmatrix} \Delta S_1 \\ \Delta S_2 \\ \dots \\ \Delta S_n \end{bmatrix}, A = \begin{bmatrix} a_1^P & a_1^A & a_1^D \\ a_2^P & a_2^A & a_2^D \\ \dots & \dots & \dots \\ a_n^P & a_n^A & a_n^D \end{bmatrix}, I = \begin{bmatrix} I_1 \\ I_2 \\ \dots \\ I_n \end{bmatrix}, \Delta E = \begin{bmatrix} \Delta P \\ \Delta A \\ \Delta D \end{bmatrix}$$

ΔS : Ο πίνακας των χαρακτηριστικών της ομιλίας (προσωδικά)

ΔP : Οι μεταβολές της «Ευχαρίστησης» με πεδίο ορισμού [-100%, 100%];

ΔA : Οι μεταβολές της «Διέγερσης» με πεδίο ορισμού [-100%, 100%];

ΔD : Οι μεταβολές της «Επικράτησης» με πεδίο ορισμού [-100%, 100%];

A : Ο πίνακας των συντελεστών της εξίσωσης

I : Ο πίνακας των σταθερών

Είναι φανερό ότι χρησιμοποιώντας το συγκεκριμένο μοντέλο μπορούμε: α) να εκφράσουμε ασθενή συναισθήματα και β) να συμπεριλάβουμε το χρόνο στις συναισθηματικές μεταβολές-εναλλαγές [130].

Στην παρούσα μελέτη χρησιμοποιούνται τρία βασικά προσωδικά χαρακτηριστικά: τονικό ύψος, ρυθμός και ένταση. Ο τρόπος που μεταβάλλονται από τη βάση αναφοράς (baseline) παρουσιάζεται στην ακόλουθη εξίσωση [142]:

$$\begin{bmatrix} \Delta Pitch \\ \Delta Rate \\ \Delta Volume \end{bmatrix} = \begin{bmatrix} 0.1 & 0.3 & -0.1 \\ 0.2 & 0.5 & 0 \\ 0 & 0.33 & 0 \end{bmatrix} \cdot \begin{bmatrix} \Delta P \\ \Delta A \\ \Delta D \end{bmatrix} \quad (2)$$

1.7.2 Φωνή και συναισθήματα: Οι παράγοντες της γλώσσας και του πολιτισμού

Στην εργασία [144] προτείνεται ο καθολικός χαρακτήρας των συναισθημάτων. Τα βασικά ευρήματα των Scherer et al. [145] είναι ότι κριτές-αξιολογητές από εννέα χώρες, που μιλούσαν διαφορετικές γλώσσες, μπόρεσαν να αντιληφθούν τέσσερα διαφορετικά συναισθήματα και την ουδέτερη κατάσταση από φωνητικές απεικονίσεις που χρησιμοποιούσαν ομιλία ανεξάρτητης περιεχομένου. Οι διαφοροποιήσεις των ποσοστιαίων συναισθηματικών μεταβολών και τα πρότυπα λάθους (error patterns) είναι όμοια μεταξύ των χωρών. Επίσης, υπάρχουν ενδείξεις ότι γλωσσολογικές και/ή πολιτισμικές ομοιότητες μπορούν να παίξουν σημαντικό ρόλο στη φωνητική αναγνώριση των συναισθημάτων. Επεκτείνοντας αυτή τη δουλειά, οι Pell et al. [146] παρουσιάζουν τα πειραματικά αποτελέσματα της μελέτης αναγνώρισης έξι συναισθημάτων για τέσσερις γλώσσες (Αγγλικά, Γερμανικά, Ινδικά και Αραβικά) και υποστηρίζουν ότι η φωνητική αναγνώριση των συναισθημάτων δεν επηρεάζεται από τη γλωσσική ή από γλωσσολογικές ομοιότητες. Επίσης, όπως αναφέρθηκε και προηγουμένως, ο Schröder [142] υποστηρίζει ότι η φωνητική έκφραση των συναισθημάτων είναι παρόμοια μεταξύ των γλωσσών.

Οι ακροατές - ανεξάρτητα γλώσσας - ερμηνεύουν τη σημασιολογία των φράσεων ορθά, πιθανόν με μία εξαίρεση τη φράση που εννοεί «βαρεμάρα» [147]. Στην ίδια εργασία υπήρχαν διαφορές μεταξύ διαφορετικών χωρών και δεν μπορούσαν να εκτιμήσουν εάν όλες οι αποκλίσεις βασιζόντουσαν μόνο στις πολιτισμικές διαφορές. Έτσι, εκτιμούν ότι μία διαπολιτισμική καθολική προσομοίωση δε θα λειτουργήσει όπως αναμένεται και τα ευρήματα υποδηλώνουν ότι τα αποτελέσματα βασισμένα στην ανάλυση δεδομένων από διαφορετικούς πολιτισμούς δεν μπορούν να εφαρμοστούν ανεπιφύλακτα. Αυτή η υπόθεση δεν μπορεί να στηριχθεί πλήρως, διότι οι αποκλίσεις μπορεί να οφείλονται σε άλλους παράγοντες πέραν των πολιτισμικών διαφορών (τα ερεθίσματα ήταν διαφορετικά μεταξύ των διαφορετικών πολιτισμών).

- Μπορεί οι διαφορές να οφείλονται στον ίδιο τον συνθέτη ομιλίας. Η σύνθεση υλοποιήθηκε με τη χρήση διφώνων διαφορετικών ομιλητών και η ποιότητα της ομιλίας μπορεί να ποικίλει για κάθε γλώσσα.
- Η προσωδία προέρχεται από διαφορετικούς ομιλητές και οι κανόνες παράγουν διαφορετικά αποτελέσματα.
- Η μετάφραση των προτάσεων μπορεί να είχε ως αποτέλεσμα διαφορετική σημασιολογία και έτσι να προκύπτει διαφορετικό συναίσθημα [147].

1.8 Στόχος Διατριβής

Η ακουστική απόδοση της οπτικής πληροφορίας των εγγράφων είναι μία σύνθετη διαδικασία. Με σκοπό να επιτευχθεί η συγκεκριμένη διεργασία, χρησιμοποιούνται μεθοδολογίες, όπως η χρήση συνθετικών ήχων μικρής διάρκειας (earcons) πριν ή/και μετά το κείμενο (που παρουσιάζεται με συγκεκριμένο τυπογραφικό τρόπο), ή με τη χρήση προσωδιακών μεταβολών στη συνθετική ομιλία. Αυτή η μέθοδος μπορεί να θεωρηθεί ως «*άμεσος τρόπος ακουστικής απόδοσης*» (direct acoustic rendition) των τυπογραφικών χαρακτηριστικών λόγω της άμεσης αντιστοίχισης μεταξύ τυπογραφίας και προσωδίας. Η σύνθεση Document-to-Audio (DtA) [128] θεωρείται ως η επόμενη γενιά των συστημάτων μετατροπής Κειμένου-σε-Ομιλία, υποστηρίζοντας την εξαγωγή της σημασιολογίας από τα μεταδεδομένα του εγγράφου (π.χ η χρήση της έντονης ή πλάγιας γραφής σε ελληνικές εφημερίδες μπορεί να υποδηλώνει έμφαση, σημείωση, τίτλο ταινίας κ.α.) [33] και την κατάλληλη ακουστική απεικόνιση της μορφοποίησης του κειμένου [148] [34] [35] [36] μέσω της μοντελοποίησης των παραμέτρων του σήματος της συνθετικής ομιλίας [149] [129], που θεωρείται ως «*έμμεσος τρόπος ακουστικής απόδοσης*» (indirect acoustic rendition).

Η W3C προτείνει το πρότυπο Aural Style Sheets [72] για την ακουστική απόδοση των τυπογραφικών στοιχείων των εγγράφων του παγκόσμιου ιστού χρησιμοποιώντας συνθετική ομιλία. Δεν παρέχονται όμως κανόνες αντιστοίχισης, π.χ. πώς πρέπει να μεταβληθούν τα προσωδιακά χαρακτηριστικά σε περίπτωση που έχουμε έντονη γραφή. Απλά δίνεται η περιγραφή του τρόπου με τον οποίο μπορούν τα έγγραφα να μετατραπούν από ένα πρότυπο σε ακουστική μορφή. Οι Kallinen et al. [150] εισάγουν τον όρο-μηχανισμό «auditive “boldfacing”». Εξετάζουν την επίδραση συγκεκριμένων μηχανισμών (για παράδειγμα μειώνοντας την ομιλία κατά δύο ημιτόνια) στην απόδοση της μνήμης για ακουστικές ειδήσεις με περιεχόμενο σχετικό με την επιχειρηματική δραστηριότητα. Αυτός ο μηχανισμός βελτίωσε την απόδοση της άμεσης μνήμης (immediate memory performance), αλλά μείωσε την απόδοση της μακροπρόθεσμης μνήμης (longer term performance).

Οι Truillet et al. [151] παρουσιάζουν μια πειραματική μελέτη για την αξιολόγηση των «ηχητικών γραμματοσειρών» (sound fonts). Μελετούν την ακουστική απόδοση των σημαντικότερων λέξεων χρησιμοποιώντας συνθετική ομιλία, που επιτυγχάνεται με τη λεκτική περιγραφή του τυπογραφικού στοιχείου (π.χ. εκφωνείται η φράση «σε έντονη» με μείωση κατά 15% του τονικού ύψους πριν από τη λέξη) ή με αύξηση 13% της τιμής του προεπιλεγμένου τονικού ύψους. Με τη χρήση διαφόρων ψυχο-ακουστικών

χειρισμών (μεταβολές του τονικού ύψους, έντασης και ρυθμού της συνθετικής ομιλίας), οι Argyropoulos et al. [152] μελέτησαν την αποτελεσματικότητα της κατανόησης συγκεκριμένων πληροφοριών (τυπογραφικών στοιχείων - έντονη και πλάγια γραφή) από 30 βλέποντες και τυφλούς συμμετέχοντες. Τα αποτελέσματα αυτής της εργασίας έδειξαν ότι η ακουστική απόδοση των τυπογραφικών χαρακτηριστικών (χρησιμοποιώντας συνθετική ομιλία) είναι κατανοητή από τους ακροατές. Οι Papadopoulos et al. [153] [154] [155] εξετάζουν την καταληπτότητα (intelligibility) και κατανόηση (comprehensibility) λέξεων με τη χρήση φυσικής φωνής ή/και συνθετικής ομιλίας σε τυφλούς και βλέποντες ακροατές. Στην [129] γίνεται μια προκαταρκτική μελέτη της ακουστικής απόδοσης της πληροφορίας για την τυπογραφία ή/και πληροφοριών στίξης χρησιμοποιώντας εκφραστική συνθετική ομιλία. Σκοπός αυτής της μελέτης είναι η βελτίωση της εκφραστικότητας του ήδη υπάρχοντος συστήματος συνθετικής ομιλίας της France Telecom χρησιμοποιώντας προσωδιακούς κανόνες. Προτείνονται η χρήση τεσσάρων προσωδιακών χαρακτηριστικών: τονικό ύψος, ρυθμός, ένταση και παύση. Η ακουστική απόδοση των εγγράφων LATEX έχει προταθεί από τον Raman [156]. Ανέπτυξε μία σειρά συστημάτων κυρίως για την παροχή ηχητικών μορφών σύνθετων μαθηματικών τύπων σε έγγραφα τύπου LaTeX. Χρησιμοποίησε ήχους μη-ομιλίας (non-speech audio) για να υποδηλώσει τη φόρμουλα, καθώς και μερικές στοιχειώδεις μεταβολές του τονικού ύψους για να ομαδοποιήσει στοιχεία μέσα στη φόρμουλα (π.χ. αριθμητής/παρονομαστής) χωρίς όμως να λαμβάνει υπόψη του, π.χ. την τυπογραφία.

Οι μελέτες που παρουσιάστηκαν, απλώς προτείνουν κανόνες για την ακουστική απόδοση των μεμονωμένων τυπογραφικών στοιχείων. Δεν υπάρχει μια συστηματική προσέγγιση. Για αυτό το λόγο είναι δύσκολη η υλοποίηση ενός αυτόματου συστήματος, το οποίο να υποστηρίζει αυτή τη λειτουργικότητα. Επίσης, τα διεθνή πρότυπα απλά παρέχουν τους *μηχανισμούς* για την ακουστική απόδοση και όχι τον *τρόπο αντιστοίχισης* των τυπογραφικών στοιχείων σε προσωδιακά χαρακτηριστικά.

Στην παρούσα διατριβή προτείνεται μια μεθοδολογική διεπιστημονική προσέγγιση για τη μελέτη και σχεδίαση ακουστικής (Acoustical User Interface - AUI) αλληλεπίδρασης στη μη-οπτική διεπαφή χρήστη (non-Visual User Interface) η οποία είναι ανεξάρτητη του περιεχομένου, της φυσικής γλώσσας και της μορφής ή του τύπου της τερματικής συσκευής αλληλεπίδρασης. Η έρευνα εστιάζεται στη μελέτη της δυνατότητας πλήρους ακουστικής προσβασιμότητας στις μεταπληροφορίες δομημένων έντυπων και ηλεκτρονικών εγγράφων. Συγκεκριμένα μελετάται η κατάλληλη τυπογραφική διάκριση

που θα πρέπει να παρέχεται στον χρήστη (βλέποντα, με μειωμένη όραση ή/και τυφλό) για κάθε ένα χαρακτηριστικό του εγγράφου που αποδίδεται ακουστικά, ώστε να επιτυγχάνεται μη λανθασμένη αντίληψή του.

Κάθε ακουστική αναπαράσταση προσδιορίζεται σε τρία επίπεδα: εννοιολογικό (conceptual), δομικό (structural) και υλοποίησης (implementation). Στο εννοιολογικό επίπεδο μελετάται ο εντοπισμός της πληροφορίας που πρέπει να αναπαρασταθεί ακουστικά και περιλαμβάνει τη μελέτη των δραστηριοτήτων που υποστηρίζονται, καθώς και την εξέταση των απαιτήσεων των αντικειμένων αλληλεπίδρασης που θα χρησιμοποιηθούν. Η πληροφορία αναπαρίσταται με ακουστικό τρόπο. Στο δομικό επίπεδο εξετάζεται πώς μια ακουστική σκηνή (scene) επηρεάζεται από το συνδυασμό διαφόρων τμημάτων πληροφορίας που απαιτεί μια δραστηριότητα, και πώς καθορίζονται οι πρωτεύοντες και συνεχόμενοι ακουστικοί όροι απόδοσης των οπτικών-δομικών χαρακτηριστικών των εγγράφων. Το επίπεδο υλοποίησης έχει σκοπό τον προσδιορισμό των μεταβολών της συνθετικής ομιλίας που είναι κατάλληλες για τη διατήρηση της πληροφορίας που αναγνωρίστηκε στο εννοιολογικό επίπεδο και την υλοποίηση της δομής που εξήχθη στο δομικό επίπεδο. Παράγεται ένα ολιστικό μοντέλο ακουστικής αναπαράστασης μεταπληροφορίας των εγγράφων που εντάσσεται στο πλαίσιο της Επαυξημένης Εικονικής Πραγματικότητας (Augmented Virtual Reality). Επίσης αναπτύσσονται κατευθυντήριες γραμμές για τους σχεδιαστές της διεπαφής χρήστη σε κάθε ένα από τα τρία επίπεδα που αναφέρθηκαν παραπάνω.

Τα αποτελέσματα κατατάσσονται σύμφωνα με δύο βασικά κριτήρια:

- ποιοι μηχανισμοί επιτυγχάνουν μεγαλύτερο βαθμό ακουστικής αντίληψης (αναγνώριση, διάκριση και κατανόηση) και
- ποιοι είναι συγχρόνως πιο ευκολομνημόνευτοι (μελέτη της ακουστικής μνήμης με βάση συγκεκριμένες ακουστικές αποδόσεις των οπτικών χαρακτηριστικών στα έγγραφα).

Η μελέτη του προβλήματος της ακουστικοποίησης των οπτικών πληροφοριών ενός εγγράφου, χωρίζεται ουσιαστικά σε τέσσερα βασικά μέρη:

1. Προσδιορισμός των μεταπληροφοριών, οπτικών-τυπογραφικών στοιχείων που αναπαριστώνται ακουστικά με τη χρήση συνθετικής ομιλίας.
2. Προτεινόμενη μεθοδολογία για την πολυτροπική πρόσβαση των εγγράφων.

3. Χρήση των συναισθημάτων-συναισθηματικής κατάστασης ως μέσο που βοηθάει στην ακουστική απεικόνιση της οπτικής πληροφορίας με τον συνθέτη ομιλίας.
4. Αξιολόγηση της παρούσας μεθοδολογίας μέσω ψυχοακουστικών πειραμάτων αντίληψης-κατανόησης των τυπογραφικών στοιχείων με ακουστικό τόπο και βελτιστοποίηση της εκπαιδευτικής διαδικασίας με τη χρήση της παρούσας μεθοδολογίας.

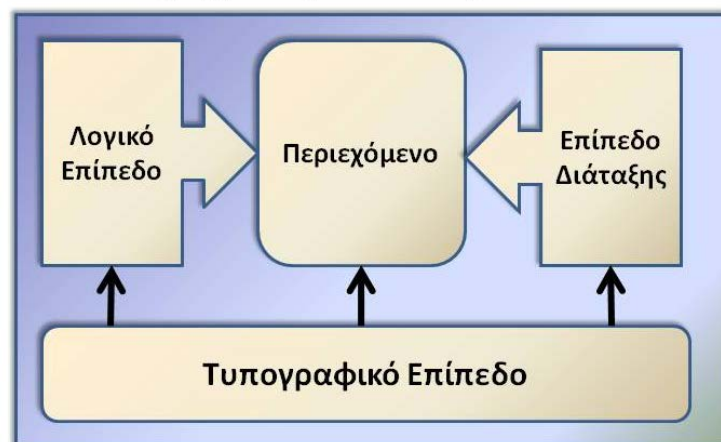
2. ΠΟΛΥΤΡΟΠΙΚΗ ΠΡΟΣΒΑΣΗ ΣΤΑ ΕΓΓΡΑΦΑ

2.1 Εισαγωγή

Όπως αναφέρθηκε και στη θεωρητική ανασκόπηση, η συζήτηση για την προσβασιμότητα των εγγράφων θέτει ερωτήματα κατά πόσο μπορεί ένα έγγραφο να είναι προσβάσιμο με οπτικό, ακουστικό ή απτικό τρόπο. Μελέτες όπως των [48] [156] [148], προσπαθούν να δημιουργήσουν προσβάσιμα έγγραφα με ακουστικό τρόπο χρησιμοποιώντας συνθετική ομιλία ή συνδυάζοντας τεχνητούς ήχους μικρής διάρκειας (earcons) [157] [158] [159], ακουστικά εικονίδια (auditory icons) [160] και τρισδιάστατους ήχους [161]. Μία ad-hoc προσέγγιση για πολυτροπική πρόσβαση, στα τεχνικά κυρίως έγγραφα σε μορφοποίηση TEX, παρουσιάστηκε πρόσφατα στην [162]. Εισάγουμε το εξής «αρχιτεκτονικό» μοντέλο του εγγράφου (Σχήμα 6):

1. *Λογικό Επίπεδο*: συνδέει το περιεχόμενο με στοιχεία όπως η επικεφαλίδα, τίτλος/υπότιτλος, κεφάλαιο, παράγραφος, πίνακες, λίστες, υποσημειώσεις και περιεχόμενα.
2. *Επίπεδο Διάταξης*: συνδέει το περιεχόμενο με στοιχεία που σχετίζονται με τη διάταξη στη σελίδα και σε περιοχές μέσα στις σελίδες, όπως περιθώρια, στήλες και ευθυγράμμιση (alignment).
3. *Τυπογραφικό Επίπεδο*: περιλαμβάνει στοιχεία γραμματοσειράς (font), όπως το είδος, το μέγεθος, το χρώμα, το χρώμα υποβάθρου και τον τρόπο γραφής (font style) όπως έντονη, πλάγια, υπογραμμισμένη.

Έγγραφο Κειμένου



Σχήμα 6: Γενικό μοντέλο «αρχιτεκτονικής» κειμενικού εγγράφου.

Τα επίπεδα είναι συμπληρωματικά - αλληλένδετα και όχι ανεξάρτητα. Η τυπογραφία μπορεί να εφαρμοστεί και στο λογικό επίπεδο και στο επίπεδο διάταξης του εγγράφου καθώς και στο περιεχόμενο. Για παράδειγμα, η υποσημείωση (λογικό επίπεδο) μπορεί να είναι σε πλάγια γραφή ή με μικρότερο μέγεθος γραμματοσειράς από ότι το σώμα κειμένου. Η απόσταση μεταξύ διαδοχικών γραμμών ενός κειμένου, που ονομάζεται *leading* (επίπεδο διάταξης) μπορεί να επηρεαστεί από το είδος της γραμματοσειράς. Η τυπογραφία μπορεί να εφαρμοστεί απευθείας στο σώμα του κειμένου, για παράδειγμα μια λέξη σε έντονη γραφή χρησιμοποιείται για να υποδηλώσει έμφαση ή την εισαγωγή ενός καινούριου όρου.

Όλα τα επίπεδα, που μπορεί να αναφέρονται και ως «σήματα», αποτελούν τους μηχανισμούς που [23]:

- Στοχεύουν στην καθοδήγηση της προσοχής του αναγνώστη κατά τη διαδικασία ανάγνωσης.
- Εξυπηρετούν συγκεκριμένες γνωσιακές διαδικασίες που πραγματοποιούνται κατά την ανάγνωση.
- Καθιστούν απόλυτα κατανοητή την πληροφορία του κειμένου.
- Μπορεί να επηρεάζουν τη μνήμη – αποστήθιση του κειμένου.
- Καθοδηγούν την επιλεκτική πρόσβαση μεταξύ και μέσα στα κείμενα.
- Επηρεάζουν τα συναισθήματα – συναισθηματική κατάσταση του αναγνώστη.

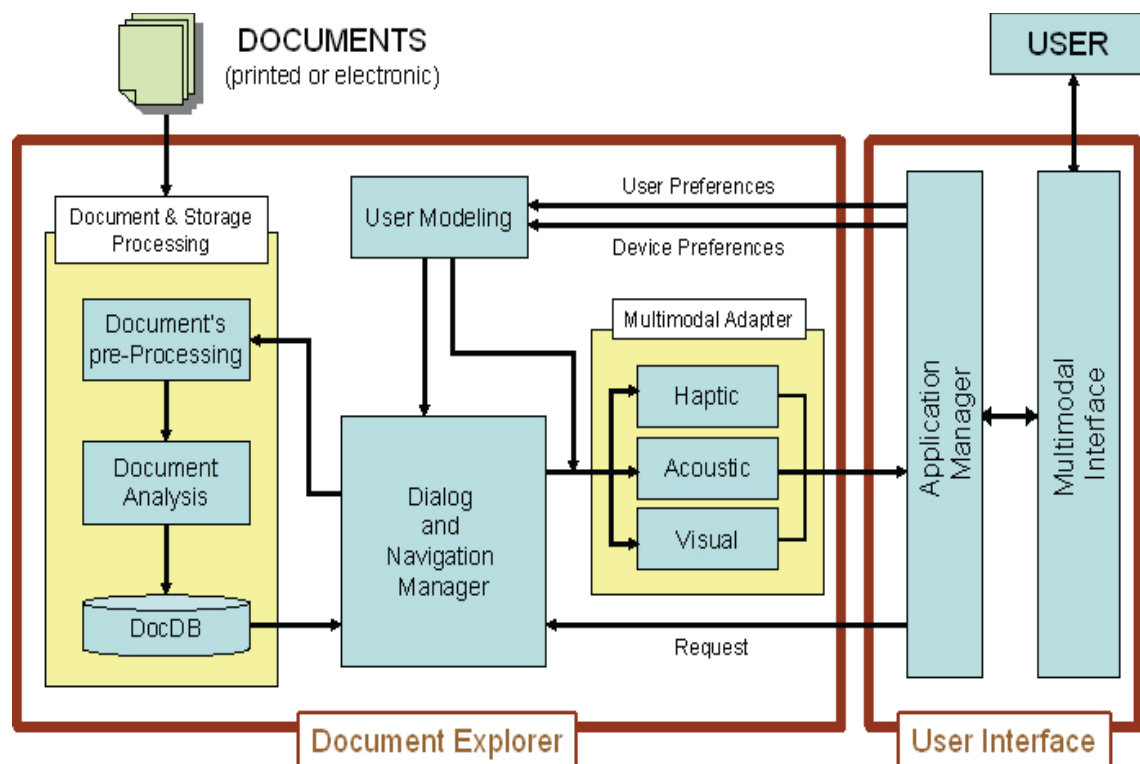
Ο πρωταρχικός ρόλος της χρήσης των στοιχείων παρουσίασης στα έγγραφα και πιο συγκεκριμένα των τυπογραφικών, είναι η διάκριση συγκεκριμένων κομματιών του κειμένου και η δημιουργία ενός αισθητικά όμορφου αποτελέσματος παρουσίασης του περιεχομένου, με σκοπό π.χ. την επαυξημένη απόδοση κατά την ανάγνωση και την προσέλκυση του αναγνώστη. Οι συγγραφείς χρησιμοποιούν την τυπογραφία με συγκεκριμένο τρόπο, όπως π.χ. τα επιστημονικά έγγραφα χρησιμοποιούν αυστηρούς τυπογραφικούς κανόνες. Αντίθετα, στις εφημερίδες και τα περιοδικά, ο πρωταρχικός λόγος χρήσης και εφαρμογής της τυπογραφίας ανήκει στο τμήμα σχεδιασμού της εφημερίδας και όχι στους συγγραφείς των άρθρων.

Σε αυτό το κεφάλαιο παρουσιάζουμε ένα πρωτότυπο σύστημα βασισμένο στην τεχνολογία XML για την πολυτροπική παραγωγή, παρουσίαση και πλοήγηση σε πραγματικό χρόνο προσβάσιμων εγγράφων, ακολουθώντας τις οδηγίες και τα πρότυπα που παρουσιάστηκαν στην παράγραφο 1.4. Αυτή η προσέγγιση περιλαμβάνει την

πρόταση για μία ενιαία μεθοδολογία για την πολυτροπική απόδοση της μορφοποίησης, δομής, διάταξης του κειμένου και των μη-κειμενικών στοιχείων του.

2.2 Η προτεινόμενη αρχιτεκτονική για πολυτροπική πρόσβαση στα έγγραφα

Η προτεινόμενη αρχιτεκτονική αποτελείται από δύο βασικά μέρη (Σχήμα 7): α) Τη Διεπαφή Χρήστη – ΔΧ (User Interface) και β) τον Εξερευνητή Εγγράφου – ΕΕ (Document Explorer).



Σχήμα 7: Το προτεινόμενο σύστημα για την πολυτροπική πρόσβαση στα έγγραφα.

Ο δεύτερος είναι υπεύθυνος για την ανάλυση των εγγράφων κατά τη διάρκεια της παραγωγής και εξερεύνησής τους (κατά τη διαδικασία πλοήγησης). Η ΔΧ είναι υπεύθυνη για την πολυτροπική αλληλεπίδραση με τα έγγραφα. Συγκεντρώνει τις προτιμήσεις του χρήστη και τα αιτήματα της συσκευής που χρησιμοποιεί, όπως επίσης τις εντολές πλοήγησης και εκτελεί τη διαδικασία παρουσίασης. Αυτά τα μέρη του συστήματος υλοποιούνται ακολουθώντας τη λογική του μοντέλου Πελάτη – Εξυπηρετητή (Client – Server). Ο ΕΕ μπορεί να φιλοξενηθεί σε ένα υπολογιστικό σύστημα που έχει το ρόλο του εξυπηρετητή, διότι έχει πολλές απαιτήσεις σε πόρους λόγω των εργασιών που εκτελεί. Αντίθετα, η διεπαφή χρήστη μπορεί να φιλοξενηθεί σε ένα κοινό υπολογιστή (για παράδειγμα, ένας προσωπικός υπολογιστής, ένα PDA,

κινητές έξυπνες συσκευές - smartphones). Αυτός ο τρόπος υλοποίησης ακολουθεί τις οδηγίες Web Content Accessibility Guidelines, με σκοπό να είναι ανεξάρτητος συσκευής και λογισμικού. Η υλοποίηση και η επικοινωνία των αρθρωμάτων βασίζεται στην τεχνολογία XML.

Οι υποστηριζόμενες διαδικασίες είναι:

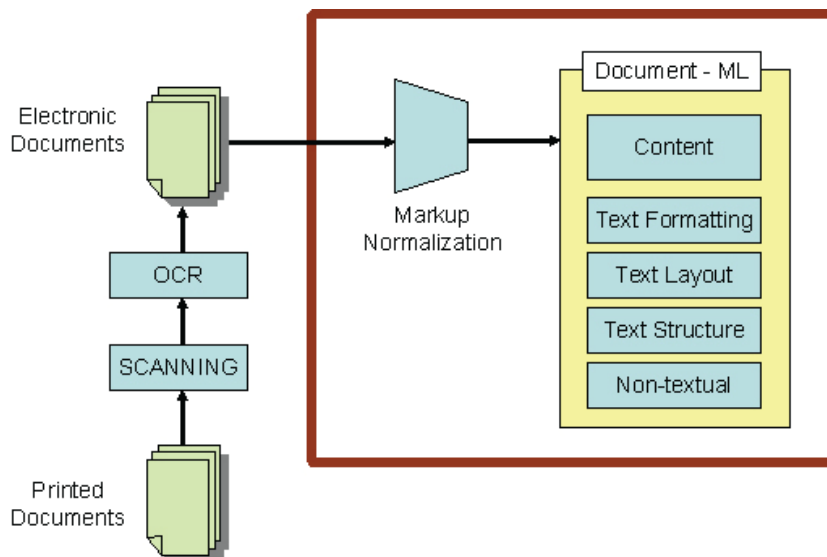
- Παραγωγή
- Πλοήγηση
- Παρουσίαση

2.3 Παραγωγή πολυτροπικά προσβάσιμων εγγράφων

Το άρθρωμα Επεξεργασίας και Αποθήκευσης Εγγράφων είναι υπεύθυνο για την παραγωγή προσβάσιμων εγγράφων. Επεξεργάζεται ένα έγγραφο και μετά δημιουργεί την έξοδο σύμφωνα με το πρότυπο DAISY/NISO. Η έξοδος αποθηκεύεται σε μία βάση δεδομένων (DocDB), ώστε το άρθρωμα Διαχείρισης Διαλόγου και Πλοήγησης (Dialog and Navigation Manager module) να έχει γρήγορη πρόσβαση στο περιεχόμενο.

2.4 Προ-επεξεργασία εγγράφων

Το Σχήμα 8 παρουσιάζει την μονάδα προ-επεξεργασίας.



Σχήμα 8: Το άρθρωμα Προ-Επεξεργασίας των Εγγράφων.

Μπορεί να χειρίζεται είτε έντυπα είτε ηλεκτρονικά έγγραφα. Τα έντυπα σαρώνονται και επεξεργάζονται από μία εφαρμογή λογισμικού Οπτικής Αναγνώρισης Χαρακτήρων. Το άρθρωμα Κανονικοποίησης Επισημείωσης (Markup Normalization module) επεξεργάζεται τα ψηφιοποιημένα ή τα ηλεκτρονικά έγγραφα με σκοπό η μορφή του αρχείου να ακολουθεί τις προδιαγραφές. Το αρχείο εξόδου (δηλαδή το Document-ML) περιλαμβάνει: το περιεχόμενο του εγγράφου, τη μορφοποίηση του κειμένου, τη δομή του κειμένου, και τα μη-κειμενικά στοιχεία (π.χ. εικόνες, ήχους κ.α.).

2.5 Ανάλυση των εγγράφων

Για να εξυπηρετηθούν οι απαιτήσεις για πολυτροπική προσβασιμότητα, η αρχιτεκτονική επισημειώνει το έγγραφο στο:

- Σημασιολογικό επίπεδο.
- Συναισθηματικό επίπεδο.

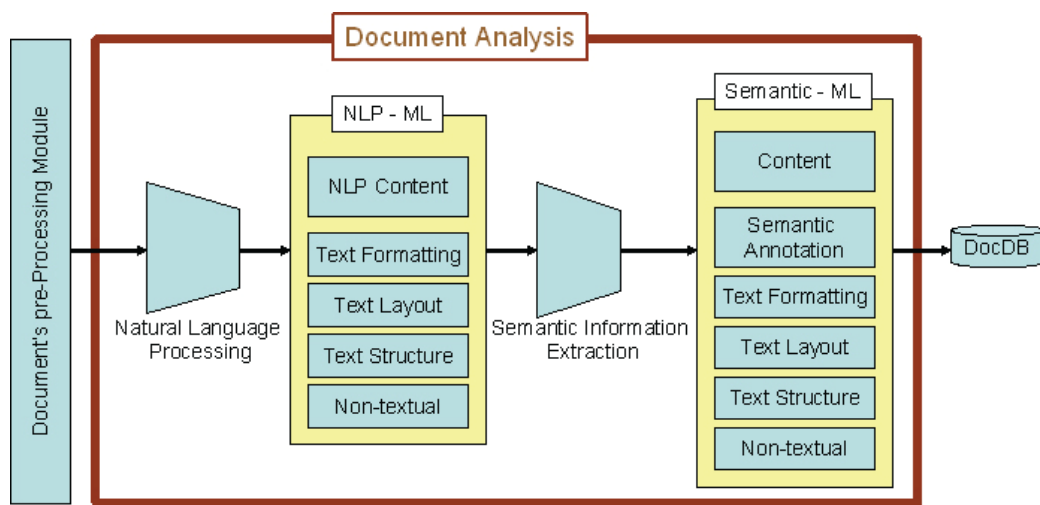
Και τα δύο επίπεδα δεν περιέχουν πληροφορίες που σχετίζονται με την παρουσίαση του εγγράφου. Το σημασιολογικό επίπεδο στοχεύει στην καταγραφή της κατανόησης του εγγράφου από τον αναγνώστη. Το συναισθηματικό επίπεδο εξάγει την επαγόμενη συναισθηματική κατάσταση του αναγνώστη από τα οπτικά χαρακτηριστικά του εγγράφου. Τα δεδομένα εξόδου αυτών των δύο επιπέδων μπορούν να μεταφερθούν σε κάθε αισθητηριακό τρόπο. Στις παρακάτω παραγράφους παρουσιάζεται ο τρόπος με τον οποίο το αρχικό οπτικό ερέθισμα επηρεάζει την πολυτροπική παρουσίαση, μέσω αυτών των επιπέδων, με έμφαση τον ακουστικό τρόπο.

2.5.1 Σημασιολογικό Επίπεδο

Υπάρχουν διάφορες εργασίες για την αυτόματη αναγνώριση της λογικής δομής των εγγράφων π.χ. [163] [164] [165] [166] [167] [168]. Οι περισσότερες παραδοσιακές προσεγγίσεις χρησιμοποιούν ντετερμινιστικές μεθόδους (π.χ. δένδρα αποφάσεων) [169] [170] [165], που μπορεί να υστερούν λόγω θορύβου και αβεβαιότητας που μειώνει την απόδοση τους. Επιπρόσθετα, τέτοιες προσεγγίσεις δημιουργούν μοντέλα που δεν είναι ευέλικτα σε αλλαγές του θεματικού τομέα και δεν μπορούν να εξελιχτούν με την παρουσία νέων στοιχείων. Με σκοπό την άρση αυτών των περιορισμών, οι Fourli-Kartsouni et al. [33] χρησιμοποιούν μία πιθανολογική (probabilistic) προσέγγιση βασισμένη σε Bayesian Δίκτυα, εκπαιδευμένα από μια σειρά επισημειωμένων εγγράφων. Τα δίκτυα προσφέρουν μια σημαντική ανοχή στο θόρυβο και την αβεβαιότητα και μπορούν να εκπαιδεύονται από παραδείγματα, με σκοπό την

προσαρμογή τους σε παρόντα και μελλοντικά στοιχεία. Μπορούν να εκπαιδευτούν για την αντιστοίχιση μεταξύ μορφής- δομής του κειμένου και λογικών στοιχείων. Οι κανόνες αντιστοίχισης (που σε μερικές περιπτώσεις είναι 1:N) μπορούν να προκύψουν από μια σειρά πειραμάτων (π.χ. η έντονη γραφή μπορεί να αντιστοιχηθεί ως έμφαση, αλλά και ως έντονη έμφαση).

Το άρθρωμα της Ανάλυσης Εγγράφου (Document Analysis module) (Σχήμα 9) προσθέτει σημασιολογική επισήμειωση - ετικέτες στο έγγραφο χρησιμοποιώντας τη μεθοδολογία που προτείνεται στην [33]. Παράγει ένα XML αρχείο (Semantic-ML) που περιλαμβάνει τα στοιχεία του εγγράφου που περιγράφηκαν στην αρχιτεκτονική του εγγράφου και αποθηκεύει την έξοδο στη βάση δεδομένων (DocDB).



Σχήμα 9: Το άρθρωμα της Ανάλυσης των Εγγράφων: η προσέγγιση της εξαγωγής Σημασιολογικής Πληροφορίας.

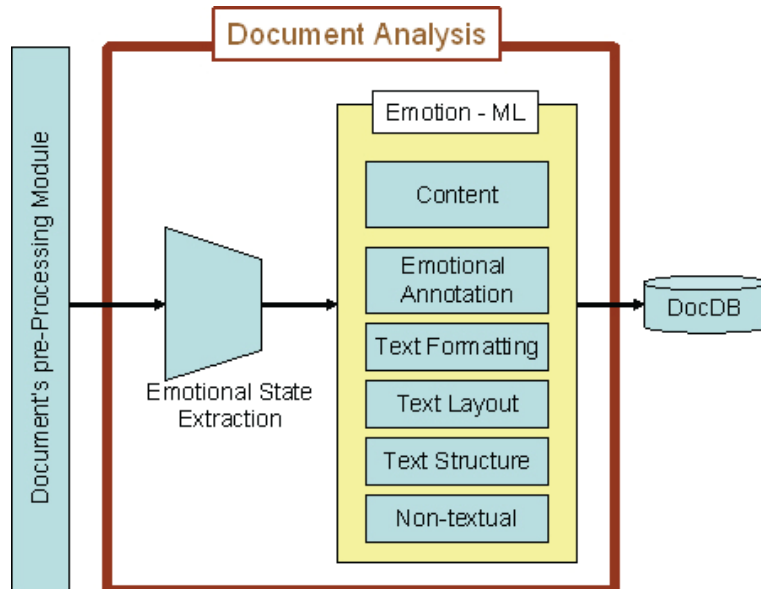
2.5.2 Συναισθηματικό Επίπεδο

Χρησιμοποιώντας την αυτόματη εξαγωγή της συναισθηματικής κατάστασης των αναγνωστών [171] [94], το άρθρωμα της Ανάλυσης Εγγράφου υλοποιεί την αντιστοίχιση των στοιχείων του εγγράφου σε μεταβολές των διαστάσεων συναισθηματικής κατάστασης «Ευχαρίστηση», «Διέγερση» και «Επικράτηση» ή συγκεκριμένων συναισθημάτων (Σχήμα 10). Το άρθρωμα παράγει ένα αρχείο τύπου XML (Emotion-ML) με επισήμειωση της συναισθηματικής κατάστασης.

2.6 Πολυτροπική πλοήγηση στα προσβάσιμα έγγραφα

Το υποσύστημα Διαχείρισης Διαλόγου και Πλοήγησης (Dialog and Navigation Manager) χειρίζεται τις διεργασίες πλοήγησης και προωθεί το περιεχόμενο στον Πολυτροπικό Προσαρμογέα (Multimodal Adapter). Ο διαχειριστής υποστηρίζει όλες τις λειτουργίες

πλοήγησης, όπως προτείνονται από το πρότυπο DAISY/NISO, χρησιμοποιώντας το αρχείο ελέγχου πλοήγησης (Navigation Control File) - αυτό το αρχείο περιγράφει την ιεραρχία του εγγράφου.



Σχήμα 10: Το άρθρωμα της Ανάλυσης των Εγγράφων: η προσέγγιση της εξαγωγής Συναισθηματικής Πληροφορίας

Για την υλοποίηση του διαχειριστή, με σκοπό τη βέλτιστη απόδοση, προτείνεται η χρήση ενός προσαρμόσιμου, διαλογικού συστήματος έξυπνου πράκτορα (agent-based) [172] [173] ο οποίος συμβάλει :

- στη Χρήση του Συστήματος, αναλαμβάνοντας μέρη διεργασιών ρουτίνας, προσαρμόζοντας τη διεπαφή, δίνοντας συμβουλή για τη χρήση του συστήματος και να ελέγχει το διάλογο
- στην Ανάκτηση Πληροφορίας, βοηθώντας τους χρήστες να βρουν πληροφορίες και προσαρμόζοντας την παρουσίαση της πληροφορίας.

Ένα τέτοιο διαλογικό σύστημα επιτρέπει σε κάθε άρθρωμα να αλληλεπιδρά και είναι ικανό για συλλογιστική-συμπερασμό. Είναι ευέλικτο επιτρέποντας στο χρήστη να έχει πλήρη έλεγχο του διαλόγου [174] και υποστηρίζει πολυτροπική είσοδο στο διάλογο [175].

Οι προτιμήσεις του χρήστη και τα αιτήματα της συσκευής επεξεργάζονται από το άρθρωμα Μοντελοποίησης του Χρήστη (User Modeling module) και διαχειρίζονται από το υποσύστημα του Διαχειριστή Διαλόγου και Πλοήγησης. Το άρθρωμα της μοντελοποίησης του χρήστη δημιουργεί το προφίλ του χρήστη σύμφωνα με τα αιτήματα της αλληλεπίδρασης και της συσκευής. Για παράδειγμα, ο χρήστης μπορεί να επιλέξει τον επιθυμητό τρόπο παρουσίασης του περιεχομένου σύμφωνα με τις ανάγκες ή τον τρόπο που δίνεται μια εντολή, μαζί με τις απαιτήσεις της συσκευής του χρήστη. Το προφίλ του χρήστη είναι μια συλλογή από ενέργειες και προτιμήσεις που συλλέγονται κατά την αλληλεπίδραση με το Διαχειριστή Εφαρμογής (Application Manager). Αυτό μπορεί να περιλαμβάνει τις ανάγκες του χρήστη, τον τρόπο αλληλεπίδρασης με τη διεπαφή, τους παράγοντες πλοήγησης, περιβαλλοντικούς παράγοντες κ.α. Το αρχείο XML που παράγεται (User-ML) προωθείται στο Διαχειριστή Διαλόγου και Πλοήγησης. Έτσι, σύμφωνα με τις προτιμήσεις του χρήστη, ο διαχειριστής διαλόγου χειρίζεται τον τρόπο με τον οποίο το περιεχόμενο χρησιμοποιείται και παρουσιάζεται (ακουστικό, οπτικό, απτικό τρόπο) μέσω του Ακουστικού ή/και του Οπτικού ή/και του Απτικού Προσαρμογέα (Haptic Adapter).

2.7 Πολυτροπική παρουσίαση των προσβάσιμων εγγράφων

Η διεπαφή χρήστη αποτελείται από δύο μέρη: Την Πολυτροπική Διεπαφή (Multimodal Interface) και το Διαχειριστή Εφαρμογής (Application Manager) (Σχήμα 7). Η Πολυτροπική Διεπαφή είναι υπεύθυνη για την αλληλεπίδραση του χρήστη με το έγγραφο υποστηρίζοντας και τους τρεις τρόπους αλληλεπίδρασης: οπτικό, ακουστικό και απτικό. Υποστηρίζει:

- Διεργασίες εισόδου: (εντολές λειτουργίας ή πλοήγησης): επιτυγχάνονται με τη χρήση συσκευών (π.χ. πληκτρολόγιο, ποντίκι, κουμπιά) ή εφαρμογών εισόδου (π.χ. αυτόματη αναγνώριση φωνής)
- Διεργασίες Εξόδου: ο χρήστης λαμβάνει το ζητούμενο περιεχόμενο ή εντολές-οδηγίες του συστήματος (system prompts) με σειριακό τρόπο και ιεραρχική πλοήγηση.

Η λειτουργικότητα των παραπάνω διεργασιών μπορεί να αλλάξει σύμφωνα με τις προτιμήσεις του χρήστη.

Ο Διαχειριστής Εφαρμογής συλλέγει και χειρίζεται τα ερωτήματα της συσκευής και τις προτιμήσεις του χρήστη, που παρέχονται από την πολυτροπική διεπαφή. Η έξοδος του διαχειριστή εφαρμογής χρησιμοποιείται από: α) το άρθρωμα Μοντελοποίησης του

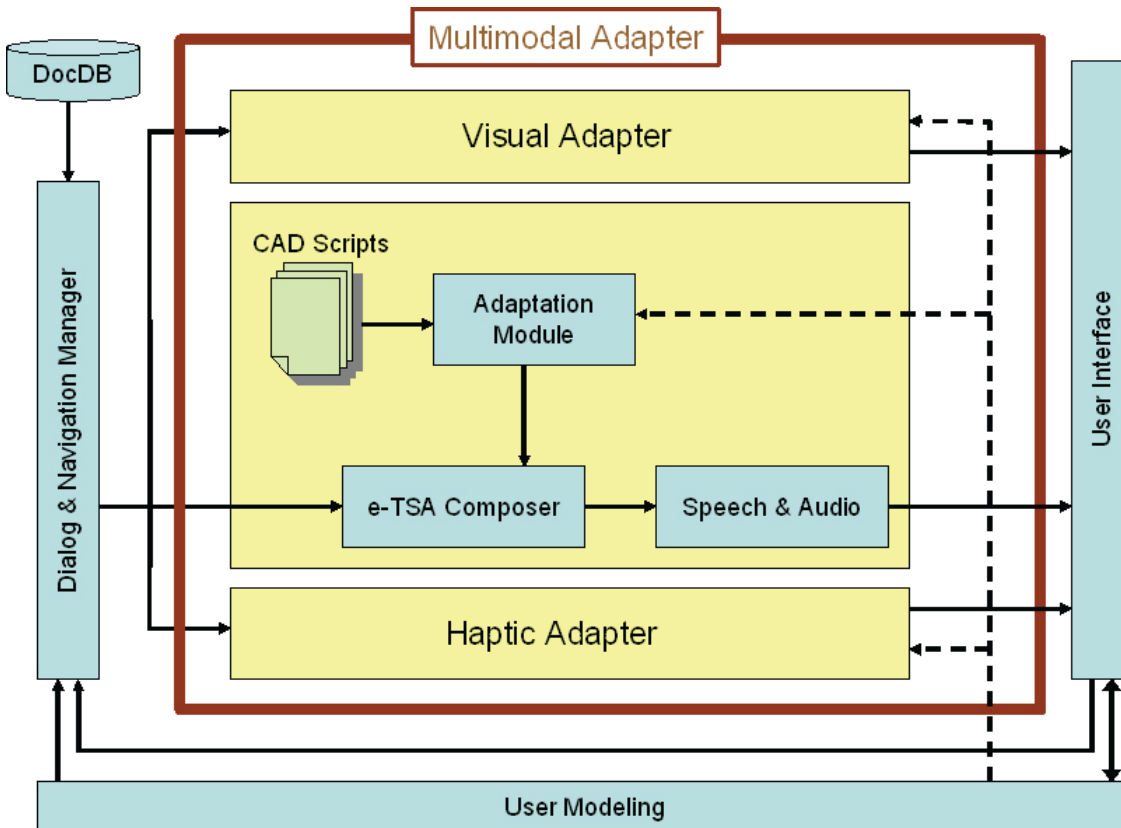
Χρήστη και β) το Διαχειριστή Διαλόγου και Πλοήγησης (εντολές πλοήγησης και εφαρμογής).

Ο Διαχειριστής Διαλόγου και Πλοήγησης τροφοδοτεί τον πολυτροπικό προσαρμογέα με το έγγραφο που πρέπει να παρουσιαστεί υπό τη μορφή Semantic-ML ή Emotion-ML. Ο αντίστοιχος προσαρμογέας ενεργοποιείται και παράγει δεδομένα εξόδου σύμφωνα με τις προτιμήσεις του χρήστη και τις ερωτήσεις της συσκευής. Η διεπαφή του χρήστη διαχειρίζεται την έξοδο με σκοπό την παρουσίαση του περιεχομένου με οπτικό, ακουστικό ή απτικό τρόπο.

2.8 Ακουστική παρουσίαση των πολυτροπικά προσβάσιμων εγγράφων

Η αντιστοίχιση με ακουστικό τρόπο επιτυγχάνεται χρησιμοποιώντας την πλατφόρμα μετατροπής Εγγράφου σε Ακουστική μορφή (Document-to-Audio - DtA) [148]. Στον Ακουστικό Προσαρμογέα, το άρθρωμα προσαρμογής (Σχήμα 11) συνδυάζει πληροφορίες για τις προτιμήσεις του χρήστη (User Modeling) και τους κανόνες για την ακουστική αντιστοίχιση (CAD script), ώστε το αποτέλεσμα να μπορεί να χρησιμοποιηθεί από τον συνθέτη e-TSA. Τα σενάρια Cluster Auditory Definition παρέχουν τους κανόνες αντιστοίχισης για την πλατφόρμα DtA. Οι κανόνες για τη συσχέτιση των στοιχείων του εγγράφου και την ακουστική αναπαράσταση, εξαρτώνται από τη μεθοδολογία που ακολουθείται, χρησιμοποιώντας τη σημασιολογία [35] [37] [38] ή τη συναισθηματική προσέγγιση μέσω εκφραστικής συνθετικής ομιλίας [132] [135] [142].

Η πλατφόρμα DtA δίνει μεγαλύτερη προτεραιότητα στις προτιμήσεις του χρήστη από ότι τους προεπιλεγμένους κανόνες CAD. Για παράδειγμα, ο χρήστης μπορεί να χρειάζεται να ακούσει γρήγορα το περιεχόμενο του βιβλίου, αλλά μερικά στοιχεία πρέπει να διαβάζονται πιο αργά. Το άρθρωμα προσαρμογής δίνει μεγαλύτερη προτεραιότητα στους κανόνες που παρέχονται από τη μοντελοποίηση του χρήστη. Χρησιμοποιώντας την πλατφόρμα DtA, τα έγγραφα αντιστοιχίζονται σε συγκεκριμένα ακουστικά χαρακτηριστικά που υλοποιούνται με τον ακουστικό συνθέτη (auditory synthesizer) (η μορφή της εξόδου μπορεί να είναι π.χ. MPEG4, SMIL ή WAV). Όπως φαίνεται στο Σχήμα 14 της παραγράφου 2.9, το άρθρωμα του προσαρμογέα μπορεί να δεχτεί και άλλους κανόνες, για παράδειγμα την ακουστικοποίηση των μαθηματικών εκφράσεων [51].



Σχήμα 11: Το διάγραμμα της Πολυτροπικής Απεικόνισης των εγγράφων, με έμφαση στον ακουστικό τρόπο

Ο όρος «εναλλαγή» (alteration) στην παρούσα εργασία χρησιμοποιείται για τις αλλαγές στην τυπογραφία και ορίζεται ως:

«ένα γεγονός που συμβαίνει όταν κάτι περνάει από μία κατάσταση ή φάση σε μια άλλη ή η πράξη του να γίνεται κάτι διαφορετικό χωρίς να αλλάζει σε κάτι άλλο.»

Και μεταβολή (variation) χρησιμοποιείται για τα προσωδιακά χαρακτηριστικά ως:

«ένα περιστατικό - γεγονός αλλαγής, το ποσοστό ή το μέγεθος της αλλαγής, κάτι λίγο διαφορετικό από άλλα του ίδιου τύπου.»

Στη μελέτη μας χρησιμοποιούμε διαφορετικούς όρους για τις τυπογραφικές και προσωδιακές αλλαγές, με σκοπό να διαφοροποιούνται οι έννοιες. Ο τρόπος γραφής (διακριτές τιμές) είναι απλό κείμενο, έντονη, πλάγια γραφή κτλ. και το μέγεθος γραμματοσειράς μετρίεται σε εικονοστοιχεία, δηλαδή σε ακέραιες τιμές. Οι τιμές αυτές «εναλλάσσονται» μέσα στο έγγραφο. Αντίθετα, τα προσωδιακά χαρακτηριστικά όπως τονικό ύψος, ρυθμός και ένταση είναι συνεχείς τιμές και μεταβάλλονται από ένα βασικό επίπεδο-τιμή κατά την ομιλία.

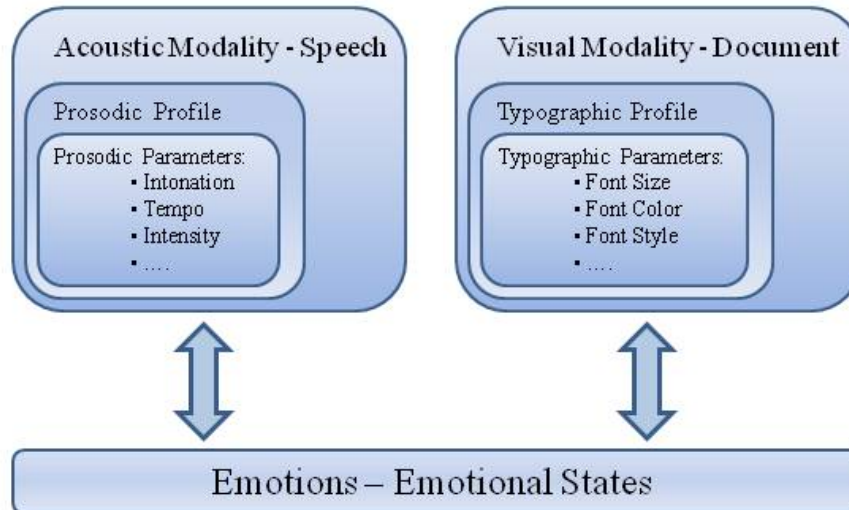
Καθώς διαβάζουμε ένα έγγραφο, τα τυπογραφικά στοιχεία εναλλάσσονται με το πέρασμα του χρόνου. Κατ' επέκταση τα επαγόμενα συναισθήματα και συναισθηματικές καταστάσεις του αναγνώστη εναλλάσσονται και αυτά. Ο συλλογισμός αυτός είναι παρόμοιος με αυτόν που χρησιμοποιείται για την εκφραστική συνθετική ομιλία από τον Schräder [142] με βάση τη διαστατική προσέγγιση των συναισθημάτων. Είναι άξιο αναφοράς ότι οι συνήθεις εναλλαγές των τυπογραφικών στοιχείων δεν είναι ακραίες, για παράδειγμα, μέγεθος γραμματοσειράς από 12pt σε 14pt ή από απλό κείμενο σε πλάγια γραφή. Για αυτό η διαστατική προσέγγιση των συναισθημάτων - είτε στην τυπογραφία είτε στην ομιλία - θεωρείται ως η κατάλληλη προσέγγιση για την αντιστοίχιση των τυπογραφικών εναλλαγών σε προσωδιακές μεταβολές. Η διαστατική προσέγγιση των συναισθημάτων στην ομιλία βασίζεται στην απεικόνιση και μεταβολή των συναισθημάτων με συνεχείς τιμές, για αυτό το λόγο μπορούμε να χρησιμοποιήσουμε το συγκεκριμένο μοντέλο ώστε να απεικονιστεί η ανάλογη συμπεριφορά των τυπογραφικών στοιχείων των εγγράφων στον ακουστικό τρόπο.

Όπως αναφέρθηκε στο 1ο Κεφάλαιο, οι συγγραφείς χρησιμοποιούν την τυπογραφία με συγκεκριμένο τρόπο ανάλογα με τον τύπο του εγγράφου. Για παράδειγμα, στα επιστημονικά έγγραφα χρησιμοποιούνται αυστηροί τυπογραφικοί κανόνες, σε αντίθεση με τα περιοδικά και τις εφημερίδες που τον πρωταρχικό λόγο έχει το τμήμα σχεδιασμού της εφημερίδας και όχι οι συγγραφείς των άρθρων. Σε αυτό το σημείο εισάγουμε τον όρο «*τυπογραφική όψη*» (typographic profile) ενός συναισθήματος με τρόπο παρόμοιο με αυτόν που παρουσιάζεται η έννοια «*προσωδιακή όψη*» (prosodic profile) στην [130]. Ο τρόπος με τον οποίο χρησιμοποιούνται τα τυπογραφικά στοιχεία, συνθέτει την τυπογραφική όψη του εγγράφου η οποία καθορίζει το χώρο - διάστημα που τοποθετείται κάθε συναίσθημα και αντίστροφα. Στο Σχήμα 12 παρουσιάζονται οι ομοιότητες των προσωδιακών και τυπογραφικών όψεων και η διασύνδεσή τους.

Οι προσεγγίσεις τυπογραφία-σε-συναισθήματα και συναισθήματα-σε-ομιλία, μπορούν να αντιστοιχήσουν τις τυπογραφικές εναλλαγές σε προσωδιακές μεταβολές και είμαστε σε θέση να παρατηρήσουμε τις παραμέτρους της ομιλίας που μεταβάλλονται. Ο συνδυασμός αυτών των μεθοδολογιών παρέχει το μοντέλο του τρόπου μεταβολής των προσωδιακών χαρακτηριστικών σύμφωνα με τις εναλλαγές των τυπογραφικών στοιχείων. Έτσι, τα ανθρώπινα συναισθήματα αποτελούν το μέσο για την αλληλεπίδραση-επικοινωνία των δύο καναλιών / αισθητηριακών τρόπων.

Υπάρχει μία τάση για μαθηματική περιγραφή της τυπογραφικής επίδρασης στα συναισθήματα και των συναισθημάτων στην ομιλία. Για αυτό, μπορούμε να προτείνουμε

ένα γενικό μαθηματικό χειρισμό των υπάρχοντων ή μελλοντικών μοντέλων υπό την προτεινόμενη μεθοδολογία.



Σχήμα 12: Η προσωδιακή και τυπογραφική όψη.

Η τυπογραφία επηρεάζει τα συναισθήματα/συναισθηματική κατάσταση σύμφωνα με την εξίσωση 3:

$$\Delta E = A * \Delta T + B \quad (3)$$

όπου:

A: ο πίνακας των συντελεστών της εξίσωσης ($n \times m$)

B: ο πίνακας των σταθερών ($n \times 1$)

ΔT : ο πίνακας που περιγράφει τις τυπογραφικές εναλλαγές ($m \times 1$)

ΔE : ο πίνακας που περιγράφει τις επαγόμενες συναισθηματικές μεταβολές ($n \times 1$)

$$\Delta E = E_{\text{final}} - E_{\text{baseline}} \quad (4)$$

Ο πίνακας ΔE που περιγράφει, είτε τα διακριτά συναισθήματα είτε τις μεταβολές των επαγόμενων συναισθηματικών καταστάσεων.

όπου:

$E_{baseline}$, αποτελεί τις επαγόμενες συναισθηματικές μεταβολές (από την ουδέτερη κατάσταση) για την αντίστοιχη τυπογραφική τιμή αναφοράς και

E_{final} , τις συναισθηματικές μεταβολές (από την ουδέτερη κατάσταση) για το αντίστοιχο τυπογραφικό στοιχείο.

Για παράδειγμα, έστω ότι έχουμε τη μαθηματική περιγραφή των έξι βασικών συναισθημάτων από τα τυπογραφικά στοιχεία. Ο πίνακας των συναισθηματικών μεταβολών είναι διαστάσεων 6×1 , των συντελεστών $6 \times m$ και των σταθερών 6×1 . Η περίπτωση των συναισθηματικών καταστάσεων περιγράφεται στην παράγραφο 3.3.5. Ο πίνακας ΔT περιγράφει τις τυπογραφικές εναλλαγές. Οι τιμές (συνεχείς) μπορούν να περιγραφούν, χρησιμοποιώντας είτε απλής είτε πολλαπλών μεταβλητών πολυωνυμικές εξισώσεις (για παράδειγμα, $P = a_n \cdot t^n + \dots + a_2 \cdot t^2 + a_1 \cdot t + b$ και $P = a_n \cdot t_n + \dots + a_2 \cdot t_2 + a_1 \cdot t_1 + b$ αντίστοιχα). Εάν τα συναισθήματα/συναισθηματικές καταστάσεις είναι διακριτές μεταβολές λόγω των τυπογραφικών εναλλαγών, τότε ο πίνακας των συντελεστών είναι $A = 0$ και οι συναισθηματικές μεταβολές είναι ίσες με τον πίνακα των σταθερών B .

Παρόμοια με τον προηγούμενο συλλογισμό, οι συναισθηματικές μεταβολές επηρεάζουν τα χαρακτηριστικά της ομιλίας (προσωδιακά), όπως παρουσιάζονται στην εξίσωση (5) :

$$\Delta S = A' * \Delta E + B' \quad (5)$$

όπου:

A' : ο πίνακας των συντελεστών της εξίσωσης ($m \times n$)

B' : ο πίνακας των σταθερών ($m \times 1$)

ΔE : ο πίνακας που περιγράφει τις συναισθηματικές μεταβολές ($n \times 1$)

$\Delta S = S_{final} - S_{baseline}$ ($m \times 1$)

Ο πίνακας ΔS περιγράφει τις προσωδιακές μεταβολές,

όπου:

$S_{baseline}$, η προσωδιακή βάση αναφοράς και

S_{final} , η προσωδιακή μεταβολή (από τη βάση αναφοράς) για την αντίστοιχη συναισθηματική μεταβολή.

Συνδυάζοντας τις εξισώσεις 3 και 5 προκύπτει:

$$(3), (5) \Rightarrow \Delta S = A' * (A * \Delta T + B) + B' = A' * A * \Delta T + A' * B + B' \quad (6)$$

$$\Delta S = C * \Delta T + D \quad (7)$$

όπου:

$C = A' * A$: ο πίνακας συντελεστών ($m \times m$)

$D = A' * B + B'$: ο πίνακας των σταθερών ($m \times 1$)

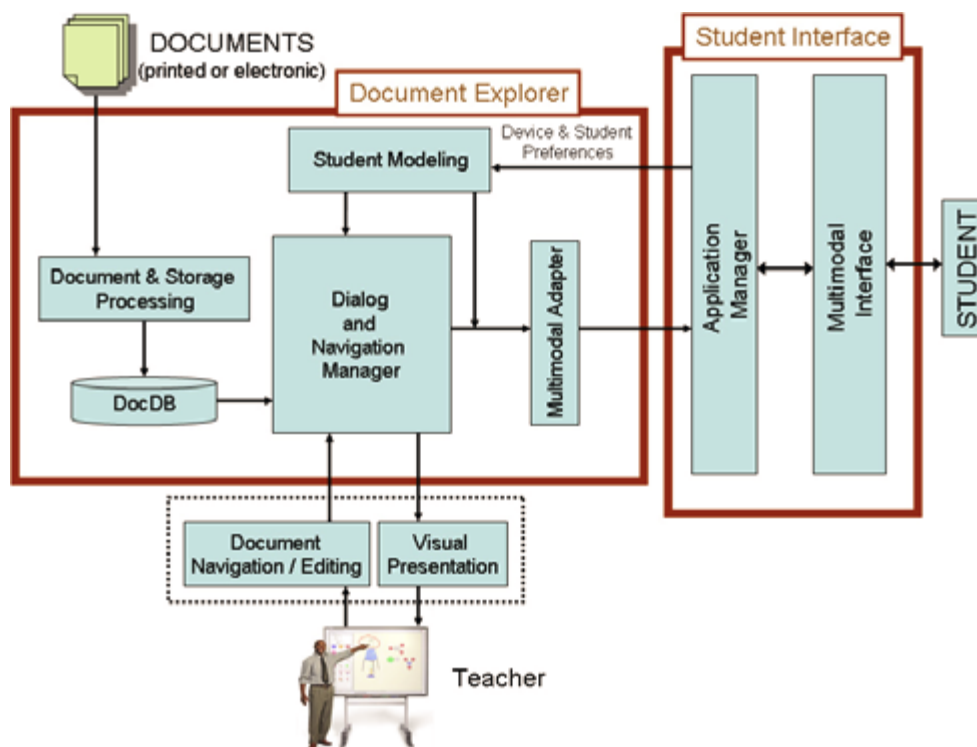
Η παραπάνω μαθηματική περιγραφή διατυπώνει το συνδυασμό των δύο μοντέλων (εξισώσεις 3 και 5) σε μία εξίσωση (7). Τα δύο μοντέλα περιγράφουν την συναισθηματική συσχέτιση των δύο διαφορετικών αισθητηριακών τρόπων – οπτική και ακουστική - και τα συναισθήματα/συναισθηματική κατάσταση αποτελούν το μέσο για τη διασύνδεση αυτών. Έτσι, οι προσωδιακές μεταβολές συσχετίστηκαν άμεσα με τις τυπογραφικές εναλλαγές.

2.9 Εφαρμογή της προτεινόμενης αρχιτεκτονικής κατά την διδακτική διαδικασία

Μία εφαρμογή της προτεινόμενης αρχιτεκτονικής, είναι η χρήση για την πρόσβαση των μαθητών στην παρουσίαση του διδάσκοντα κατά τη διδακτική διαδικασία και την ακουστικοποίηση των μαθηματικών εκφράσεων και εννοιών. Όπως φαίνεται και στο Σχήμα 13:

- Ο ΕΕ χειρίζεται: α) την επεξεργασία των εγγράφων και την προετοιμασία τους για την πολυτροπική παρουσίασή τους, β) διεργασίες πλοήγησης, γ) τη μοντελοποίηση του προφίλ του μαθητή και της απαιτήσεις της συσκευής που χρησιμοποιεί.
- Η Διεπαφή του μαθητή (Student Interface) υποστηρίζει: α) την οπτική, ακουστική ή/και την απτική παρουσίαση των εγγράφων στους μαθητές και β) τη συλλογή των προτιμήσεων του μαθητή και της συσκευής
- Οι διεργασίες πλοήγησης υποστηρίζονται μόνο από τη διεπαφή του καθηγητή για την οπτική παρουσίαση του εγγράφου.

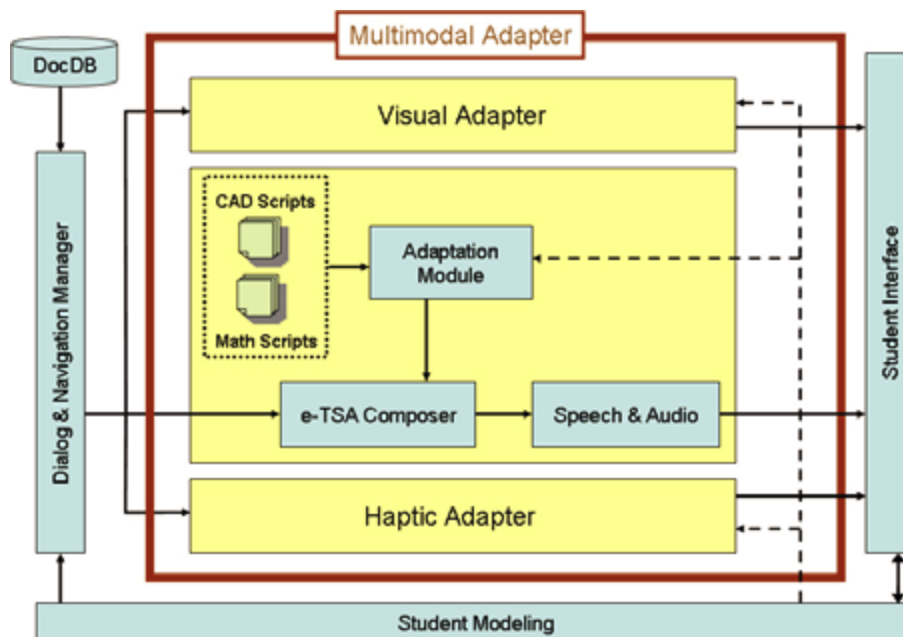
Η αλληλεπίδραση με τα έγγραφα διαχωρίζεται: α) στη διεργασία παρουσίασης και β) στη διεργασία πλοήγησης, που τις διαχειρίζεται ο Διαχειριστής Διαλόγου και Πλοήγησης. Ο διδάσκοντας χειρίζεται τις διεργασίες πλοήγησης και π.χ. ο υπολογιστής του μαθητή υποστηρίζει την πολυτροπική παρουσίαση μέσω κατάλληλων συσκευών – λογισμικών Υποστηρικτικής Τεχνολογίας. Σύμφωνα με τις προτιμήσεις του μαθητή και τις προδιαγραφές της συσκευής που χρησιμοποιεί, οι διεργασίες παρουσίασης επιτυγχάνονται με την επιλογή του κατάλληλου αρθρώματος για κάθε αισθητηριακό τρόπο (οπτικό, απτικό, ακουστικό ή ο συνδυασμός τους). Ο διαχειριστής τροφοδοτεί τα επιλεγμένα αρθρώματα με το περιεχόμενο για το επόμενο στάδιο επεξεργασίας σε κάθε αισθητηριακό τρόπο.



Σχήμα 13: Η τροποποιημένη αρχιτεκτονική του προτεινόμενου συστήματος για πολυτροπική προσβασιμότητα. Η λειτουργικότητα του έχει προσαρμοστεί για τη χρήση του στη διδακτική διαδικασία.

Στο Σχήμα 14 παρουσιάζεται ο πολυτροπικός προσαρμογέας. Ο προσαρμογέας αποτελείται από τρία μέρη, ένα για κάθε αισθητηριακό τρόπο (οπτικό, απτικό και ακουστικό). Ο μαθητής μπορεί να τροποποιήσει την οπτική παρουσίαση του περιεχομένου του εγγράφου, συμπεριλαμβανομένων και των μαθηματικών εκφράσεων.

Η MathML επιτρέπει την αλλαγή, π.χ. του μεγέθους γραμματοσειράς και του χρώματος της έκφρασης ακολουθώντας τις προτιμήσεις του μαθητή. Η απτική απόδοση των μαθηματικών εκφράσεων υποστηρίζεται από τη μετατροπή της Presentation ή Content MathML σε οποιοδήποτε μαθηματική κωδικοποίηση Braille (Mathematical Braille code), για παράδειγμα, με τη χρήση της Universal Math Conversion Library (UMCL) [176]. Η ακουστική απόδοση επιτυγχάνεται με τη χρήση της πλατφόρμας Εγγράφου-σε-Ακουστική μορφή. Το άρθρωμα προσαρμογής συνδυάζει της πληροφορίες των προτιμήσεων του μαθητή και τους κανόνες της ακουστικής αντιστοίχισης των οπτικών στοιχείων του εγγράφου (CAD script) και των μαθηματικών εκφράσεων (Math script). Τα σενάρια CAD και Math, περιλαμβάνουν τους κανόνες αντιστοίχισης των σημασιολογικών/λογικών μεταδεδομένων και των μαθηματικών εκφράσεων (χρησιμοποιώντας Presentation ή Content MathML) αντίστοιχα σε ακουστικά χαρακτηριστικά ώστε τα αποτελέσματα να μπορούν να χρησιμοποιηθούν από το σύστημα TtS.



Σχήμα 14: Ο πολυτροπικός προσαρμογέας τροποποιημένος έτσι ώστε να εξυπηρετεί και την ακουστική απόδοση των μαθηματικών εκφράσεων.

3. ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΕΠΑΓΩΜΕΝΗΣ ΣΥΝΑΙΣΘΗΜΑΤΙΚΗΣ ΚΑΤΑΣΤΑΣΗΣ ΤΟΥ ΑΝΑΓΝΩΣΤΗ

3.1 Εισαγωγή

Στο προηγούμενο κεφάλαιο προτείνεται η ακουστικοποίηση των τυπογραφικών εναλλαγών του εγγράφου μέσω της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη και την αντιστοίχισή της σε προσωδιακές μεταβολές. Ουσιαστικά, αυτή η μεθοδολογία χωρίζεται σε δύο μέρη, α) τη μοντελοποίηση της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη από τα τυπογραφικά στοιχεία και β) τη χρήση εκφραστικής συνθετικής ομιλίας για την ακουστική απόδοση των τυπογραφικών εναλλαγών.

Σε αυτό το κεφάλαιο αναλύεται το πρώτο μέρος της προτεινόμενης μεθοδολογία, δηλαδή η μαθηματική μοντελοποίηση της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη από τα τυπογραφικά στοιχεία και η μεθοδολογία που ακολουθήθηκε για αυτή τη μοντελοποίηση. Χρησιμοποιείται η διαστατική περιγραφή των συναισθημάτων (συναισθηματική κατάσταση) του αναγνώστη, με τη χρήση της δοκιμασίας SAM που παρουσιάστηκε εκτενώς στο εδάφιο 1.5.3.

Πραγματοποιήθηκε ένα πιλοτικό πείραμα σε μικρό αριθμό συμμετεχόντων (15 άτομα) και μετρήθηκε η αξιοπιστία των απαντήσεών τους σύμφωνα με την ανάλυση που παρουσιάζεται στην εργασία [177]. Με την ολοκλήρωση του πιλοτικού πειράματος, εξήχθησαν συμπεράσματα που βοήθησαν στη διόρθωση της πειραματικής διαδικασίας (π.χ. μειώθηκε ο αριθμός των ερεθισμάτων, γιατί παρατηρήθηκε κούραση των συμμετεχόντων με αποτέλεσμα την πιθανή αλλοίωση των αποτελεσμάτων). Προχωρήσαμε στο κύριο πείραμα όπου συμμετείχαν 30 άτομα. Γίνεται αναλυτική περιγραφή των συνθηκών διεξαγωγής και της ανάλυσης των αποτελεσμάτων (στατιστική και μοντελοποίηση).

Τέλος, παρουσιάζεται το σύστημα DocEmoX, το οποίο υλοποιεί το μαθηματικό μοντέλο που προέκυψε από την πειραματική διαδικασία. Παρουσιάζονται οι προδιαγραφές, οι απαιτήσεις, ο τρόπος υλοποίησης και η αξιολόγηση της ορθής λειτουργίας του.

3.2 Πιλοτικό πείραμα

Μέσα από το πιλοτικό πείραμα θέλουμε να έχουμε τις πρώτες ενδείξεις σε μικρό αριθμό ατόμων (15 συμμετέχοντες). Τα σημαντικότερα συμπεράσματα αναμένονται να εξαχθούν από: α) τον έλεγχο της ορθής συνεισφοράς των συμμετεχόντων στα αποτελέσματα, μετρώντας την αξιοπιστία των απαντήσεων τους και β) την καταγραφή

των μειονεκτημάτων του σχεδιασμού της διαδικασίας με σκοπό τη διόρθωσή τους στο κύριο μέρος.

Γενικά, αναμένεται να επιβεβαιώσουμε ή όχι ότι οι χρωματικοί συνδυασμοί γραμματοσειράς/υποβάθρου, το είδος της γραμματοσειράς και ο τρόπος γραφής θα επηρεάζουν τις συναισθηματικές διαστάσεις «Ευχαρίστηση», «Διέγερση» και «Επικράτηση». Επίσης, αναμένεται ότι το μέγεθος της γραμματοσειράς, διατηρώντας τα υπόλοιπα τυπογραφικά χαρακτηριστικά σταθερά, θα επηρεάζει και τις τρεις διαστάσεις. Αναλυτικά οι υποθέσεις της πειραματικής διαδικασίας παρουσιάζονται στην παράγραφο 3.3.1. όπου γίνεται λεπτομερής περιγραφή της.

3.2.1 Μεθοδολογία

Συμμετέχοντες

Συμμετείχαν 15 άτομα, ηλικίας 18 ως 33 ετών (μέση ηλικία 26,2 έτη), 8 άνδρες και 7 γυναίκες. Στο σύνολό τους ήταν προπτυχιακοί και μεταπτυχιακοί φοιτητές του τμήματος Πληροφορικής και Τηλεπικοινωνιών του Πανεπιστημίου Αθηνών. Κανένας δεν είχε κάποιο πρόβλημα στην όρασή του αναφορικά με τη διάκριση των χρωμάτων (π.χ. δυσχρωματοψία) και η οπτική οξύτητα τους ήταν φυσιολογική ή διορθωμένη.

Ερεθίσματα

Μελετήθηκαν πενήντα τέσσερις (54) τυπογραφικοί συνδυασμοί, σε συναισθηματικά ουδέτερο ελληνικό κείμενο (παρόμοιο με αυτό που χρησιμοποιήθηκε στην εργασία [13]). Αυτοί είναι:

- Απλό κείμενο σε μεγέθη γραμματοσειράς 9pt, 10pt, 11pt, 12pt, 13pt, 14pt, 15pt, 16pt, 18pt, 26pt, και 32pt και τύπο γραμματοσειράς «Times New Roman».
- Απλό κείμενο, έντονη, πλάγια, έντονη-πλάγια γραφή σε μέγεθος 16pt., και τύπο γραμματοσειράς «Arial» και «Times New Roman», σε χρωματικούς συνδυασμούς, όπως προτείνεται στις [16] και [13] (Πίνακας 1).

Μερικοί τυπογραφικοί συνδυασμοί αποκλείστηκαν, διότι έχουν μελετηθεί στο παρελθόν [16]. Κατά τη διάρκεια της πειραματικής διαδικασίας, επαναλάβαμε την προβολή μερικών ερεθισμάτων, με σκοπό το συσχετισμό της εξάρτησης του προηγούμενα

προβεβλημένου συνδυασμού τυπογραφικών στοιχείων, στην παρούσα συναισθηματική απόκριση του αναγνώστη. Όλα τα ερεθίσματα προβλήθηκαν σε τυχαία σειρά.

Πίνακας 1: Οι επτά χρωματικοί συνδυασμοί που χρησιμοποιήθηκαν

α/α	Χρώμα Γραμματοσειράς	Χρώμα Υποβάθρου
1	Μαύρο	Άσπρο
2	Κίτρινο	Μπλε
3	Άσπρο	Μαύρο
4	Άσπρο	Μπλε
5	Κόκκινο	Πράσινο
6	Μαύρο	Γκρι
7	Πράσινο	Κίτρινο

Τα τυπογραφικά στοιχεία επιλέχθηκαν σύμφωνα με τις εργασίες [16] και [21]. Τα αποτελέσματα, επίσης, βασίζονται στη στατιστική ανάλυση ελληνικών, αγγλικών βιβλίων και εφημερίδων (η αμερικάνικη έκδοση Metro NY, οι ελληνικές εφημερίδες «Το Βήμα» και «Τα Νέα» και τα σχολικά βιβλία του Α'Βάθμιας και Β'Βάθμιας εκπαίδευσης) σε έντυπη και ηλεκτρονική μορφή. Τα επιλεγμένα τυπογραφικά στοιχεία είναι τα πιο συχνά χρησιμοποιούμενα στοιχεία σε αυτού του είδους τα έγγραφα. Για παράδειγμα, σε σώμα κειμένων της εφημερίδας MetroNY [94] το 10% των λέξεων περίπου είναι σε έντονη γραφή, το 2.4% σε πλάγια ενώ η υπογράμμιση ήταν σχεδόν μηδενική και το υπόλοιπο είναι απλό κείμενο. Το 70% του κειμένου είναι σε «Times New Roman» ενώ το υπόλοιπο σε «Arial». Επίσης, σε αντίστοιχη μελέτη που πραγματοποιήθηκε σε 36 σχολικά βιβλία της Α'Βάθμιας και Β'Βάθμιας εκπαίδευσης [178], στην οποία η πιο συχνά χρησιμοποιούμενη γραμματοσειρά είναι η «Times New Roman» σε σχέση με την «Arial», ο πιο συχνός σε χρήση τρόπος γραφής (πέρα από το απλό κείμενο) είναι η έντονη, η πλάγια και η έντονη-πλάγια γραφή στο σώμα του κειμένου. Το μέγεθος γραμματοσειράς διαφοροποιούταν ανάλογα και τη χρήση του (π.χ. τίτλος, υπότιτλος, λεζάντα, υποσημείωση). Αναλυτικά για το μέγεθος γίνεται αναφορά στο εδάφιο 3.3.1. Μερικά στοιχεία αποκλείστηκαν, όπως για παράδειγμα η υπογράμμιση, διότι είτε χρησιμοποιείται για να υποδηλώσει σύνδεσμο είτε η χρήση της είναι περιορισμένη.

Τα ερεθίσματα ήταν μία παράγραφος με 46 περίπου λέξεις, χρησιμοποιώντας τα προαναφερθέντα τυπογραφικά στοιχεία, από την οποία αποκλείστηκε οποιαδήποτε

πληροφορία εξαρτημένη από το περιεχόμενο και το συναίσθημα - αναλυτικά παρουσιάζεται στο εδάφιο 3.3.2. Προβλήθηκαν σε μία οθόνη LCD, 17 ιντσών, με ανάλυση 1024x768, χρησιμοποιώντας 32 bit βάθος χρώματος και MS Internet Explorer 7 σε πλήρη οθόνη. Για την ανάπτυξη της αυτοματοποιημένης πειραματικής διαδικασίας χρησιμοποιήθηκε PHP [179] σε ένα Apache Web Server [180] και MySQL βάση δεδομένων [181] υπό τις οδηγίες που παρέχονται από το IAPS [182]. Με τη γλώσσα PHP αναπτύχθηκαν δυναμικές ιστοσελίδες για την αυτοματοποιημένη παρουσίαση των ερεθισμάτων, καθώς επίσης και για την αυτόματη καταχώρηση των απαντήσεων των συμμετεχόντων. Οι απαντήσεις αποθηκεύονταν στη βάση δεδομένων (MySQL).

3.2.2 Πειραματική διαδικασία

Κατά την έναρξη της πειραματικής διαδικασίας ζητήθηκε από τους συμμετέχοντες να διαβάσουν με προσοχή τις οδηγίες που τους παρέχονταν και εάν χρειαζόταν να ρωτήσουν για διευκρινήσεις τον επιβλέποντα του πειράματος. Επίσης, τους ζητήθηκε να συμπληρώσουν μια ηλεκτρονική φόρμα με τις προσωπικές πληροφορίες (δημογραφικά στοιχεία) και μία δήλωση ότι συμφωνούν να συμμετάσχουν στην πειραματική διαδικασία. Εν συνεχεία, συμμετείχαν σε μία δοκιμαστική έκδοση του πειράματος με σκοπό την εξοικείωσή τους με αυτή.

Το κάθε ερέθισμα προβλήθηκε για περίπου 15 δευτερόλεπτα και στην αμέσως επόμενη σελίδα οι συμμετέχοντες έπρεπε να αξιολογήσουν τη συναισθηματική τους κατάσταση σε μία κλίμακα 9 σημείων με τους άξονες PAD χρησιμοποιώντας τις φιγούρες που παρέχονται από το τεστ SAM.

Το ερωτηματολόγιο της πειραματικής διαδικασίας αυτοαξιολόγησης SAM αποτελείται από τρεις ομάδες πέντε φιγούρων, για τον κάθε άξονα «Ευχαρίστηση», «Διέγερση» και «Επικράτηση». Στο Σχήμα 2 παρουσιάζεται η οθόνη που φαινόταν σε κάθε συμμετέχοντα μετά την προβολή του ερεθίσματος. Οι συμμετέχοντες έχουν τη δυνατότητα να αξιολογούν και ενδιάμεσες συναισθηματικές καταστάσεις (οι κύκλοι ανάμεσα στις φιγούρες). Αυτή η προσέγγιση επιλέχθηκε με γνώμονα τη μεγαλύτερη λεπτομέρεια στα αποτελέσματα. Η διάρκεια κάθε συνεδρίας ήταν περίπου 20-30 λεπτά, εξαρτώμενη από το χρόνο απάντησης των συμμετεχόντων.

3.2.3 Αποτελέσματα

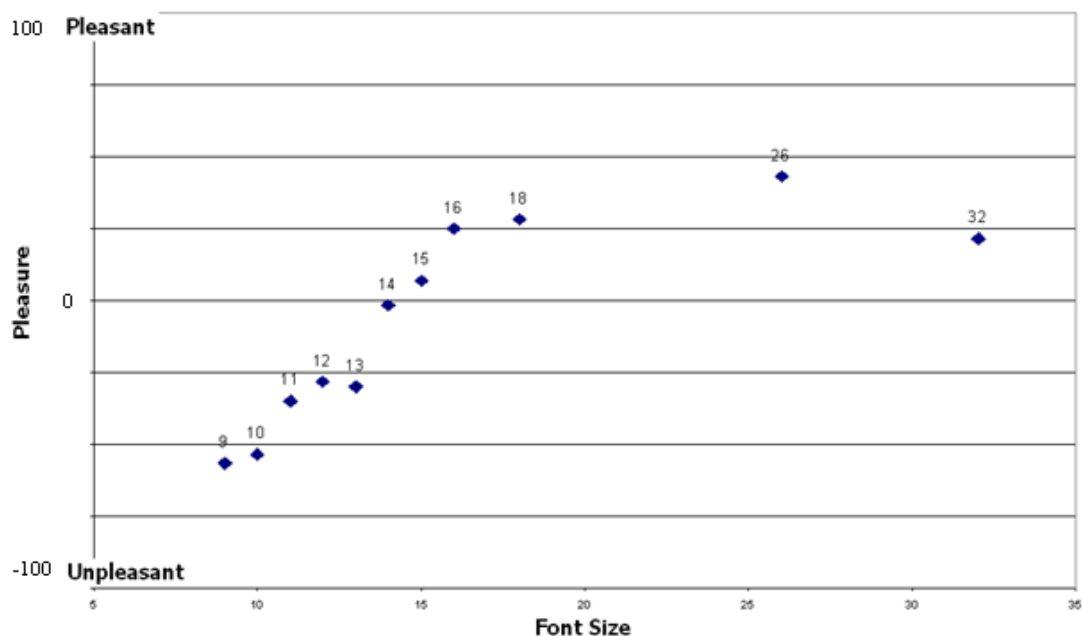
Επαγόμενη συναισθηματική κατάσταση ως προς το μέγεθος της γραμματοσειράς.

Τα αποτελέσματα παρουσιάζονται στα Σχήματα 15α, 15β και 15γ όπου φαίνονται οι μέσες τιμές της κάθε διάστασης ως προς το μέγεθος γραμματοσειράς. Η διάσταση «Ευχαρίστηση» αυξάνεται με την αύξηση του μεγέθους γραμματοσειράς. Για μεγέθη μεγαλύτερα των 26pt ο αναγνώστης αρχίζει να αισθάνεται λιγότερο ευχαριστημένος.

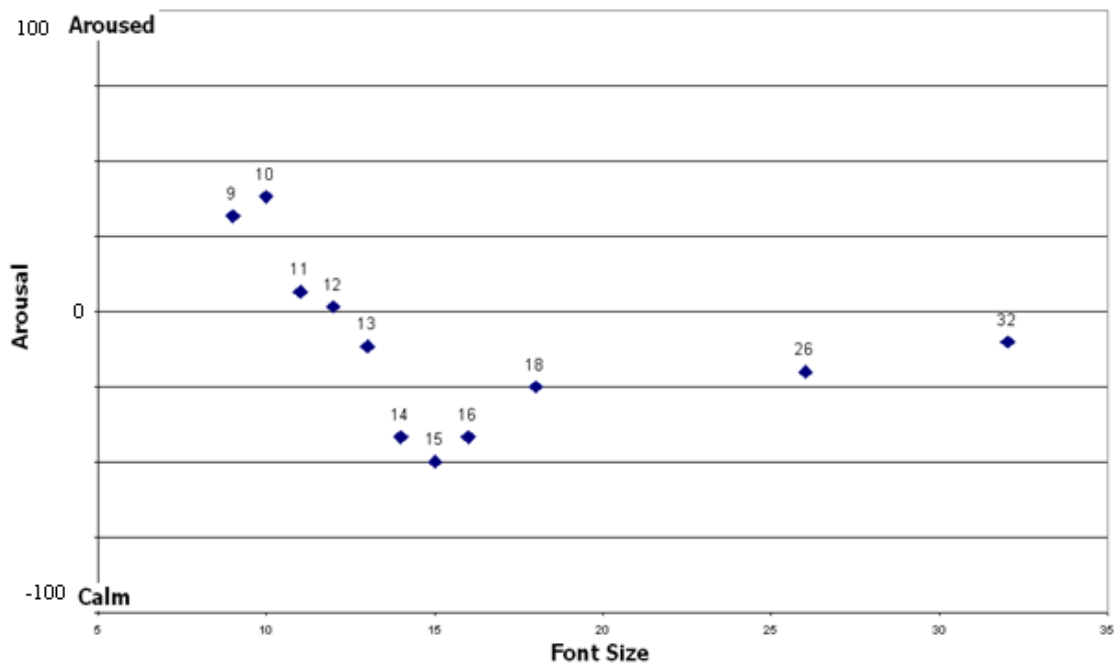
Οι διαστάσεις «Διέγερση» και «Επικράτηση» έχουν παρόμοιες μεταβολές. Η διάσταση «Διέγερση» μειώνεται με την αύξηση του μεγέθους γραμματοσειράς μέχρι 15pt, ενώ από αυτή την τιμή και πάνω έχουμε αύξησή της. Αντίστοιχη, είναι και η συμπεριφορά της διάστασης «Επικράτηση», με μόνη διαφορά ότι η αλλαγή της παρατηρείται για μέγεθος γραμματοσειράς 18pt.

Επαγόμενη συναισθηματική κατάσταση ως προς τον τρόπο γραφής.

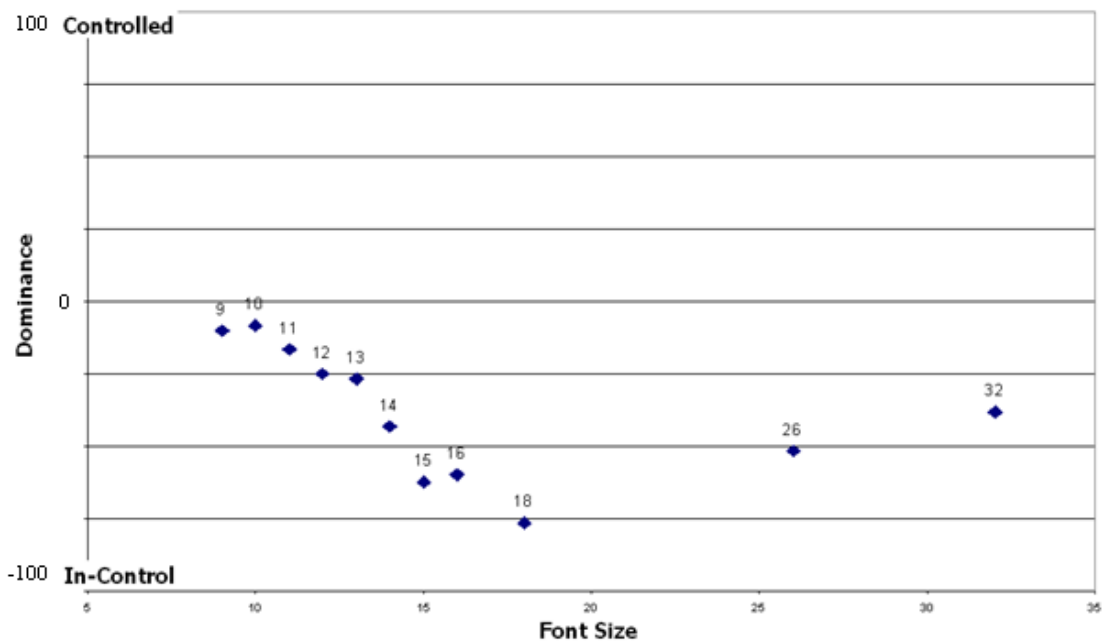
Ο τρόπος γραφής που χρησιμοποιείται είναι το απλό κείμενο, έντονη, πλάγια και έντονη-πλάγια γραφή. Μερικά από αυτά τα ερεθίσματα παρουσιάστηκαν επανειλημμένα για δύο ή τρεις φορές κατά τη διάρκεια της πειραματικής διαδικασίας. Παρατηρήθηκε ότι οι μεταβολές στη συναισθηματική κατάσταση ήταν μικρές και διαφορετικές σε κάθε μία περίπτωση επανάληψης. Αυτό δημιούργησε την υποψία ότι το τρέχον ερέθισμα μπορεί να επηρεάζεται από προηγούμενα προβεβλημένα ερεθίσματα.



(α)



(β)



(γ)

Σχήμα 15: Τα πειραματικά αποτελέσματα της συναισθηματικής κατάστασης α) «Ευχαρίστηση» β), «Διέγερση» (γ), «Επικράτηση» ως προς το μέγεθος της γραμματοσειράς.

Όμως, δεν είναι ασφαλές να βγάλουμε συμπεράσματα για τις συγκεκριμένες μεταβολές, επειδή το δείγμα μας δεν είναι ικανοποιητικά μεγάλο, σε συνδυασμό με τις μικρές μεταβολές που παρατηρούνται. Έτσι, χρησιμοποιώντας τα αποτελέσματα του πιλοτικού πειράματος, είναι δύσκολο να εξάγουμε ασφαλή συμπεράσματα και να μοντελοποιήσουμε την επαγόμενη συναισθηματική απόκριση του αναγνώστη κατά τις μεταβολές του τρόπου γραφής. Επίσης, να τονιστεί ότι σύμφωνα με τις οδηγίες του IAPS, οι αξιολογήσεις των συμμετεχόντων είναι σταθερές είτε κατά την πειραματική διαδικασία είτε μεταξύ διαφορετικών πειραμάτων [182]. Η ανάλυση των αποτελεσμάτων έδειξε ότι δεν υπήρχαν στατιστικώς σημαντικές διαφορές στις απαντήσεις για τα επαναλαμβανόμενα ερεθίσματα. Επομένως, οδηγούμαστε στο συμπέρασμα ότι δεν υπάρχει επιρροή στην αξιολόγηση της συναισθηματικής κατάστασης από το προηγούμενο ερέθισμα.

Επαγόμενη συναισθηματική κατάσταση ως προς το χρώμα και τον τύπο γραμματοσειράς.

Οι συνδυασμοί του χρώματος της γραμματοσειράς και του υποβάθρου στην παρούσα πειραματική διαδικασία παρουσιάζονται στον Πίνακα 1 και τα αποτελέσματα στα Σχήματα 16α και 16β.

Όπως προαναφέρθηκε, οι συγκεκριμένοι χρωματικοί συνδυασμοί χρησιμοποιούνται στη μελέτη του Laarni [16] και παρατηρήθηκαν ομοιότητες στα αποτελέσματα:

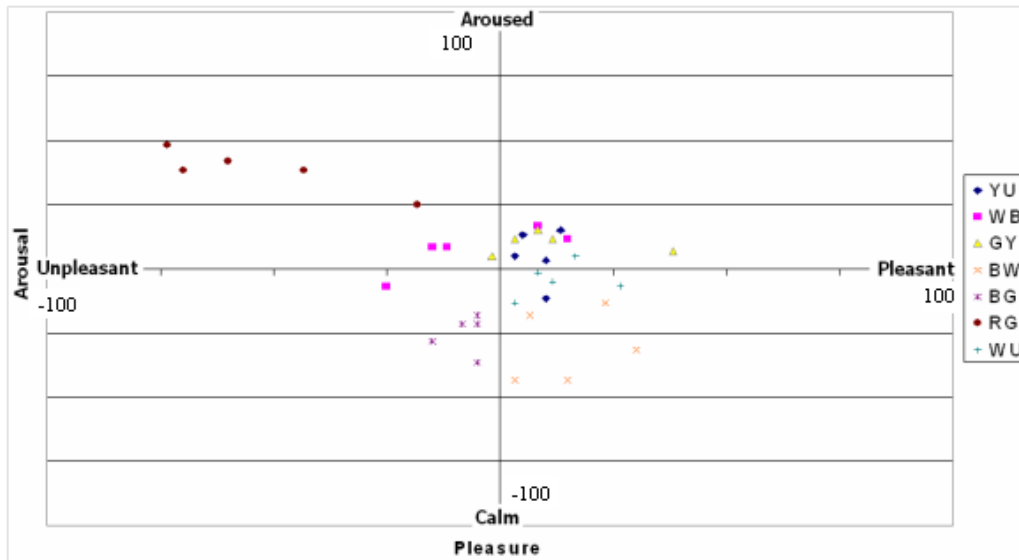
- Ο χρωματικός συνδυασμός κόκκινο σε πράσινο (RG) είναι ο πιο έντονος συναισθηματικά και ο λιγότερο ευχάριστος.
- Ο χρωματικός συνδυασμός μαύρο σε άσπρο (BW) έχει τη μικρότερη μέση τιμή διέγερσης.
- Οι συνδυασμοί άσπρο σε μπλε (WU), πράσινο σε κίτρινο (GY) και μαύρο σε άσπρο (BW) είναι οι πιο ευχάριστοι.

Τα παραπάνω αποτελέσματα μας δίνουν ενδείξεις ότι η επίδραση των χρωματικών συνδυασμών γραμματοσειράς και υποβάθρου στη συναισθηματική κατάσταση του αναγνώστη, είναι ανεξάρτητη της γλώσσας.

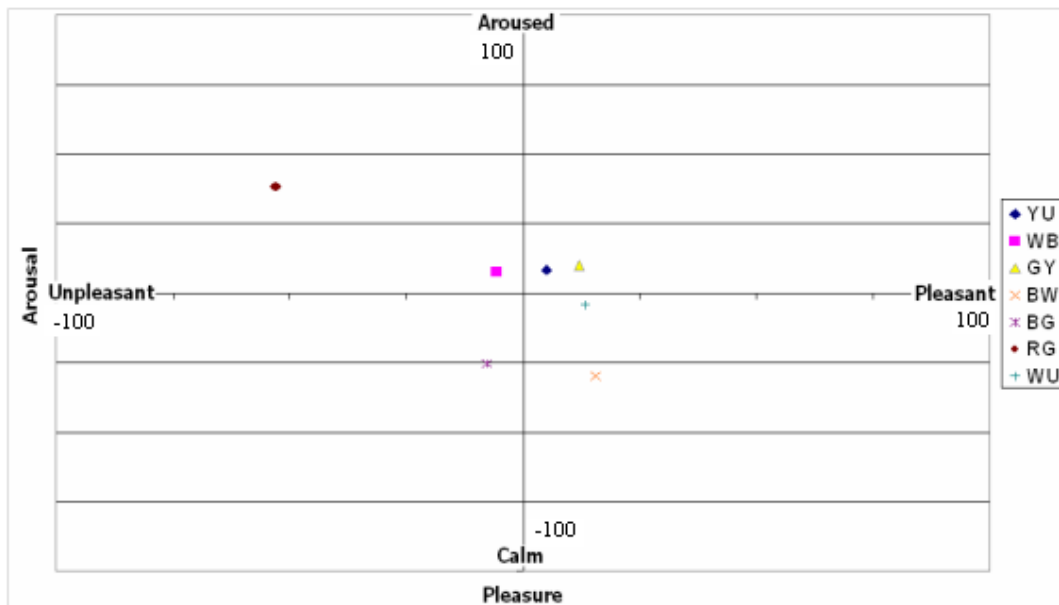
Αποτελέσματα αξιολόγησης

Χρησιμοποιώντας την ανάλυση που περιγράφεται στην [177] δίνεται η δυνατότητα να μελετηθεί το μέτρο της εμπιστοσύνης και αξιοπιστίας των αξιολογητών-συμμετεχόντων (confidence measure for individual evaluators). Η μετρούμενη συναισθηματική

κατάσταση συμβολίζεται με $x_{n,k}^{(i)}$ για τη διάσταση $i \in \{P, A, D\}$, ο δείκτης του ερεθίσματος $1 \leq n \leq N$ και ο δείκτης του αξιολογητή $1 \leq k \leq K$. Η βαθμολογία αξιοπιστίας (confidence score) προκύπτει από τη συσχέτιση r_k μεταξύ μετρούμενων συναισθηματικών καταστάσεων $\{x_{n,k}^{(i)}\}_{n=1,\dots,N}$ και των μέσων τιμών $\{x_n^{MLE,(i)}\}_{n=1,\dots,N}$:



(α)



(β)

Σχήμα 16: Τα πειραματικά αποτελέσματα των διαστάσεων «Ευχαρίστησης» και «Διέγερσης» με α) τις τιμές κάθε ερεθίσματος και β) τις μέσες τιμές των ερεθισμάτων για κάθε χρωματικό συνδυασμό.

$$r_k^{(i)} = \frac{\sum_{n=1}^N (x_{n,k}^{(i)} - \mu_k^{(i)}) \cdot (x_n^{MLE,(i)} - \mu^{MLE,(i)})}{\sqrt{\sum_{n=1}^N (x_{n,k}^{(i)} - \mu_k^{(i)})^2} \cdot \sqrt{\sum_{n=1}^N (x_n^{MLE,(i)} - \mu^{MLE,(i)})^2}} \quad (8)$$

με

$$\mu_k^{(i)} = \frac{1}{N} \cdot \sum_{n=1}^N x_{n,k}^{(i)} \quad (9)$$

$$\mu^{MLE,(i)} = \frac{1}{N} \cdot \sum_{n=1}^N x_n^{MLE,(i)} \quad [\text{MLE} - \text{Maximum Likelihood Estimator}] \quad (10)$$

και

$$x_n^{MLE,(i)} = \frac{1}{K} \cdot \sum_{k=1}^K x_{n,k}^{(i)} \quad (11)$$

Στην περίπτωση του ιδανικού αξιολογητή η τιμή συσχέτισης γίνεται $r_k^{(i)} = 1, \forall k, i \in \{P, A, D\}$. Μη αξιόπιστοι αξιολογητές περιγράφονται ως $r_k^{(i)} < 0, \forall k, i \in \{P, A, D\}$. Οι τιμές που προκύπτουν από την παρούσα πειραματική διαδικασία είναι προφανώς κάτω από το επίπεδο του ιδανικού αξιολογητή, αλλά ταυτόχρονα πολύ υψηλότερες από το κατώτατο όριο, $r_k^{(i)} \gg -1, \forall k, i \in \{P, A, D\}$.

Ο Πίνακας 2 παρουσιάζει τη βαθμολογία αξιοπιστίας r του κάθε αξιολογητή, για κάθε μία συναισθηματική διάσταση χρησιμοποιώντας την εξίσωση 8.

3.2.4 Συμπεράσματα

Τα αποτελέσματα της μελέτης αξιοπιστίας των συμμετεχόντων κατά την πιλοτική πειραματική διαδικασία είναι ενθαρρυντικά για την περαιτέρω χρήση της πειραματικής διαδικασίας. Παρατηρούμε ότι όλες οι απαντήσεις των συμμετεχόντων συνεισφέρουν στα αποτελέσματα του πειράματος και για τις τρεις διαστάσεις ($r_k^{(i)} \gg 0, \forall k, i \in \{P, A, D\}$) με μέσες τιμές $\bar{r}_k^{(P)} = 0.59$, $\bar{r}_k^{(A)} = 0.51$, $\bar{r}_k^{(D)} = 0.53$. Αυτό μας ωθεί σε περαιτέρω μελέτη με τη χρήση μεγαλύτερου αριθμού συμμετεχόντων, προκειμένου να εξαχθούν περισσότερα και πιο ασφαλή συμπεράσματα.

Μερικά συμπεράσματα αναφορικά με την υλοποίηση του συγκεκριμένου πιλοτικού πειράματος και τη μεταβολή της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη είναι:

Πίνακας 2: Βαθμολογία αξιοπιστίας κάθε συμμετέχοντα για κάθε μία συναισθηματική διάσταση.

k	$r_k^{(P)}$	$r_k^{(A)}$	$r_k^{(D)}$	Ηλικία	Γένος
1	0,61	0.78	0.38	28	γυναίκα
2	0.46	0.09	0.36	19	άνδρας
3	0.58	0.73	0.66	23	γυναίκα
4	0.77	0.48	0.6	24	άνδρας
5	0.67	0.46	0.78	26	γυναίκα
6	0.67	0.14	0.36	26	γυναίκα
7	0.39	0.78	0.49	33	άνδρας
8	0.43	0.37	0.61	31	άνδρας
9	0.56	0.71	0.66	31	γυναίκα
10	0.68	0.65	0.66	28	άνδρας
11	0.57	0.47	0.3	25	γυναίκα
12	0.71	0.73	0.48	29	γυναίκα
13	0.44	0.41	0.2	18	άνδρας
14	0.69	0.69	0.77	27	άνδρας
15	0.63	0.19	0.53	26	άνδρας
$\bar{r}_k^{(i)}$	0.59	0.51	0.53		

- Χρησιμοποιώντας διάφορα μεγέθη γραμματοσειράς, οι τιμές για τη διάσταση «Ευχαρίστηση» αυξάνονται μέχρι 18pt και μειώνονται λίγο μέχρι 32pt. Σε αντίθεση, οι τιμές των διαστάσεων «Διέγερση» και «Επικράτηση» μειώνονται μέχρι το μέγεθος 15pt και 18pt αντίστοιχα, και παρατηρείται μια μικρή αύξηση μέχρι το μέγεθος 32pt.
- Τα αποτελέσματα για τον τρόπο γραφής και είδους γραμματοσειράς χρειάζεται να μελετηθούν περαιτέρω και πιο διεξοδικά. Δεν μπορούν να εξαχθούν ασφαλή συμπεράσματα λόγω α) των μικρών μεταβολών των τιμών κάθε συναισθηματικής διάστασης και β) του μικρού αριθμού συμμετεχόντων στην πειραματική διαδικασία.

Οι οδηγίες που παρέχονται από το IAPS είναι για την έκδοση της πειραματικής διαδικασίας «paper-and-pencil» SAM. Στην παρούσα μελέτη, υλοποιήθηκε ως μία εφαρμογή σε υπολογιστή (Web application). Αυτό επέτρεψε να διεξαχθεί εύκολα το

πείραμα και να συγκεντρωθούν γρήγορα και εύκολα τις απαντήσεις των συμμετεχόντων για τη στατιστική τους μελέτη.

Επίσης, προέκυψαν κάποια συμπεράσματα χρήσιμα για το σχεδιασμό του πειράματος είτε σε επίπεδο διαδικασίας είτε σε τεχνικό επίπεδο. Τα ερεθίσματα ήταν αρκετά σε αριθμό, με συνέπεια μερικοί συμμετέχοντες να κουράζονται προς το τέλος της διαδικασίας. Έτσι, επανεξετάστηκε ο αριθμός των ερεθισμάτων, ο χρόνος προβολής τους, καθώς και το αν χρειάζεται να διασπαστεί η ενιαία πειραματική διαδικασία σε μικρότερες, δίνοντας στους συμμετέχοντες τη δυνατότητα να ξεκουραστούν και να συνεχίσουν απρόσκοπτα. Οι αποφάσεις που πρέπει να παρθούν φαίνονται στο σχεδιασμό της κύριας πειραματικής διαδικασίας (παράγραφος 3.3).

3.3 Πειραματική διαδικασία

Μετά την ολοκλήρωση του πιλοτικού πειράματος και την καταγραφή των προβλημάτων, έγινε διόρθωση και βελτιστοποίηση της πειραματικής διαδικασίας. Σε αυτό το μέρος του κεφαλαίου παρέχεται η λεπτομερής περιγραφή του σχεδιασμού του πειράματος, η διαδικασία διεξαγωγής του, η στατιστική ανάλυση των αποτελεσμάτων και η μαθηματική μοντελοποίησή τους. Τέλος, γίνεται μια εκτενής συζήτηση επί των αποτελεσμάτων.

3.3.1 Ερευνητικό Μοντέλο - Υπόθεση

Περιλαμβάνονται τέσσερις υποθέσεις ($H_i, i = 1,2,3,4$) των οποίων οι ανεξάρτητες μεταβλητές ορίζονται ως:

1. Ο χρωματικός συνδυασμός γραμματοσειράς/υποβάθρου.
2. Το μέγεθος γραμματοσειράς.
3. Ο τύπος γραμματοσειράς.
4. Ο τρόπος γραφής.

Οι τρεις συναισθηματικές διαστάσεις «Ευχαρίστηση», «Διέγερση» και «Επικράτηση» αποτελούν τις εξαρτημένες μεταβλητές.

Υπόθεση 1 (H₁).

Οι χρωματικοί συνδυασμοί γραμματοσειράς/υποβάθρου, το είδος της γραμματοσειράς και ο τρόπος γραφής επηρεάζουν τις συναισθηματικές διαστάσεις «Ευχαρίστηση», «Διέγερση» και «Επικράτηση».

Επιλέχθηκαν οι παρακάτω επτά χρωματικοί συνδυασμοί γραμματοσειράς/υποβάθρου: Κίτρινο σε μπλε (Y/U), άσπρο σε μαύρο (W/B), άσπρο σε μπλε (W/U), κόκκινο σε πράσινο (R/G), μαύρο σε γκρι (B/G), μαύρο σε άσπρο (B/W) και πράσινο σε κίτρινο (G/Y). Οι συγκεκριμένοι χρωματικοί συνδυασμοί προτάθηκαν από τους Hill και Scharff [13] και επίσης χρησιμοποιήθηκαν από τον Laarni [16]. Με σκοπό να μελετηθεί ο τρόπος που οι χρωματικοί συνδυασμοί γραμματοσειράς/υποβάθρου επηρεάζουν τη συναισθηματική κατάσταση, χρησιμοποιώντας διάφορους τύπους γραμματοσειράς και τρόπους γραφής, μελετάμε τους συνδυασμούς χρησιμοποιώντας μέγεθος γραμματοσειράς 16 px (5.28 mm) (η μαθηματική συσχέτιση των μονάδων px και pt περιγράφεται στο εδάφιο 4.2), κείμενο σε απλή, έντονη και έντονη-πλάγια γραφή για την περίπτωση του «Times New Roman» και έντονη, έντονη-πλάγια γραφή για «Arial».

Οι τιμές για την χρωματική φωτεινότητα και αντίθεση (Πίνακας 3β) προέκυψαν από τη σύσταση της W3C [71] [72] σύμφωνα με τις εξισώσεις:

$$Brightness = \frac{R \times 299 + G \times 587 + B \times 114}{1000} \quad (12)$$

$$Contrast = |R_2 - R_1| + |G_2 - G_1| + |B_2 - B_1| \quad (13)$$

όπου: R - κόκκινο, G - πράσινο, B – μπλε δίνονται στην κλίμακα RGB.

Το εύρος τιμών της χρωματικής φωτεινότητας εκτείνεται από 0 έως 255 και της αντίθεσης από 0 έως 765. Η σύσταση της W3C για την ελάχιστη τιμή της χρωματικής φωτεινότητας είναι 125 και για τη χρωματική αντίθεση είναι 500. Έτσι, οι χρωματικοί συνδυασμοί μαύρο σε άσπρο, άσπρο σε μαύρο, κίτρινο σε μπλε και άσπρο σε μπλε θεωρούνται υψηλής αντίθεσης χρωματικοί συνδυασμοί, ενώ οι συνδυασμοί κόκκινο σε πράσινο, μαύρο σε γκρι και πράσινο σε κίτρινο θεωρούνται χαμηλής αντίθεσης. Ενδεικτικά αποτελέσματα που προέκυψαν από άλλες μελέτες (που παρουσιάζονται εκτενώς στην παράγραφο 3.3.5) μας οδηγούν στο συμπέρασμα ότι χρωματικοί

συνδυασμοί υψηλής αντίθεσης αναμένεται να είναι πιο ευχάριστοι, πιο ήρεμοι και να μειώνουν την τιμή της «Επικράτησης».

Πίνακας 3: Οι χρωματικοί συνδυασμοί γραμματοσειράς / υποβάθρου που χρησιμοποιούνται στην παρούσα εργασία: (α) τα χρώματα σύμφωνα με τον χρωματικό χώρο αντίληψης του Munsell (perceptual color space) και (β) στον χρωματικό χώρο RGB μαζί με τις αντίστοιχες τιμές διαφοράς χρωματικής φωτεινότητας του ζεύγους χρωμάτων / χρωματικής αντίθεσης.

(α)

Χρωματικός συνδυασμός γραμματοσειράς/υποβάθρου	Χρώμα Γραμματοσειράς	Χρώμα Υποβάθρου
Μαύρο σε Άσπρο (B/W)	N 0.0/0.0	8.6RP 9.8/0.6
Άσπρο σε Μαύρο (W/B)	8.6RP 9.8/0.6	N 0.0/0.0
Κίτρινο σε Μπλε (Y/U)	0.2GY 9.5/12.7	7.0PB 3.2/29.4
Κόκκινο σε Πράσινο (R/G)	7.9R 5.2/20.7	9.6GY 4.5/11.5
Μαύρο σε Γκρι (B/G)	N 0.0/0.0	N 5.2/0.0
Πράσινο σε Κίτρινο (G/Y)	9.6GY 4.5/11.5	0.2GY 9.5/12.7
Άσπρο σε Μπλε (W/U)	8.6RP 9.8/0.6	7.0PB 3.2/29.4

(β)

Χρωματικός συνδυασμός γραμματοσειράς/υποβάθρου	Χρώμα Γραμματοσειράς			Χρώμα Υποβάθρου			Χρώμα	
	R	G	B	R	G	B	ΔBrightness	Contrast
Μαύρο σε Άσπρο (B/W)	0	0	0	255	255	255	255	765
Άσπρο σε Μαύρο (W/B)	255	255	255	0	0	0	255	765
Κίτρινο σε Μπλε (Y/U)	255	255	0	0	0	255	197	765
Κόκκινο σε Πράσινο (R/G)	255	0	0	0	128	0	1	383
Μαύρο σε Γκρι (B/G)	0	0	0	128	128	128	128	384
Πράσινο σε Κίτρινο (G/Y)	0	128	0	255	255	0	151	382
Άσπρο σε Μπλε (W/U)	255	255	255	0	0	255	226	510

Υπόθεση 2 (H₂).

Το μέγεθος της γραμματοσειράς με μαύρο χρώμα, σε άσπρο υπόβαθρο και «Times New Roman» επηρεάζει τις συναισθηματικές διαστάσεις «Ευχαρίστηση», «Διέγερση» και «Επικράτηση».

Οι Bernard et al. [183] έδειξαν ότι κείμενο σε μέγεθος 12pt στην οθόνη του υπολογιστή είναι προτιμότερο από ότι μέγεθος 10pt. Οι Schmidt et al. [184] συμπέραναν ότι αύξηση του μεγέθους γραμματοσειράς (από 7.5 pt σε 14 pt) προκαλεί αύξηση στην προτίμηση του χρήστη και συνεισφέρει στην ευκολία αλληλεπίδρασης με τις ιστοσελίδες. Οι Alsumait et al. [185] μελέτησαν τις προτιμήσεις των παιδιών Αραβικής καταγωγής, χρησιμοποιώντας προγράμματα ηλεκτρονικής εκμάθησης. Συμπέραναν ότι το μέγεθος γραμματοσειράς 14 pt είναι προτιμότερο από το 12 pt και το είδος γραμματοσειράς «Simplified Arabic» (εύκολο και γρήγορο κατά την ανάγνωση) και «Arial» (ελκυστικό) ήταν προτιμότερα, μεταξύ πέντε διαφορετικών ειδών γραμματοσειράς.

Διατηρώντας σταθερό το χρωματικό συνδυασμό γραμματοσειράς/υποβάθρου (μαύρο σε άσπρο), τον τύπο γραμματοσειράς (Times New Roman) και τον τρόπο γραφής, μελετήθηκαν οι μεταβολές των τριών συναισθηματικών διαστάσεων ως προς το μέγεθος γραμματοσειράς. Τα μεγέθη γραμματοσειράς που χρησιμοποιήθηκαν είναι (σε παρένθεση δίνεται το αντίστοιχο φυσικό μέγεθος): 10 px (3.3 mm), 11 px (3.63 mm), 12 px (3.96 mm), 13 px (4.29 mm), 14 px (4.62 mm), 15 px (4.95 mm), 18 px (5.94 mm), 26 px (8.58 mm), και 32 px (10.56 mm). Αυτά τα μεγέθη γραμματοσειράς επιλέχθηκαν ως τα συχνότερα χρησιμοποιούμενα, στη στατιστική ανάλυση που παρουσιάστηκε προηγουμένως στο εδάφιο 3.2.1 [94].

Υπόθεση 3 (H₃).

Ο τύπος γραμματοσειράς, χρησιμοποιώντας απλό κείμενο, έντονη, πλάγια ή έντονη-πλάγια γραφή επηρεάζει τις συναισθηματικές διαστάσεις «Ευχαρίστηση», «Διέγερση» και «Επικράτηση».

Οι Bernard et al. [183] μελέτησαν σε οθόνη υπολογιστή τις προτιμήσεις του χρήστη, χρησιμοποιώντας ως τύπους γραμματοσειράς τις «Arial» και «Times New Roman». Συμπέραναν ότι ο προτιμότερος τύπος γραμματοσειράς είναι ο «Arial», χάρη στην υψηλού βαθμού αναγνωσιμότητα στην οθόνη του υπολογιστή. Παρόλα αυτά, αποτελέσματα από την ίδια εργασία [183] υποδεικνύουν ότι το 57% των συμμετεχόντων στην καθημερινή τους ζωή χρησιμοποιούν 12pt «Times New Roman», ενώ μόνο το 8.6% χρησιμοποιούν 12pt «Arial». Αποτελέσματα των Ling και van Schaik [186]

αποκαλύπτουν ότι οι συμμετέχοντες έδειξαν προτίμηση σε «Arial», παρά σε «Times New Roman».

Αντίθετα, τα αποτελέσματα του Laarni [16] έδειξαν ότι η γραμματοσειρά «Times New Roman», μαύρη σε άσπρο υπόβαθρο, θεωρείται ως η πιο ευανάγνωστη και ευχάριστη. Η γραμματοσειρά «Arial» θεωρείται ως η πιο ευανάγνωστη και ευχάριστη για τις περιπτώσεις «πλάγια, άσπρη γραφή σε μπλε υπόβαθρο» και «απλό, άσπρο κείμενο σε μαύρο υπόβαθρο». Επιπλέον, ως πιο ευχάριστοι και λιγότερα έντονοι χρωματικοί συνδυασμοί, προτείνονται οι παρακάτω επιλογές: (α) άσπρο, απλό κείμενο σε μπλε υποβάθρου, «Times New Roman» και (β) μαύρο, πλάγια γραφή σε λευκό υπόβαθρο, «Times New Roman».

Στην παρούσα μελέτη, τα υπό εξέταση είδη γραμματοσειράς είναι «Times New Roman» και «Arial» σε απλή, έντονη, πλάγια ή έντονη-πλάγια γραφή, με χρωματικό συνδυασμό μαύρο σε άσπρο υπόβαθρο με σταθερό το μέγεθος της γραμματοσειράς (16px). Έχουν επιλέξει αυτοί οι συγκεκριμένοι τύποι γραμματοσειράς, ως οι πιο αντιπροσωπευτικοί από τις ομάδες-οικογένειες τύπου γραμματοσειρών «Serif» και «Sans-Serif», αντίστοιχα. Η επιλογή των συγκεκριμένων οικογενειών γραμματοσειράς προέκυψε μετά από εκτεταμένη στατιστική ανάλυση σε ένα σώμα εγγράφων από Αγγλικά και Ελληνικά βιβλία και εφημερίδες σε έντυπη και ηλεκτρονική μορφή [94]. Σύμφωνα με τη στατιστική ανάλυση, περί το 83% του κειμένου (μαύρο σε άσπρο υπόβαθρο) στο σώμα του εγγράφου είναι σε «Times New Roman» και μόνο το 17% είναι σε «Arial».

Η επιβεβαίωση ή όχι της υπόθεσης 3 για κάθε μία από τις τρεις διαστάσεις θα δείξει: (α) την εξάρτηση της συναισθηματικής κατάστασης από τον τύπο της γραμματοσειράς και (β) τη μεταβολή της συναισθηματικής κατάστασης κατά την εναλλαγή του τύπου της γραμματοσειράς.

Υπόθεση 4 (H₄).

Ο τρόπος γραφής, χρησιμοποιώντας «Arial» ή «Times New Roman», επηρεάζει τις συναισθηματικές διαστάσεις «Ευχαρίστηση», «Διέγερση» και «Επικράτηση».

Όπως αναφέρθηκε στην υπόθεση 3, οι συνδυασμοί «μαύρο, απλό κείμενο σε άσπρο υπόβαθρο, Times New Roman» και «μαύρο, πλάγιο κείμενο σε άσπρο υπόβαθρο, Times New Roman» είναι καλές επιλογές για επαγωγή ευχάριστων συναισθημάτων. Στην παρούσα μελέτη, χρησιμοποιούνται τέσσερις τρόποι γραφής: απλό κείμενο, έντονη, πλάγια, έντονη-πλάγια γραφή σε «Times New Roman» και «Arial», με μαύρο χρώμα γραμματοσειράς σε άσπρο υπόβαθρο και διατηρώντας το ίδιο μέγεθος μεταξύ

των συναισθημάτων (16 ρχ). Μελετάται: (α) η επιρροή του τρόπου γραφής στις τρεις διαστάσεις συναισθηματικής κατάστασης και (β) οι μεταβολές τους.

3.3.2 Μεθοδολογία

Εφόσον οι μεταβολές των συναισθημάτων και των συναισθηματικών καταστάσεων είναι συνεχείς και μη-ακραίες (continuous - non-extreme), επιλέχθηκε η διαστατική προσέγγιση και η πειραματική διαδικασία αυτοαξιολόγησης με τη χρήση φιγούρων - Self Assessment Manikin Test. Σε αυτή τη φάση χρησιμοποιήθηκε η αυτοματοποιημένη πειραματική διαδικασία που σχεδιάστηκε και υλοποιήθηκε για το πιλοτικό πείραμα - σύμφωνα με τις οδηγίες-προδιαγραφές του International Affective Picture System [182] - με μικρές αλλαγές στις παραμέτρους της, όπως για παράδειγμα στον αριθμό των ερεθισμάτων. Η πειραματική διαδικασία πραγματοποιήθηκε με τη χρήση ενός υπολογιστή, τα ερεθίσματα προβάλλονταν σε μία οθόνη τεχνολογίας LCD και οι απαντήσεις του χρήστη καταγράφονταν απευθείας στον Η/Υ. Οι προδιαγραφές της οθόνης ήταν: 17 ίντσες, τεχνολογίας LCD, με ανάλυση 1024x768 και μέγιστη φωτεινότητα (luminance) 300 cd/m², Hewlett-Packard L1706. Η φωτεινότητα της οθόνης κατά τη διάρκεια του τεστ ήταν 140 cd/m² και το επίπεδο φωτισμού του δωματίου (room illuminance) μετρήθηκε στα 105 lux, με τη χρήση ενός ψηφιακού μετρητή φωτεινότητας (digital illuminance meter - E-SUN LX-101) στο ύψος του παρατηρητή, στο κατακόρυφο επίπεδο, και με την οθόνη κλειστή. Η απόσταση μεταξύ του κέντρου της οθόνης και των ματιών του συμμετέχοντα ήταν 70±5 cm.

1. Οι συμμετέχοντες συμπλήρωσαν μία φόρμα με τα προσωπικά τους στοιχεία για τη συλλογή δημογραφικών πληροφοριών: ηλικία, φύλο, επίπεδο μόρφωσης, προβλημάτων - απώλειας όρασης, καθώς και πόση ώρα χρησιμοποιούν καθημερινά τον Η/Υ. Επίσης, συμπλήρωσαν μια δήλωση συγκατάθεσης για τη συμμετοχή τους στο πείραμα.
2. Κάθε ερέθισμα παρουσιαζόταν με τη χρήση συγκεκριμένης παραγράφου με συναισθηματικά ουδέτερο περιεχόμενο.
3. Κάθε ερέθισμα παρουσιαζόταν στον χρήστη σε τυχαία σειρά με σταθερή χρονική διάρκεια (15 δευτερόλεπτα).
4. Μετά την προβολή κάθε ερεθίσματος, ο συμμετέχοντας συμπλήρωνε τις απαντήσεις του με τη χρήση των φιγούρων.

5. Οι απαντήσεις καταχωρούνταν με τέτοιο τρόπο, ώστε να δίνεται η δυνατότητα εύκολης ανάκτησής τους κατά το στάδιο της ανάλυσης.

Ο σκοπός της πειραματικής διαδικασίας ήταν να αξιολογηθούν οι συναισθηματικές καταστάσεις των συμμετεχόντων ατόμου είχαν διαβάσει μία σύντομη παράγραφο που παρουσιαζόταν στην οθόνη σε διάφορους συνδυασμούς τυπογραφικών στοιχείων. Η διάρκεια ήταν περίπου 20 με 30 λεπτά. Κατά τη διενέργεια του πειράματος ήταν παρόν και ένας ερευνητής με σκοπό (α) να παρέχει συμπληρωματικές επεξηγήσεις για τη διαδικασία και (β) να επιβλέπει την ορθή εκτέλεση του πειράματος.

Συμμετέχοντες

Στην πειραματική διαδικασία συμμετείχαν 30 άτομα, με ηλικιακό εύρος από 18 έως 33 έτη (μέση ηλικία 25.5 έτη). Ήταν προπτυχιακοί και μεταπτυχιακοί φοιτητές του Τμήματος Πληροφορικής και Τηλεπικοινωνιών, 16 άνδρες και 14 γυναίκες, και η μητρική τους γλώσσα ήταν η ελληνική. Όλοι τους δήλωσαν ότι είχαν φυσιολογική ή διορθωμένη στο φυσιολογικό οπτική οξύτητα. Όποιος υποψήφιος συμμετέχοντας είχε αχρωματοψία – δυσχρωματοψία, αποκλείστηκε από την πειραματική διαδικασία.

Ερεθίσματα

Ετέθει ο συμβιβασμός να μην είναι μεγάλος ο αριθμός των ερεθισμάτων, με σκοπό να μη κουράζονται οι συμμετέχοντες. Επιλέχθηκαν σαράντα επτά συνδυασμοί γραμματοσειράς και τρόπου γραφής, χρησιμοποιώντας ελληνικό κείμενο (μία παράγραφος με 46 λέξεις περίπου) από το οποίο αποκλείστηκε οποιαδήποτε πληροφορία εξαρτημένη από το περιεχόμενο και το συναίσθημα (η παράγραφος αξιολογήθηκε ως συναισθηματικά ουδέτερη στην πιλοτική εργασία [94]). Το κείμενο ήταν παρόμοιο με αυτό που χρησιμοποιήθηκε από τους Hill and Scharff [13]. Όλα τα ερεθίσματα προβλήθηκαν σε τυχαία σειρά, σε πλήρη οθόνη, χρησιμοποιώντας 32 bit βάθος χρώματος και MS Internet Explorer έκδοση 7. Στο Σχήμα 17 παρουσιάζονται δύο παραδείγματα ερεθισμάτων που χρησιμοποιήθηκαν.

Οι φιγούρες (The manikins)

Η δοκιμασία SAM περιελάμβανε κλίμακα 9 σημείων για κάθε συναισθηματική διάσταση «Ευχαρίστηση», «Διέγερση» και «Επικράτηση». Στο Σχήμα 2 φαίνεται ένα στιγμιότυπο της οθόνης που παρουσιάστηκε σε κάθε συμμετέχοντα, μετά την προβολή ενός ερεθίσματος (οι λέξεις «Ευχαρίστηση», «Διέγερση» και «Επικράτηση» δεν εμφανιζόντουσαν). Οι συμμετέχοντες είχαν τη δυνατότητα να αξιολογήσουν τις επαγόμενες συναισθηματικές καταστάσεις (οι κύκλοι μεταξύ των φιγούρων) για κάθε

ερέθισμα. Αυτή η προσέγγιση (κλίμακα 9 σημείων) επιλέχθηκε με σκοπό τη μεγαλύτερη ανάλυση-λεπτομέρεια στα αποτελέσματα.

3.3.3 Πειραματική διαδικασία

Κατά την έναρξη της πειραματικής διαδικασίας, ζητήθηκε από τους συμμετέχοντες να διαβάσουν με προσοχή τις οδηγίες που παρέχονται από το IAPS [182] και να ρωτήσουν τον εκπαιδευτή-επιβλέποντα για πιθανές διευκρινήσεις. Στη συνέχεια, τους ζητήθηκε να συμπληρώσουν μία ηλεκτρονική φόρμα με τα προσωπικά τους στοιχεία και μία δήλωση ότι συμφωνούν να συμμετάσχουν στην πειραματική διαδικασία. Έπειτα, συμμετείχαν σε μία δοκιμαστική έκδοση του πειράματος με σκοπό να εξοικειωθούν με τη διαδικασία. Κάθε ερέθισμα προβλήθηκε για 15 δευτερόλεπτα, πριν προχωρήσουν στην αξιολόγηση 9 σημείων χρησιμοποιώντας τις φιγούρες που παρέχονται από τη δοκιμασία SAM.

Σε αυτή την έρευνα προσπαθούμε να αξιολογήσουμε πώς το μέγεθος των γραμμάτων ενός κειμένου, το είδος της γραμματοσειράς και ο τρόπος γραφής (απλή, πλάγια, έντονη, έντονη-πλάγια) που χρησιμοποιείται στο κείμενο και οι χρωματικοί συνδυασμοί του κειμένου και του υπόβαθρου του επηρεάζουν τα συναισθήματα του αναγνώστη.

(α)

Σε αυτή την έρευνα προσπαθούμε να αξιολογήσουμε πώς το μέγεθος των γραμμάτων ενός κειμένου, το είδος της γραμματοσειράς και ο τρόπος γραφής (απλή, πλάγια, έντονη, έντονη-πλάγια) που χρησιμοποιείται στο κείμενο και οι χρωματικοί συνδυασμοί του κειμένου και του υπόβαθρου του επηρεάζουν τα συναισθήματα του αναγνώστη.

(β)

Σχήμα 17: Δύο δείγματα ερεθισμάτων όπως παρουσιάστηκαν κατά την πειραματική διαδικασία. (α) Ελληνικό κείμενο σε έντονη γραφή, «Times New Roman» με μέγεθος γραμματοσειράς 16px και (β) ελληνικό απλό κείμενο, «Times New Roman» με μέγεθος γραμματοσειράς 26px.

3.3.4 Αποτελέσματα

Υπόθεση 1 (H₁).

Οι μέσες τιμές και το τυπικό σφάλμα των απαντήσεων των συμμετεχόντων για κάθε ένα από τα 35 ερεθίσματα (από τα 47 συνολικά - περιέχουν τους χρωματικούς συνδυασμούς και αριθμημένα σύμφωνα με τις παρακάτω πέντε ομάδες) παρουσιάζονται στα Σχήματα 18α, 18β και 18γ για την «Ευχαρίστηση», «Διέγερση» και «Επικράτηση», αντίστοιχα.

1. Ερεθίσματα 1–7: Απλό κείμενο, «Times New Roman» (P/TR): Y/U (Κίτρινο σε Μπλε), W/B (Άσπρο σε Μαύρο), W/U (Άσπρο σε Μπλε), R/G (Κόκκινο σε Πράσινο), B/G (Μαύρο σε Γκρι), B/W (Μαύρο σε Άσπρο), G/Y (Πράσινο σε Κίτρινο).
2. Ερεθίσματα 8–14: Έντονη γραφή, «Times New Roman» (B/TR): Y/U, W/B, W/U, R/G, B/G, B/W, G/Y.
3. Ερεθίσματα 15–21: Έντονη γραφή, «Arial» (B/A): Y/U, W/B, W/U, R/G, B/G, B/W, G/Y.
4. Ερεθίσματα 22–28: Έντονη-Πλάγια γραφή, «Times New Roman» (BI/TR): Y/U, W/B, W/U, R/G, B/G, B/W, G/Y.
5. Ερεθίσματα 29–35: Έντονη-Πλάγια γραφή, «Arial» (BI/A): Y/U, W/B, W/U, R/G, B/G, B/W, G/Y.

όπου:

P/TR = Plain text/Times New Roman

B/TR = Bold/Times New Roman

B/A = Bold/Arial

BI/TR = Bold-Italics/Times New Roman

BI/A = Bold-Italics/Arial

Y/U = Yellow on Blue

W/B = White on Black

W/U = White on Blue

R/G = Red on Green

B/G = Black on Gray

B/W = Black on White

G/Y = Green on Yellow

Διενεργήθηκαν τρεις πολυπαραγοντικές επαναλαμβανόμενων μετρήσεων (multifactorial repeated measures) ANOVA (1 three-way και 2 two-way r-m ANOVA) για κάθε μία από τις εξαρτημένες μεταβλητές («Ευχαρίστηση», «Διέγερση» και «Επικράτηση»). Συγκεκριμένα:

Three-Way ANOVA

Διενεργήθηκε 3-way r-m ANOVA (για κάθε μια από τις εξαρτημένες μεταβλητές), χρησιμοποιώντας τρεις ανεξάρτητες μεταβλητές: χρωματικός συνδυασμός γραμματοσειράς / υποβάθρου (7 επίπεδα - χρωματικοί συνδυασμοί όπως αναφέρθηκε παραπάνω), τύπος γραμματοσειράς (2 επίπεδα – «Times New Roman» και «Arial»), και τρόπος γραφής (2 επίπεδα – έντονη και έντονη-πλάγια γραφή). Τα αποτελέσματα παρουσιάζονται στον Πίνακα 4.

Συνοψίζοντας προκύπτει για τα αποτελέσματα ότι:

1. «Ευχαρίστηση»: είναι στατιστικώς σημαντικά για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου, $F(6,24)=7.668$, $p<0.05$, για τον τύπο γραμματοσειράς, $F(1,29)=22.146$, $p<0.05$, και μη-σημαντικά για τον τρόπο γραφής $F(1,29)=0.680$, $p>0.05$. Η αλληλεπίδραση είναι στατιστικώς σημαντική για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου με τον τύπο γραμματοσειράς, $F(6,24)=7.406$, $p<0.05$, για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου με τον τρόπο γραφής, $F(6,24)=6.436$, $p<0.05$ και μη-σημαντική για τον τύπο γραμματοσειράς με τον τρόπο γραφής, $F(1,29)=0.785$, $p>0.05$. Η αλληλεπίδραση μεταξύ των τριών μεταβλητών (χρωματικοί συνδυασμοί γραμματοσειράς/υποβάθρου με τον τύπο γραμματοσειράς και με τον τρόπο γραφής) είναι επίσης στατιστικώς σημαντική, $F(6,24)=5.213$, $p<0.05$.
2. «Διέγερση»: είναι στατιστικώς μη-σημαντικά για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου, $F(6,24)=1.950$, $p>0.05$, για τον τύπο γραμματοσειράς, $F(1,29)=0.321$, $p>0.05$ και για τον τρόπο γραφής, $F(1,29)=2.147$, $p>0.05$. Η αλληλεπίδραση είναι στατιστικώς σημαντική για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου με τον τύπο γραμματοσειράς, $F(6,24)=5.408$, $p<0.05$, για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου με τον τρόπο γραφής, $F(6,24)=4.481$, $p<0.05$ και μη-σημαντική για τον τύπο γραμματοσειράς με τον τρόπο γραφής, $F(1,29)=1.416$, $p>0.05$. Η αλληλεπίδραση μεταξύ των τριών μεταβλητών

(χρωματικοί συνδυασμοί γραμματοσειράς/υποβάθρου με τον τύπο γραμματοσειράς και με τον τρόπο γραφής) είναι επίσης στατιστικώς σημαντική, $F(6,24)=6.051$, $p<0.05$.

3. «Επικράτηση»: είναι στατιστικώς σημαντικά για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου, $F(6,24)=4.945$, $p<0.05$, για τον τρόπο γραφής, $F(1,29)=5.481$, $p<0.05$, και μη-σημαντικά για τον τύπο γραμματοσειράς, $F(1,29)=2.698$, $p>0.05$. Η αλληλεπίδραση είναι στατιστικώς σημαντική για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου με τον τύπο γραμματοσειράς, $F(6,24)=5.559$, $p<0.05$, για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου με τον τρόπο γραφής, $F(6,24)=2.929$, $p<0.05$ και μη-σημαντική για το τύπο γραμματοσειράς με τον τρόπο γραφής, $F(1,29)=0.334$, $p>0.05$. Η αλληλεπίδραση μεταξύ των τριών μεταβλητών (χρωματικοί συνδυασμοί γραμματοσειράς/υποβάθρου με τον τύπο γραμματοσειράς και με τον τρόπο γραφής) είναι επίσης στατιστικώς σημαντική, $F(6,24)=2.917$, $p<0.05$.

Two-Way ANOVA

Διενεργήθηκε 2-way r-m ANOVA για κάθε μία από τις εξαρτημένες μεταβλητές, χρησιμοποιώντας δύο ανεξάρτητες μεταβλητές, το χρωματικό συνδυασμό γραμματοσειράς/υποβάθρου (7 επίπεδα χρωματικών συνδυασμών όπως αναφέρθηκε παραπάνω) και τον τρόπο γραφής (3 επίπεδα – απλό κείμενο, έντονη και έντονη-πλάγια γραφή για «Times New Roman» και 2 επίπεδα – έντονη και έντονη/πλάγια γραφή για «Arial»). Τα αποτελέσματα παρουσιάζονται λεπτομερώς στον Πίνακα 5.

Συνοψίζοντας προκύπτει για τα αποτελέσματα ότι:

«Ευχαρίστηση»

1. «Times New Roman»: στατιστικώς σημαντικά για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου, $F(6,24)=5.552$, $p<0.05$, και μη-σημαντικά για τον τρόπο γραφής, $F(2,28)=0.265$, $p>0.05$. Η αλληλεπίδραση είναι στατιστικώς σημαντική για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου με τον τρόπο γραφής, $F(12,18)=8.749$, $p<0.05$.
2. «Arial»: στατιστικώς σημαντικά για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου, $F(6,24)=9.196$, $p<0.05$, και μη-σημαντικά για τον τρόπο γραφής, $F(1,29)=2.312$, $p>0.05$. Η αλληλεπίδραση είναι στατιστικώς

σημαντική για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου με τον τρόπο γραφής, $F(6,24)=2.652$, $p<0.05$.

«Διέγερση»

1. «*Times New Roman*»: στατιστικώς σημαντικά για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου, $F(6,24)=2.795$, $p<0.05$, και μη-σημαντικά για τον τρόπο γραφής, $F(2,28)=1.447$, $p>0.05$. Η αλληλεπίδραση είναι στατιστικώς σημαντική για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου με τον τρόπο γραφής, $F(12,18)=4.397$, $p<0.05$.
2. «*Arial*»: στατιστικώς σημαντικά για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου, $F(6,24)=5.107$, $p<0.05$, και μη-σημαντικά για τον τρόπο γραφής, $F(1,29)=3.186$, $p>0.05$. Η αλληλεπίδραση είναι στατιστικώς σημαντική για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου με τον τρόπο γραφής, $F(6,24)=2.953$, $p<0.05$.

«Επικράτηση»

1. «*Times New Roman*»: στατιστικώς σημαντικά για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου, $F(6,24)=3.55$, $p<0.05$, και για τον τρόπο γραφής, $F(2,28)=4.043$, $p<0.05$. Η αλληλεπίδραση είναι στατιστικώς σημαντική για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου με τον τρόπο γραφής, $F(12,18)=6.11$, $p<0.05$.
2. «*Arial*»: στατιστικώς μη-σημαντικά για τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου, $F(6,24)=2.323$, $p>0.05$, και για τον τρόπο γραφής, $F(1,29)=1.934$, $p>0.05$. Η αλληλεπίδραση είναι στατιστικώς μη-σημαντική για χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου με τον τρόπο γραφής, $F(6,24)=1.818$, $p>0.05$.

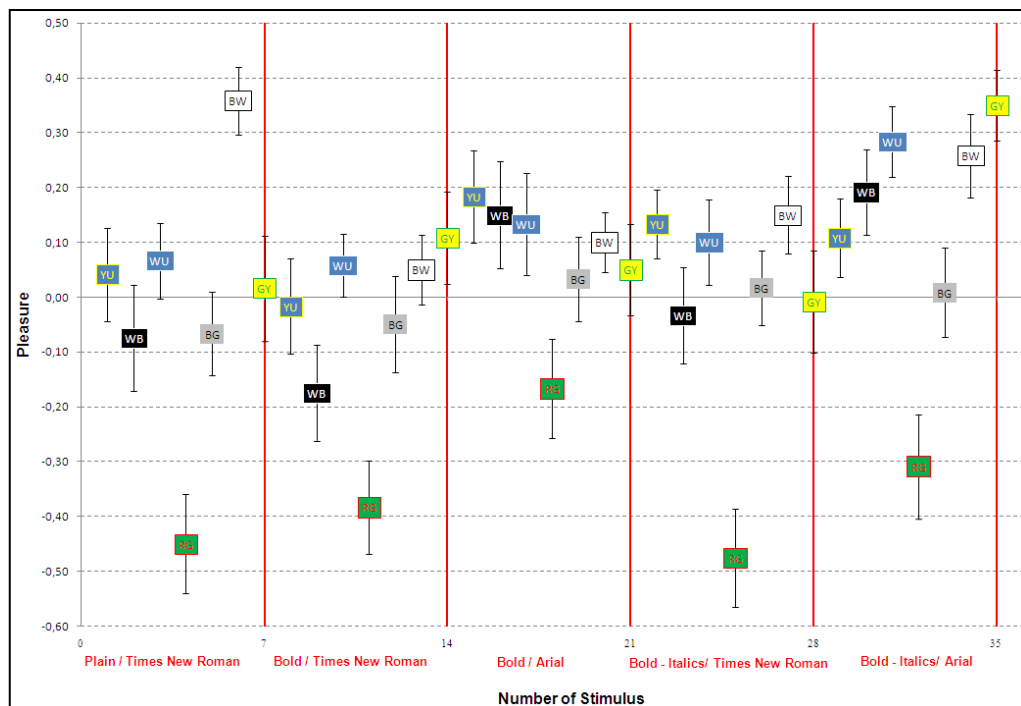
Γενικά, παρατηρούμε ότι και οι τρεις συναισθηματικές διαστάσεις επηρεάζονται από τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου.

Έλεγχος και Σύγκριση Μέσων Τιμών για τους χρωματικούς συνδυασμούς

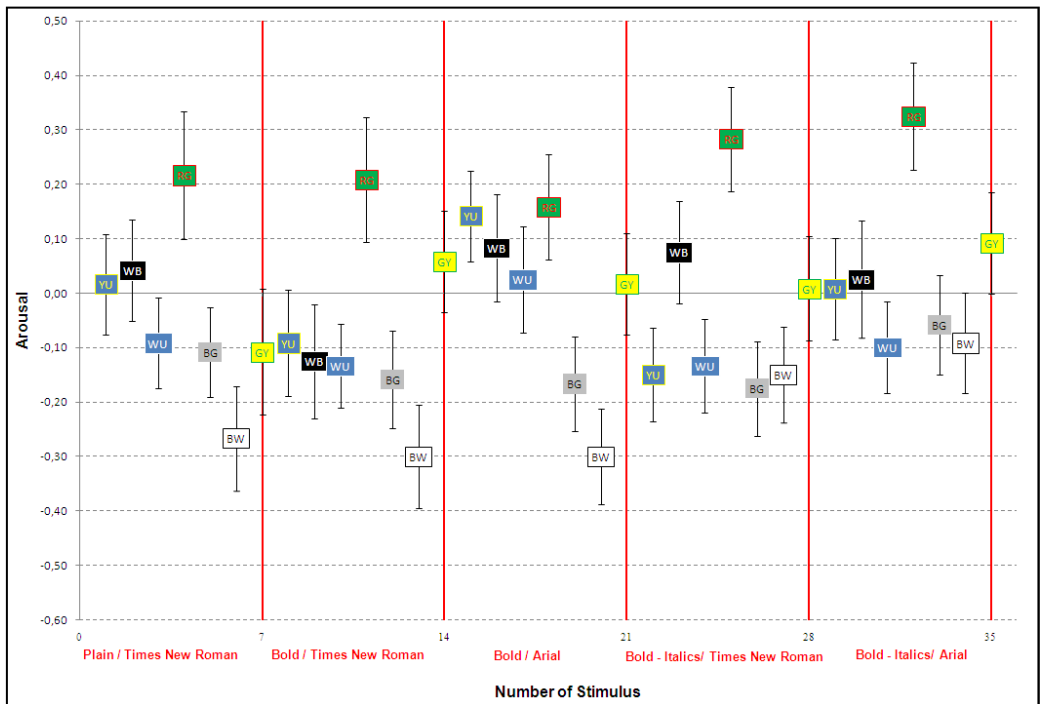
Για κάθε ανάλυση της διακύμανσης προέκυψαν και τα αντίστοιχα αποτελέσματα των στατιστικώς σημαντικών διαφορών για τις μέσες τιμές των χρωματικών συνδυασμών (pairwise comparison, $p<.05$). Παρατηρήθηκε γενικά, ότι ο χρωματικός συνδυασμός R/G έχει στατιστικώς σημαντικές διαφορές από του υπολοίπους συνδυασμούς (και για τις τρεις διαστάσεις).

Μερικά σημαντικά αποτελέσματα για την:

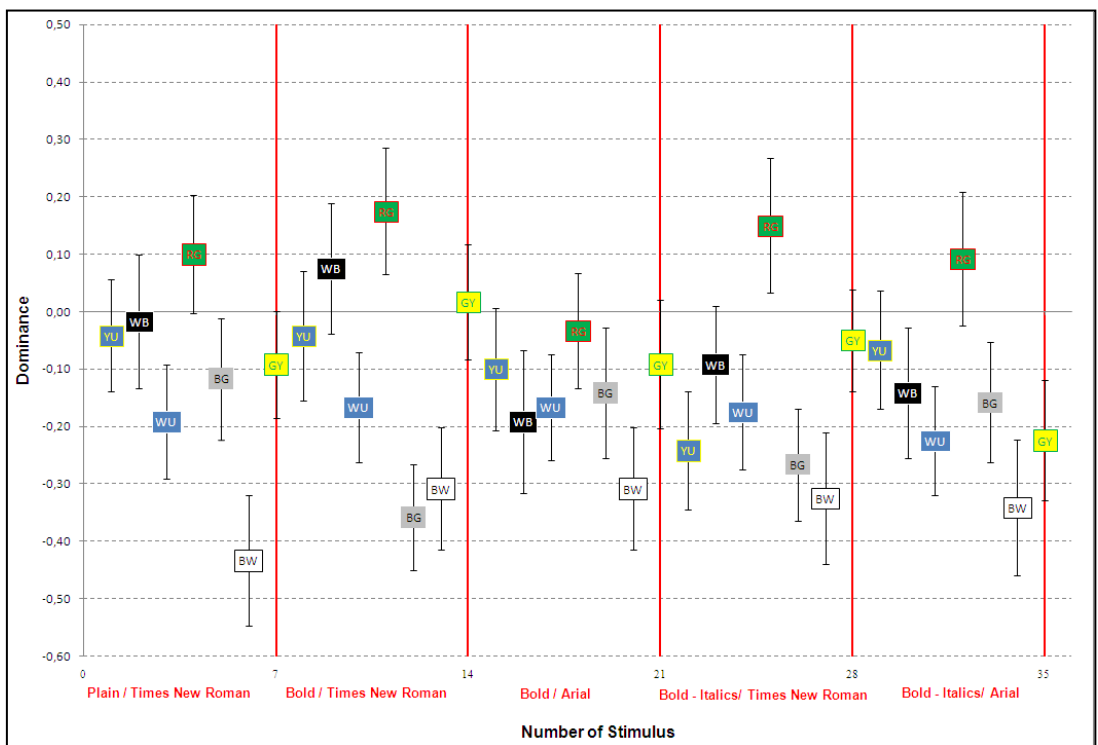
- «Ευχαρίστηση», εμφανίστηκαν στατιστικώς σημαντικές διαφορές για τον χρωματικό συνδυασμό B/W από τους W/B και B/G (με την χρήση «Times New Roman»).
- «Διέγερση», πέρα του συνδυασμού R/G, εμφανίστηκαν διαφορές για την B/W από τους W/B (με την χρήση «Times New Roman» και «Arial»), G/Y και Y/U (με την χρήση «Arial»). Επίσης για την περίπτωση της «Arial», σημαντικές διαφορές παρατηρήθηκαν για τον Y/U από τον W/U.
- «Επικράτηση», παρατηρούνται διαφορές για τον συνδυασμό B/W από τους G/Y, W/B, Y/U (με την χρήση «Times New Roman»).



(α)



(β)



(γ)

Σχήμα 18: Οι μέσες τιμές των απαντήσεων και τα αντίστοιχα τυπικά σφάλματα για όλα τα ερεθίσματα στους χρωματικούς συνδυασμούς για τον άξονα α) «Ευχαρίστηση» β) «Διέγερση» και γ) «Επικράτηση»

Πίνακας 4: Τα αποτελέσματα της three-way r-m ANOVA για κάθε μία εξαρτημένη μεταβλητή (συναισθηματικές διαστάσεις) ως προς τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου, τύπου γραμματοσειράς και τρόπου γραφής μαζί με τις αλληλεπιδράσεις.

ΕΥΧΑΡΙΣΤΗΣΗ	F	p	MSE†
Χρώμα γραμματοσειράς/υποβάθρου	$F(6,24)=7.668$	$p<0.05^*$	0.155
Τύπος Γραμματοσειράς	$F(1,29)=22.146$	$p<0.05^*$	0.111
Τρόπος Γραφής	$F(1,29)= 0.680$	$p>0.05^{**}$	0.126

Αλληλεπίδραση

Χρώμα γραμματοσειράς/υποβάθρου * Τύπος Γραμματοσειράς	$F(6,24)=7.406$	$p<0.05^*$	0.281
Χρώμα γραμματοσειράς/υποβάθρου * Τρόπος Γραφής	$F(6,24)=6.436$	$p<0.05^*$	0.191
Τύπος Γραμματοσειράς * Τρόπος Γραφής	$F(1,29)=0.785$	$p>0.05^{**}$	0.085
Χρώμα γραμματοσειράς/υποβάθρου * Τύπος Γραμματοσειράς * Τρόπος Γραφής	$F(6,24)=5.213$	$p<0.05^*$	0.252

ΔΙΕΓΕΡΣΗ	F	p	MSE†
Χρώμα γραμματοσειράς/υποβάθρου	$F(6,24)= 1.950$	$p>0.05^{**}$	0.200
Τύπος Γραμματοσειράς	$F(1,29)=0.321$	$p>0.05^{**}$	0.112
Τρόπος Γραφής	$F(1,29)= 2.147$	$p>0.05^{**}$	0.151

Αλληλεπίδραση

Χρώμα γραμματοσειράς/υποβάθρου * Τύπος Γραμματοσειράς	$F(6,24)=5.408$	$p<0.05^*$	0.290
Χρώμα γραμματοσειράς/υποβάθρου * Τρόπος Γραφής	$F(6,24)=4.481$	$p<0.05^*$	0.404
Τύπος Γραμματοσειράς * Τρόπος Γραφής	$F(1,29)=1.416$	$p>0.05^{**}$	0.076
Χρώμα γραμματοσειράς/υποβάθρου * Τύπος Γραμματοσειράς * Τρόπος Γραφής	$F(6,24)=6.051$	$p<0.05^*$	0.212

ΕΠΙΚΡΑΤΗΣΗ	F	p	MSE†
Χρώμα γραμματοσειράς/υποβάθρου	$F(6,24)=4.945$	$p<0.05^*$	0.129
Τύπος Γραμματοσειράς	$F(1,29)= 2.698$	$p>0.05^{**}$	0.096
Τρόπος Γραφής	$F(1,29)= 5.481$	$p<0.05^*$	0.090

Αλληλεπίδραση

Χρώμα γραμματοσειράς/υποβάθρου * Τύπος Γραμματοσειράς	$F(6,24)=5.559$	$p<0.05^*$	0.287
Χρώμα γραμματοσειράς/υποβάθρου * Τρόπος Γραφής	$F(6,24)=2.929$	$p<0.05^*$	0.213
Τύπος Γραμματοσειράς * Τρόπος Γραφής	$F(1,29)=0.334$	$p>0.05^{**}$	0.112
Χρώμα γραμματοσειράς/υποβάθρου * Τύπος Γραμματοσειράς * Τρόπος Γραφής	$F(6,24)=2.917$	$p<0.05^*$	0.304

* στατιστικώς σημαντικό, ** στατιστικώς μη-σημαντικό † Mean Square Error

Πίνακας 5: Τα αποτελέσματα της two-way r-m ANOVA για κάθε μία εξαρτημένη μεταβλητή (συναισθηματικές διαστάσεις) ως προς τους χρωματικούς συνδυασμούς γραμματοσειράς/υποβάθρου και τον τρόπο γραφής, την αλληλεπίδραση τους για «Times New Roman» και «Arial».

ΕΥΧΑΡΙΣΤΗΣΗ		F	p	MSE†
Times New Roman				
Χρώμα γραμματοσειράς/υποβάθρου		F(6,24)=5.552	p<0.05*	0.170
Τρόπος Γραφής		F(2,28)=0.265	p>0.05**	0.118
Αλληλεπίδραση				
Χρώμα γραμματοσειράς/υποβάθρου * Τρόπος Γραφής		F(12,18)=8.749	p<0.05*	0.322
Arial				
Χρώμα γραμματοσειράς/υποβάθρου		F(6,24)=9.196	p<0.05*	0.216
Τρόπος Γραφής		F(1,29)=2.312	p>0.05**	0.079
Αλληλεπίδραση				
Χρώμα γραμματοσειράς/υποβάθρου * Τρόπος Γραφής		F(12,18)=2.652	p<0.05*	0.240
ΔΙΕΓΕΡΣΗ		F	P	MSE†
Times New Roman				
Χρώμα γραμματοσειράς/υποβάθρου		F(6,24)=2.795	p<0.05*	0.186
Τρόπος Γραφής		F(2,28)=1.447	p>0.05**	0.097
Αλληλεπίδραση				
Χρώμα γραμματοσειράς/υποβάθρου * Τρόπος Γραφής		F(12,18)=4.397	p<0.05*	0.348
Arial				
Χρώμα γραμματοσειράς/υποβάθρου		F(6,24)=5.107	p<0.05*	0.227
Τρόπος Γραφής		F(1,29)=3.186	p>0.05**	0.090
Αλληλεπίδραση				
Χρώμα γραμματοσειράς/υποβάθρου * Τρόπος Γραφής		F(6,24)=2.953	p<0.05*	0.283
ΕΠΙΚΡΑΤΗΣΗ		F	P	MSE†
Times New Roman				
Χρώμα γραμματοσειράς/υποβάθρου		F(6,24)=3.550	p<0.05*	0.149
Τρόπος Γραφής		F(2,28)=4.043	p<0.05*	0.089
Αλληλεπίδραση				
Χρώμα γραμματοσειράς/υποβάθρου * Τρόπος Γραφής		F(12,18)=6.110	p<0.05*	0.325
Arial				
Χρώμα γραμματοσειράς/υποβάθρου		F(6,24)=2.323	p>0.05**	0.252
Τρόπος Γραφής		F(1,29)=1.934	p>0.05**	0.077
Αλληλεπίδραση				
Χρώμα γραμματοσειράς/υποβάθρου * Τρόπος Γραφής		F(6,24)=1.818	p>0.05**	0.214

* στατιστικώς σημαντικό, ** στατιστικώς μη-σημαντικό † Mean Square Error

Υπόθεση 2 (H₂)

Διενεργήθηκε 1-way r-m ANOVA για κάθε μία από τις εξαρτημένες μεταβλητές, ώστε να προσδιοριστεί εάν υπάρχουν στατιστικώς σημαντικές διαφορές στις μέσες τιμές της βαθμολογίας των συναισθηματικών καταστάσεων, καθώς αλλάζουμε το μέγεθος της γραμματοσειράς (10 επίπεδα).

1. «Ευχαρίστηση»: Τα αποτελέσματα ήταν στατιστικώς σημαντικά, $F(9,21) = 21.382$, $p < 0.05$ (MSE=0.328)
2. «Διέγερση»: Τα αποτελέσματα ήταν στατιστικώς σημαντικά, $F(9,21) = 3.359$, $p < 0.05$ (MSE=0.398)
3. «Επικράτηση»: Τα αποτελέσματα ήταν στατιστικώς σημαντικά, $F(9,21) = 6.092$, $p < 0.05$ (MSE=0.354)

Επομένως, το μέγεθος της γραμματοσειράς επηρεάζει τις συναισθηματικές διαστάσεις και η υπόθεση 2 επιβεβαιώνεται για την «Ευχαρίστηση» (H_{2α}), «Διέγερση» (H_{2β}) και «Επικράτηση» (H_{2γ}).

Χρησιμοποιώντας το Tukey test, παρουσιάστηκαν σημαντικές διαφορές μεταξύ των μεγεθών γραμματοσειράς, για παράδειγμα, για μέγεθος γραμματοσειράς 10px και 14px, οι μέσες τιμές για τη διάσταση «Ευχαρίστηση» είναι στατιστικώς σημαντικά διαφορετικές ($p=0.000121 < 0.05$).

Στον Πίνακα 6 παρουσιάζονται οι μέσες τιμές των απαντήσεων για κάθε μέγεθος γραμματοσειράς μαζί με τα αντίστοιχα τυπικά σφάλματα. Παρατηρούμε ότι καθώς μεταβάλλεται το μέγεθος της γραμματοσειράς, η διάσταση «Ευχαρίστηση» αυξάνεται μέχρι το μέγεθος 26px. Για τιμές μεγαλύτερες του μεγέθους 26px, υπάρχει μία τάση μείωσης. Οι τιμές της διάστασης «Διέγερση» μειώνονται μέχρι το μέγεθος γραμματοσειράς 15px και της διάστασης «Επικράτηση» μέχρι 18px. αντίστοιχα. Πέρα από αυτές τις τιμές υπάρχει μία αυξητική τάση.

Υποθέσεις 3 και 4 (H₃ και H₄)

Διενεργήθηκε 2-way repeated measures ANOVA για να προσδιοριστεί εάν οι μεταβολές των συναισθηματικών καταστάσεων είναι στατιστικώς σημαντικές, καθώς μεταβάλουμε τον τύπο της γραμματοσειράς («Arial» και «Times New Roman» - 2 επίπεδα), και τον τρόπο γραφής (απλό κείμενο, έντονη, πλάγια, έντονη-πλάγια γραφή - 4 επίπεδα), με σταθερό το χρωματικό συνδυασμό μαύρο χρώμα γραμματοσειράς σε άσπρο υπόβαθρο και μέγεθος γραμματοσειράς 16px. Τα αποτελέσματα παρουσιάζονται στον Πίνακα 7.

Πίνακας 6: Οι μέσες τιμές των τριών συναισθηματικών διαστάσεων ως προς το μέγεθος της γραμματοσειράς (χρησιμοποιώντας μαύρο κείμενο, «Times New Roman» σε άσπρο υπόβαθρο) μαζί με το τυπικό σφάλμα (SE).

Μέγεθος γραμματοσειράς (px)	ΕΥΧΑΡΙΣΤΗΣΗ		ΔΙΕΓΕΡΣΗ		ΕΠΙΚΡΑΤΗΣΗ	
	Μέση τιμή	SE	Μέση τιμή	SE	Μέση τιμή	SE
10	-0.47	0.08	0.08	0.12	0.06	0.12
11	-0.30	0.07	-0.05	0.10	0.04	0.11
12	-0.27	0.07	-0.06	0.10	-0.07	0.11
13	-0.23	0.08	-0.18	0.10	-0.10	0.11
14	-0.01	0.05	-0.36	0.09	-0.21	0.11
15	0.13	0.06	-0.37	0.09	-0.40	0.12
16	0.36	0.06	-0.27	0.10	-0.43	0.11
18	0.28	0.08	-0.21	0.11	-0.53	0.11
26	0.43	0.07	0.02	0.10	-0.31	0.12
32	0.20	0.07	-0.03	0.10	-0.22	0.12
Μέσο τυπικό σφάλμα		0.07		0.10		0.11

Οι μετρούμενες μεταβολές για την υπόθεση 3 (απλό κείμενο) είναι:

1. «Arial» - Ευχαρίστηση: Μέση τιμή = 0.13 (SE=0.05).
2. «Times New Roman» - Ευχαρίστηση: Μέση τιμή = 0.36 (SE=0.06).

Η αλληλεπίδραση αποκάλυψε ότι χρησιμοποιώντας «Times New Roman» έναντι «Arial» σε απλό κείμενο αυξάνεται η ευχαρίστηση κατά 23%. Σε όλες τις άλλες περιπτώσεις («Ευχαρίστηση» χρησιμοποιώντας έντονη, πλάγια, έντονη-πλάγια γραφή, «Διέγερση» χρησιμοποιώντας απλό κείμενο, έντονη, πλάγια, έντονη-πλάγια γραφή και «Επικράτηση» χρησιμοποιώντας απλό κείμενο, έντονη, πλάγια, έντονη-πλάγια γραφή) τα αποτελέσματα είναι στατιστικώς μη σημαντικά. Επομένως, η υπόθεση H3 ισχύει μόνο για την περίπτωση απλού κειμένου για την «Ευχαρίστηση» (H3α).

Η υπόθεση 4 (H4) δεν επιβεβαιώθηκε για την περίπτωση της «Διέγερσης» (H4β) και «Επικράτησης» (H4γ). Η ANOVA αποκάλυψε την επίδραση του τρόπου γραφής στην «Ευχαρίστηση» (H4α).

Τα παραπάνω αποτελέσματα μας αναγκάζουν να ξεχωρίσουμε τις μεταβολές του τρόπου γραφής, καθώς χρησιμοποιούμε «Arial» και «Times New Roman», όπως φαίνεται παρακάτω (Πίνακες 8α και 8β αντίστοιχα):

Πίνακας 7: Τα αποτελέσματα της two-way r-m ANOVA για κάθε μία εξαρτημένη μεταβλητή (συναισθηματικές διαστάσεις) ως προς τις εναλλαγές του τύπου γραμματοσειράς και του τρόπου γραφής καθώς και η αλληλεπίδραση αυτών των μεταβολών.

ΕΥΧΑΡΙΣΤΗΣΗ	F	p	MSE†
Τρόπος γραφής	$F(3,27)=4.033$	$p<0.05^*$	0.106
Τύπος γραμματοσειράς	$F(1,29)=0.477$	$p>0.05^{**}$	0.066
Αλληλεπίδραση			
Τρόπος γραφής * Τύπος γραμματοσειράς	$F(3,27)=4.150$	$p<0.05^*$	0.129
ΔΙΕΓΕΡΣΗ	F	p	MSE†
Τρόπος γραφής	$F(3,27)=2.970$	$p>0.05^{**}$	0.154
Τύπος γραμματοσειράς	$F(1,29)=0.080$	$p>0.05^{**}$	0.209
Αλληλεπίδραση			
Τρόπος γραφής * Τύπος γραμματοσειράς	$F(3,27)=0.303$	$p>0.05^{**}$	0.092
ΕΠΙΚΡΑΤΗΣΗ	F	p	MSE†
Τρόπος γραφής	$F(3,27)=1.431$	$p>0.05^{**}$	0.100
Τύπος γραμματοσειράς	$F(1,29)=0.707$	$p>0.05^{**}$	0.083
Αλληλεπίδραση			
Τρόπος γραφής * Τύπος γραμματοσειράς	$F(3,27)=0.281$	$p>0.05^{**}$	0.120

* στατιστικώς σημαντικό, ** στατιστικώς μη-σημαντικό † Mean Square Error

Πίνακας 8: Οι μέσες τιμές της «Ευχαρίστησης» και οι ποσοστιαίες μεταβολές από το απλό κείμενο, για τον τρόπο γραφής χρησιμοποιώντας α) «Arial» και β) «Times New Roman»

(α)

Arial			
	Μέση Τιμή	% μεταβολή από το απλό κείμενο	SE
Απλό κείμενο	0.13	—	0.05
Έντονη γραφή	0.10	-3	0.05
Πλάγια γραφή	0.33	20	0.07
Έντονη - Πλάγια Γραφή	0.26	13	0.08

(β)

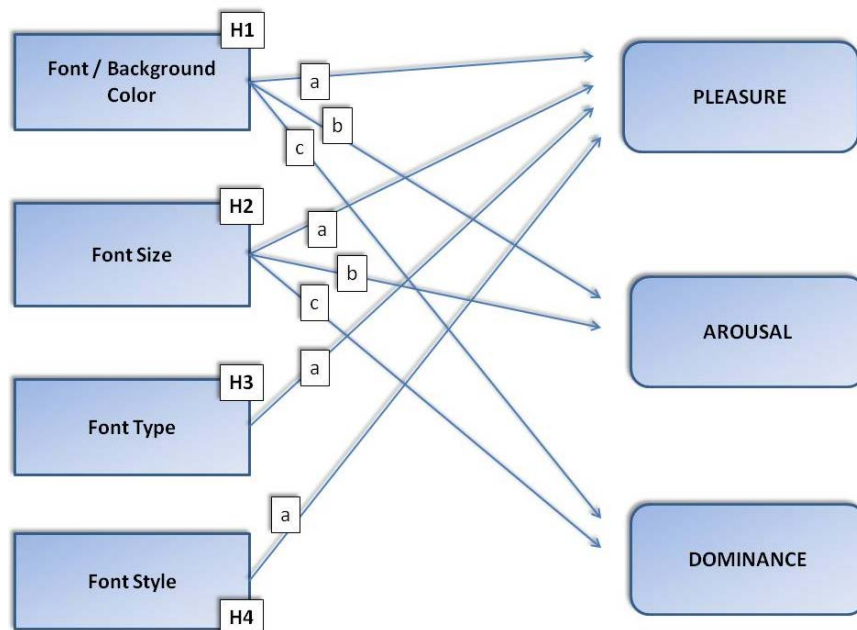
Times New Roman			
	Μέση Τιμή	% μεταβολή από το απλό κείμενο	SE
Απλό κείμενο	0.36	—	0.06
Έντονη γραφή	0.05	-31	0.06
Πλάγια γραφή	0.17	-19	0.06
Έντονη - Πλάγια Γραφή	0.15	-21	0.07

«**Arial**»: Θεωρώντας το απλό κείμενο ως το σημείο αναφοράς (baseline) μπορούμε να παρατηρήσουμε ότι η χρήση της έντονης γραφής μειώνει οριακά την «Ευχαρίστηση» κατά 3%, ενώ η πλάγια και έντονη-πλάγια γραφή την αυξάνει κατά 20% και 13% αντίστοιχα. Επομένως, μεταβάλλοντας τον τρόπο γραφής αυξάνεται η συναισθηματική διάσταση «Ευχαρίστηση» (εκτός από την περίπτωση έντονη γραφή).

«**Times New Roman**»: Σε αυτή την περίπτωση, όλες οι εναλλαγές τρόπου γραφής μειώνουν τη συναισθηματική διάσταση «Ευχαρίστηση»: έντονη γραφή κατά 31%, πλάγια γραφή κατά 19% και πλάγια-έντονη γραφή κατά 21%.

3.3.5 Μαθηματική μοντελοποίηση των αποτελεσμάτων

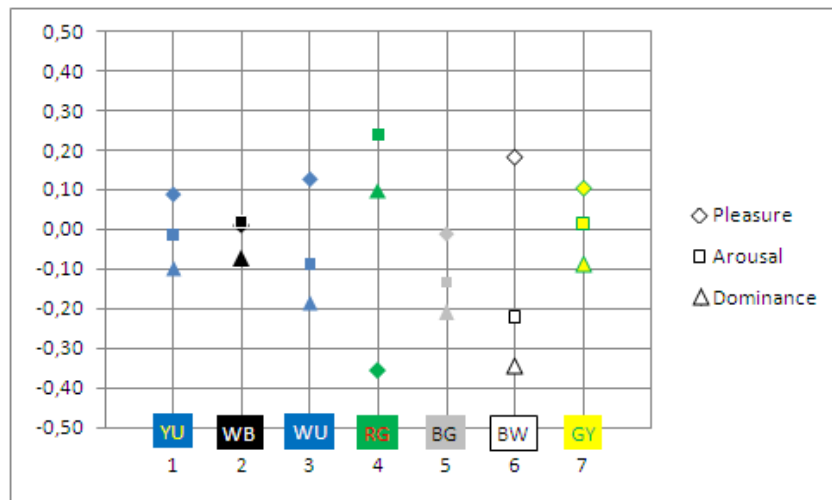
Χρησιμοποιώντας τα προηγούμενα αποτελέσματα, μπορεί να προταθεί ένα γενικό μοντέλο (Σχήμα 19) για τη συσχέτιση των ανεξάρτητων μεταβλητών με τις τρεις συναισθηματικές διαστάσεις «Ευχαρίστηση», «Διέγερση» και «Επικράτηση».



Σχήμα 19: Το προτεινόμενο μοντέλο συσχέτισμού των ανεξάρτητων μεταβλητών με τις τρεις διαστάσεις «Ευχαρίστηση», «Διέγερση» και «Επικράτηση».

Υπόθεση 1 (H₁).

Το διάγραμμα διασποράς των συναισθηματικών διαστάσεων «Ευχαρίστηση», «Διέγερση» και «Επικράτηση» για κάθε χρωματικό συνδυασμό, παρουσιάζεται στο Σχήμα 20, που διευκολύνει τη σύγκριση των μέσων τιμών και τη συζήτηση των αποτελεσμάτων.



Σχήμα 20: Οι μέσες τιμές των τριών συναισθηματικών διαστάσεων «Ευχαρίστηση», «Διέγερση» και «Επικράτηση» για κάθε χρωματικό συνδυασμό. Η κωδικοποίηση του άξονα X είναι, 1. Κίτρινο σε Μπλε, 2. Άσπρο σε Μαύρο, 3. Άσπρο σε Μπλε, 4. Κόκκινο σε Πράσινο, 5 Μαύρο σε Γκρι, 6. Μαύρο σε Άσπρο και 7. Πράσινο σε Κίτρινο.

Μερικά σημαντικά ευρήματα είναι τα εξής:

1. Δεν παρατηρήθηκαν ακραίες συναισθηματικές καταστάσεις (emotional extremes).
2. Οι χρωματικοί συνδυασμοί που αξιολογήθηκαν ως οι πιο ευχάριστοι είναι : μαύρο κείμενο σε άσπρο υπόβαθρο, άσπρο σε μπλε και πράσινο σε κίτρινο.
3. Ο λιγότερο έντονος συνδυασμός θεωρήθηκε το μαύρο κείμενο σε λευκό υπόβαθρο .
4. Ο λιγότερο ευχάριστος και περισσότερο έντονος συνδυασμός θεωρήθηκε το κόκκινο κείμενο σε πράσινο υπόβαθρο .

Επιπλέον, οι Hall και Hanna [5] έδειξαν ότι οι χρωματικοί συνδυασμοί με υψηλά επίπεδα αντίθεσης, οδηγούν σε βελτίωση της αναγνωσιμότητας και της αισθητικής. Οι χρωματικοί συνδυασμοί υψηλής αντίθεσης (μαύρο σε άσπρο, άσπρο σε μαύρο, κίτρινο σε μπλε και άσπρο σε μπλε) οδηγούν σε υψηλότερες τιμές για τη διάσταση

«Ευχαρίστηση» (H1α) και χαμηλότερες τιμές για τις «Διέγερση» και «Επικράτηση» (H1β και H1γ, αντίστοιχα). Υπάρχουν, όμως και μερικές αποκλίσεις από τις αναμενόμενες τιμές.

1. Άσπρο σε μαύρο (χρωματικός συνδυασμός υψηλής αντίθεσης) δίνει υψηλότερες τιμές για τη «Διέγερση».
2. Μαύρο σε Γκρι (χρωματικός συνδυασμός χαμηλής αντίθεσης) δίνει χαμηλότερες τιμές για τις διαστάσεις «Διέγερση» και «Επικράτηση».
3. Πράσινο σε κίτρινο (χρωματικός συνδυασμός χαμηλής αντίθεσης) δίνει υψηλότερες τιμές στη διάσταση «Ευχαρίστηση» και χαμηλότερες τιμές για τη διάσταση «Επικράτηση». Όμως, αυτά τα αποτελέσματα είναι σε συμφωνία με εκείνα που παρουσιάζει ο Laarni [16].

Τα αποτελέσματα είναι σύμφωνα με εκείνα που παρουσιάζονται στην εργασία [16]: (α) οι χρωματικοί συνδυασμοί γραμματοσειράς/υποβάθρου επηρεάζουν και τις τρεις συναισθηματικές διαστάσεις. (β) ο χρωματικός συνδυασμός κόκκινο σε πράσινο θεωρείται ως ο πιο δυσάρεστος και προκαλεί υψηλή διέγερση. (γ) ο χρωματικός συνδυασμός μαύρο σε άσπρο έχει το χαμηλότερο μέσο όρο στη διάσταση «Διέγερση». (δ) οι χρωματικοί συνδυασμοί άσπρο σε μπλε και πράσινο σε κίτρινο είναι οι πιο ευχάριστοι.

Τα αποτελέσματα αυτής της εργασίας και της [16] προέκυψαν μέσα από τη δοκιμασία SAM, αλλά βασίζονται σε συμμετέχοντες με γλωσσικές και πολιτισμικές διαφορές υποδεικνύοντας ότι η επίδραση των χρωματικών συνδυασμών γραμματοσειράς/υποβάθρου που χρησιμοποιούνται στην τυπογραφία, στις συναισθηματικές διαστάσεις, είναι πιθανώς *ανεξάρτητη γλώσσας και πολιτισμού*. Αυτό το συμπέρασμα πρέπει να μελετηθεί περαιτέρω (χρησιμοποιώντας την ίδια ή παρόμοια μεθοδολογία) μεταξύ και άλλων μη-Δυτικών (non-Western) γλωσσών και πολιτισμών. Υπάρχουν ενδείξεις για συνέπεια των αποτελεσμάτων με αυτά στην Αραβική γλώσσα [185] [187]. Επίσης, οι Cyr et al. [188] συμπέραναν ότι υπάρχουν μερικές ομοιότητες αντίδρασης στα χρωματικά σχήματα (color scheme) ιστοσελίδων σε διαφορετικούς πολιτισμούς (Καναδικός, Γερμανικός και Ιαπωνικός), αλλά παρατηρούνται και μερικές διαφοροποιήσεις. Για παράδειγμα, στους τρεις πολιτισμούς τείνει να μην τους αρέσουν ιστοσελίδες με κίτρινο χρωματικό σχήμα. Οι Γερμανοί έδειξαν προτίμηση για το μπλε και οι Καναδοί εκτιμούσαν περισσότερο το γκρι από ότι οι Γερμανοί και οι Ιάπωνες.

Υπόθεση 2 (H₂).

Βασιζόμενοι στα αποτελέσματα που παρουσιάζονται στον Πίνακα 6, μπορούμε να μοντελοποιήσουμε τη συσχέτιση των συναισθηματικών διαστάσεων με το μέγεθος της γραμματοσειράς ακολουθώντας την παρακάτω σχέση:

$$f_{\{d\}}(s) = B_3^{\{d\}} \cdot s^3 + B_2^{\{d\}} \cdot s^2 + B_1^{\{d\}} \cdot s + Intercept^{\{d\}} \quad (14)$$

όπου:

d : P, A ή D

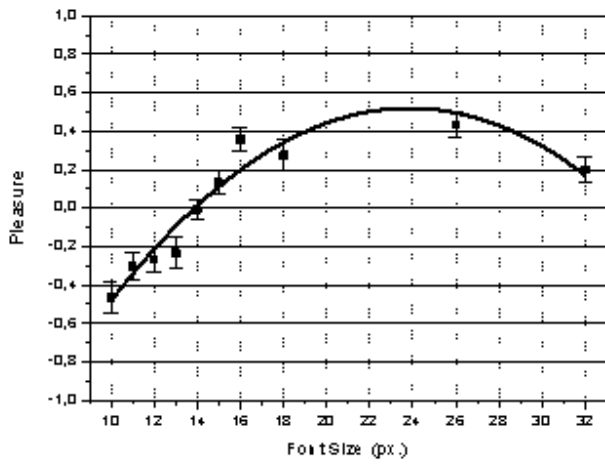
P, A και D : «Ευχαρίστηση», «Διέγερση» και «Επικράτηση» αντίστοιχα με πεδίο ορισμού $[-1,1]$

B_1, B_2 και B_3 είναι οι πολυωνυμικοί συντελεστές

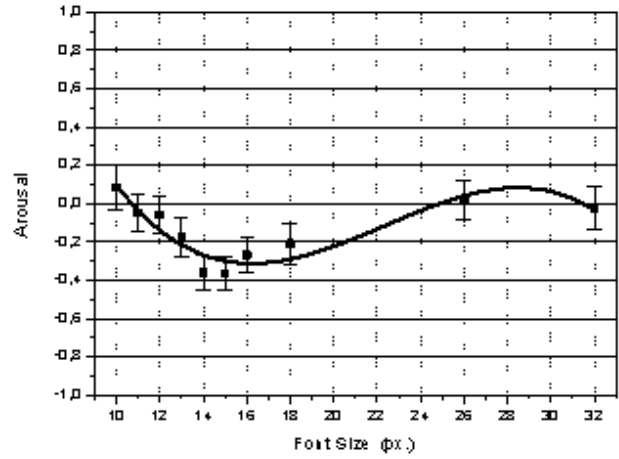
s : το μέγεθος γραμματοσειράς σε px

Ο Πίνακας 9 παρουσιάζει τους πολυωνυμικούς συντελεστές που προέκυψαν με τη χρήση του στατιστικού πακέτου Origin 8.0 [189] για την εξίσωση 14. Τα αποτελέσματα δείχνουν ότι η διάσταση «Ευχαρίστηση» έναντι του μεγέθους γραμματοσειράς, μπορεί να περιγραφεί χρησιμοποιώντας μία δευτεροβάθμια εξίσωση και οι διαστάσεις «Διέγερση» και «Επικράτηση» περιγράφονται από εξισώσεις τρίτου βαθμού. Ο διορθωμένος συντελεστής προσδιορισμού R^2 (adjusted R-square) είναι το μέτρο της ορθότητας του μοντέλου.

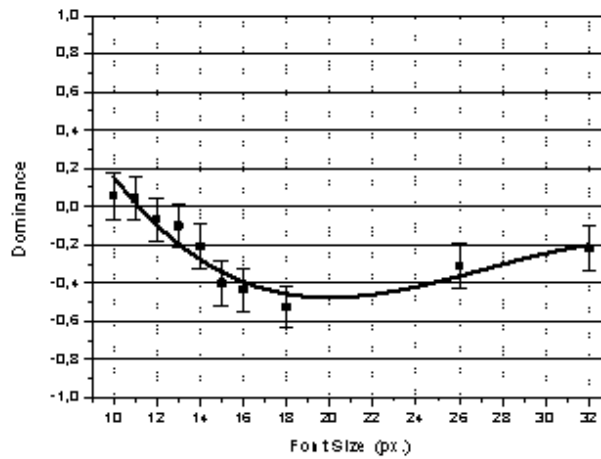
Το μαθηματικό μοντέλο παρουσιάζεται γραφικά στα Σχήματα 21α, 21β και 21γ για την «Ευχαρίστηση», «Διέγερση» και «Επικράτηση», αντίστοιχα ως προς το μέγεθος της γραμματοσειράς. Τα παραπάνω αποτελέσματα είναι σε συμφωνία με εκείνα των Bernardini et al. [190], που παρατήρησαν μείωση στη «συναισθηματική προτίμηση» όχι μόνο χρησιμοποιώντας μικρότερο μέγεθος γραμματοσειράς, αλλά και μεγαλύτερο.



(α)



(β)



(γ)

Σχήμα 21: Η γραφική παράσταση της διάστασης (α) «Ευχαρίστηση», (β) «Διέγερση» και (γ) «Επικράτηση» ως προς το μέγεθος της γραμματοσειράς.

Πίνακας 9: Οι πολυωνυμικοί συντελεστές της μαθηματικής περιγραφής του συσχετισμού των συναισθηματικών διαστάσεων με το μέγεθος γραμματοσειράς και το τυπικό σφάλμα (SE).

	Ευχαρίστηση		Διέγερση		Επικράτηση	
	SE	SE	SE	SE	SE	SE
Intercept	-2.45287	0.30771	3.58459	0.73588	2.85307	0.77138
B₁	0.24942	0.03299	-0.59264	0.12133	-0.41265	0.12828
B₂	-0.00523	0.00077	0.02866	0.00632	0.01631	0.0067
B₃	0	0	-0.000427111	0.0001029	-0.000199863	0.0001093
Adjusted R-square	0.90687		0.78103		0.84705	

Υποθέσεις 3 και 4 (H₃ και H₄).

Τα αποτελέσματα της παρούσας μελέτης καταδεικνύουν ότι μόνο η συναισθηματική διάσταση «Ευχαρίστηση» επηρεάζεται από τον τρόπο γραφής και τον τύπο της γραμματοσειράς. Πρέπει να λάβουμε υπόψη το γεγονός ότι οι διακυμάνσεις των συναισθηματικών διαστάσεων δεν είναι ακραίες κατά τις εναλλαγές του τρόπου γραφής και του τύπου της γραμματοσειράς.

Μοντέλο συνεχών - διακριτών τιμών.

Οι τιμές των τυπογραφικών στοιχείων μπορεί να είναι είτε *διακριτές* (π.χ. ο τρόπος γραφής) είτε *συνεχείς* (π.χ. το μέγεθος της γραμματοσειράς). Εάν οι διαστάσεις «Ευχαρίστηση», «Διέγερση» και «Επικράτηση» χρησιμοποιούνται ως ανεξάρτητες μεταβλητές, η αντιστοίχιση των τυπογραφικών στοιχείων σε αυτές, μπορεί να μοντελοποιηθεί χρησιμοποιώντας γραμμικές εξισώσεις (π.χ. $P = a_1^P \cdot x + b_P$) ή πολυωνυμικές εξισώσεις μίας μεταβλητής της μορφής:

$$f_{\{d\}}(x) = a_n^{\{d\}} \cdot x^n + \dots + a_3^{\{d\}} \cdot x^3 + a_2^{\{d\}} \cdot x^2 + a_1^{\{d\}} \cdot x + b_{\{d\}} \quad (15)$$

ή πολυωνυμικές εξισώσεις πολλαπλών μεταβλητών της μορφής:

$$f_{\{d\}}(x) = a_n^{\{d\}} \cdot x_n + \dots + a_3^{\{d\}} \cdot x_3 + a_2^{\{d\}} \cdot x_2 + a_1^{\{d\}} \cdot x_1 + b_{\{d\}} \quad (16)$$

Οι εξισώσεις μπορούν να περιγραφούν υπό την μορφή πινάκων:

$$E = A \cdot X + B \quad (17)$$

όπου:

$$E = \begin{bmatrix} P \\ A \\ D \end{bmatrix}, \quad X = \begin{bmatrix} x \\ x^2 \\ x^3 \\ \dots \\ x^n \end{bmatrix} \quad \text{ή} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{bmatrix}$$

$$A = \begin{bmatrix} a_1^P & a_2^P & a_3^P & \dots & a_n^P \\ a_1^A & a_2^A & a_3^A & \dots & a_n^A \\ a_1^D & a_2^D & a_3^D & \dots & a_n^D \end{bmatrix}, B = \begin{bmatrix} b_P \\ b_A \\ b_D \end{bmatrix}$$

και

$E(3 \times 1)$ ο πίνακας που περιγράφει τις τρεις συναισθηματικές διαστάσεις «Ευχαρίστηση» (P), «Διέγερση» (A) και «Επικράτηση» (D).

$X(n \times 1)$ ο πίνακας που περιγράφει την/τις ανεξάρτητη(-ες) μεταβλητή(-ες).

$A(3 \times n)$ ο πίνακας που περιγράφει τους πολυωνυμικούς συντελεστές της/των ανεξάρτητης(-ων) μεταβλητής(-ων).

$B(3 \times 1)$ ο πίνακας που περιγράφει τις σταθερές.

Οι παραπάνω εξισώσεις ισχύουν για τυπογραφικά στοιχεία συνεχών τιμών. Στην περίπτωση που τα τυπογραφικά στοιχεία είναι διακριτά, η εξίσωση (17) μετατρέπεται σε:

$$E_i = B_i \quad (18)$$

όπου:

$$E_i = \begin{bmatrix} P_i \\ A_i \\ D_i \end{bmatrix}, \quad B_i = \begin{bmatrix} b_{i,P} \\ b_{i,A} \\ b_{i,D} \end{bmatrix}$$

με $i \in Q$ όπου Q αντιστοιχεί στις τιμές των τυπογραφικών στοιχείων.

Αυτές οι δύο περιπτώσεις μπορούν να εξηγηθούν με τη χρήση των παρακάτω παραδειγμάτων.

ΠΑΡΑΔΕΙΓΜΑΤΑ

- *Παράδειγμα 1*

Ας υποθέσουμε το μοντέλο που περιγράφεται από την εξίσωση (14):

$$P = -0.00523 \cdot x^2 + 0.24942 \cdot x - 2.45287$$

$$A = -0.00042711 \cdot x^3 + 0.02866 \cdot x^2 - 0.59264 \cdot x + 3.58459 \quad (19)$$

$$D = 0.000199863 \cdot x^3 + 0.01631 \cdot x^2 - 0.41265 \cdot x + 2.85307$$

Αυτή η μοντελοποίηση μπορεί να παρουσιαστεί χρησιμοποιώντας την (17) όπου:

$$E = \begin{bmatrix} P \\ A \\ D \end{bmatrix}, \quad X = \begin{bmatrix} x \\ x^2 \\ x^3 \end{bmatrix},$$

$$A = \begin{bmatrix} 0.24942 & -0.00523 & 0 \\ -0.59264 & 0.02866 & -0.00042711 \\ -0.41265 & 0.01631 & 0.000199863 \end{bmatrix}, \quad (20)$$

$$B = \begin{bmatrix} -2.45287 \\ 3.58459 \\ 2.85307 \end{bmatrix}$$

• *Παράδειγμα 2*

Ας υποθέσουμε ότι το μοντέλο που περιγράφει τις μεταβολές της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη από τον τρόπο γραφής χρησιμοποιώντας «Arial» είναι:

1. «Ευχαρίστηση» - απλό κείμενο: 0, έντονη γραφή: -0.03, πλάγια γραφή: 0.2, έντονη-πλάγια γραφή: 0.13;
2. «Διέγερση» - απλό κείμενο: 0, έντονη γραφή: 0, πλάγια γραφή: 0, έντονη-πλάγια γραφή: 0;
3. «Επικράτηση» - απλό κείμενο: 0, έντονη γραφή: 0, πλάγια γραφή: 0, έντονη-πλάγια γραφή: 0.

Αυτή η μοντελοποίηση μπορεί να παρουσιαστεί χρησιμοποιώντας τις παρακάτω εξισώσεις (21):

$$\begin{bmatrix} E_p \\ E_b \\ E_i \\ E_{bi} \end{bmatrix} = \begin{bmatrix} B_p \\ B_b \\ B_i \\ B_{bi} \end{bmatrix}$$

$$E_p = \begin{bmatrix} P_p \\ A_p \\ D_p \end{bmatrix}, E_b = \begin{bmatrix} P_b \\ A_b \\ D_b \end{bmatrix}, E_i = \begin{bmatrix} P_i \\ A_i \\ D_i \end{bmatrix}, E_{bi} = \begin{bmatrix} P_{bi} \\ A_{bi} \\ D_{bi} \end{bmatrix} \quad (21)$$

$$B_p = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, B_b = \begin{bmatrix} -0.03 \\ 0 \\ 0 \end{bmatrix}, B_i = \begin{bmatrix} 0.2 \\ 0 \\ 0 \end{bmatrix}, B_{bi} = \begin{bmatrix} 0.13 \\ 0 \\ 0 \end{bmatrix}$$

3.3.6 Συμπεράσματα

Στην παράγραφο 3.3 παρουσιάστηκαν τα αποτελέσματα της μελέτης (θεωρητικά – πειραματικά) με σκοπό τη μοντελοποίηση των μεταβολών της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη η οποία επηρεάζεται από τα τυπογραφικά στοιχεία ενός εγγράφου. Βασιζόμενοι στα πειραματικά αποτελέσματα προτείνεται μία ομάδα κανόνων αντιστοίχισης, για τον τρόπο με τον οποίο τα τυπογραφικά στοιχεία, όπως ο τρόπος γραφής (απλό κείμενο, πλάγια, έντονη και έντονη-πλάγια γραφή), οι χρωματικοί συνδυασμοί γραμματοσειράς/υποβάθρου, ο τύπος και το μέγεθος γραμματοσειράς, επηρεάζουν τη συναισθηματική κατάσταση του αναγνώστη. Η μελέτη βασίζεται στη διαστατική προσέγγιση των συναισθημάτων και η πειραματική διαδικασία σχεδιάστηκε και υλοποιήθηκε σύμφωνα με τις πειραματικές οδηγίες του IAPS, χρησιμοποιώντας τη δοκιμασία Self-Assessment Manikin [182]. Υπάρχουν ενδείξεις ότι οι μεταβολές των συναισθηματικών καταστάσεων είναι ανεξάρτητες γλώσσας και πολιτισμού – πολιτιστικής ταυτότητας. Αυτή η υπόθεση πηγάζει από τη σύγκριση των πειραματικών αποτελεσμάτων της παρούσας μελέτης με εκείνα άλλων γλωσσών και πολιτισμών.

Μερικά σημαντικά αποτελέσματα αυτής της δουλειάς είναι ότι: καθώς αυξάνεται το μέγεθος γραμματοσειράς μέχρι 26px, αυξάνεται η συναισθηματική διάσταση «Ευχαρίστηση». Από αυτό το μέγεθος γραμματοσειράς και πάνω παρατηρείται μείωση. Αντίθετα, οι διαστάσεις «Διέγερση» και «Επικράτηση» μειώνονται μέχρι το μέγεθος 15px και 18px, αντίστοιχα. Πέρα από αυτά τα μεγέθη αυξάνονται και οι δύο διαστάσεις. Επίσης, τα αποτελέσματα καταδεικνύουν ότι μόνο η διάσταση «Ευχαρίστηση» επηρεάζεται από τον τύπο γραμματοσειράς. Ο τύπος γραμματοσειράς «Times New

Roman» θεωρείται ως πιο ευχάριστη γραμματοσειρά από ότι ο «Arial» (23% αύξηση της «Ευχαρίστηση» χρησιμοποιώντας απλό κείμενο). Τέλος, μόνο η διάσταση «Ευχαρίστηση» επηρεάζεται από τον τρόπο γραφής.

Προτείνεται ένα μοντέλο για την ποσοτική περιγραφή των αποτελεσμάτων, χρησιμοποιώντας την πολυωνυμική παλινδρόμηση για την περίπτωση των εναλλαγών του μεγέθους γραμματοσειράς (συνεχείς τιμές) και την ποσοστιαία μεταβολή για την περίπτωση του τύπου γραμματοσειράς και του τρόπου γραφής (διακριτές τιμές). Βασιζόμενοι σε αυτή τη μοντελοποίηση και την ανάγκη υλοποίησης των κανόνων για την αυτοματοποιημένη εξαγωγή της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη [94], προτείνεται ένα γενικό μοντέλο για τη μαθηματική περιγραφή του τρόπου με τον οποίο οι εναλλαγές των τυπογραφικών στοιχείων μπορούν να αντιστοιχηθούν σε μεταβολές των συναισθηματικών καταστάσεων.

Στην επόμενη παράγραφο (3.4 Το σύστημα DocEmoX), προτείνεται η χρήση του προτεινόμενου μοντέλου για την αυτόματη εξαγωγή της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη (εδάφιο 3.3.5) και την επισημείωση εγγράφων σε μορφή Open Document Format (ODF). Έτσι δίνεται η δυνατότητα της αυτόματης εισαγωγής της συναισθηματικής πληροφορίας σε ένα αρχείο .odt. Για παράδειγμα, αυτή η πληροφορία είναι δυνατόν να χρησιμοποιηθεί σε ένα σύστημα που μπορεί αυτόματα να αξιολογήσει/βαθμολογήσει αισθητικά ένα συγκεκριμένο κομμάτι κειμένου ή ακόμα και ολόκληρο το έγγραφο (προφανώς χρησιμοποιώντας τους συγκεκριμένους κανόνες ή κάποιους μελλοντικά πιο σύνθετους) με σκοπό να προταθεί (στον σχεδιαστή και/ή στον προγραμματιστή λογισμικών/ιστοσελίδων) η βελτιστοποίηση της οπτικής παρουσίασης του εγγράφου.

Άλλη μία εφαρμογή είναι η μεταδοση με ακουστικό και/ή απτικό τρόπο της πληροφορίας που θα προκύψει μετά την προσθήκη μετα-δεδομένων από τον οπτικό τρόπο (συναισθηματική κατάσταση αναγνώστη). Για παράδειγμα, συνδυάζοντας το προτεινόμενο μοντέλο (τυπογραφία σε συναισθηματική κατάσταση) και το ακουστικό μοντέλο που προτείνεται από τον Schröder [142] (συναισθηματική κατάσταση σε ομιλία), είμαστε σε θέση να μεταδώσουμε την τυπογραφική πληροφορία των εγγράφων με ακουστικό τρόπο, χρησιμοποιώντας τον έλεγχο της προσωδίας ενός συστήματος μετατροπής Κειμένου-σε-Ομιλία, [27] [191].

3.4 Το σύστημα DocEmoX

Σε αυτό το μέρος του κεφαλαίου παρουσιάζεται η σχεδίαση και ανάπτυξη ενός πρότυπου συστήματος λογισμικού, που ονομάζεται DocEmoX (**Document Emotion eXtractor**), το οποίο α) εξάγει αυτόματα την επαγόμενη συναισθηματική κατάσταση που προκαλείται από τα τυπογραφικά στοιχεία έντυπων ή ηλεκτρονικών εγγράφων και β) επισημειώνει τα έγγραφα με τις σχετικές συναισθηματικές ετικέτες (tags). Βασίζεται στη συναισθηματική μοντελοποίηση του αναγνώστη και τη διαστατική απεικόνιση των συναισθημάτων του στις τρεις διαστάσεις «Ευχαρίστηση», «Διέγερση» και «Επικράτηση».

Το σύστημα DocEmoX υλοποιήθηκε στο πλαίσιο της διπλωματικής εργασίας του φοιτητή Ευγένιου Βλάχου από το διατμηματικό μεταπτυχιακό Ηλεκτρονικής και Αυτοματισμού [192]. Τα αποτελέσματα της εργασίας παρουσιάζονται στο παρόν κείμενο, καθώς και στην εργασία [125]. .

Οι βασικές απαιτήσεις για την υλοποίηση του συστήματος περιλαμβάνουν τη συμμόρφωση στα υπάρχοντα πρότυπα (όπως το Open Document Format for Office Applications – OASIS [58], Digital Accessible Information System/National Information Standards Organization DAISY/NISO [53]) και τις αντίστοιχες συστάσεις του World Wide Web Consortium (W3C) [54].

3.4.1 Η προτεινόμενη αρχιτεκτονική

Η αρχιτεκτονική του συστήματος DocEmoX χωρίζεται σε δύο βασικά μέρη:

1. Το άρθρωμα της προ-επεξεργασίας του εγγράφου (Document pre-Processing module) – Σχήμα 22 - που είναι υπεύθυνο για: α) την ψηφιοποίηση των έντυπων εγγράφων και β) την κανονικοποίηση (normalization) κάθε ηλεκτρονικού εγγράφου στην είσοδο του αρθρώματος, ώστε να υπακούει στα πρότυπα που τίθενται από τις προδιαγραφές.
2. Το άρθρωμα της εξαγωγής της συναισθηματικής κατάστασης και επισημείωσης (Emotional State Extraction - Annotation module) – Σχήμα 24.

3.4.2 Προ-επεξεργασία των εγγράφων

Τα έγγραφα, είτε σε έντυπη είτε ηλεκτρονική μορφή, αναλύονται μέσω του αρθρώματος της προ-επεξεργασίας (Σχήμα 22). Τα έντυπα έγγραφα ψηφιοποιούνται μέσω σάρωσης και οπτικής αναγνώρισης χαρακτήρων (Optical Character Recognition – OCR) και εξάγονται σε αρχείο με επισημείωση (όπως HTML, XHTML, XML, DAISY/NISO, DOC,

DOCx, ODF, κ.α.). Στη συνέχεια, κάθε ένα επισημειωμένο ηλεκτρονικό έγγραφο επεξεργάζεται από το υπο-άρθρωμα Κανονικοποίησης Επισημείωσης (Markup Normalization sub-module), το οποίο μετατρέπει το έγγραφο σε μορφή Document-ML. Το Document-ML ενσωματώνει τα παρακάτω μεταδεδομένα, υπακούοντας στο προτεινόμενο γενικό αρχιτεκτονικό μοντέλο του εγγράφου:

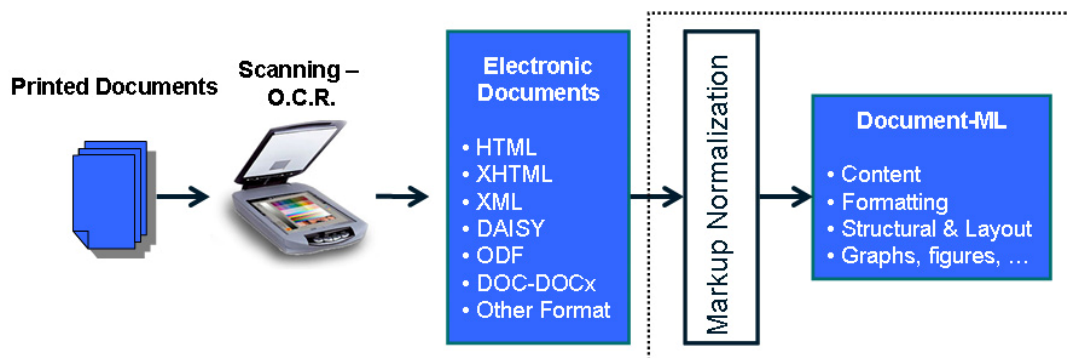
- Μορφοποίηση κειμένου (π.χ. έντονη, πλάγια, έντονη - πλάγια γραφή, μέγεθος γραμματοσειράς).
- Λογική κειμένου (π.χ. κεφάλαιο, τίτλος, υποσημείωση).
- Διάταξη κειμένου (π.χ. στήλες, όρια).
- Μη-κειμενικά (π.χ. εικόνες, γραφήματα, σχέδια).

Επιλέχθηκε το Open Document Format (ODF) [58] ως το πρότυπο στο οποίο πρέπει να υπακούει το Document-ML, για δύο βασικούς λόγους: α) το ODF είναι ένα ευρέως αποδεκτό πρότυπο ISO και β) υπάρχει μια πλειάδα εφαρμογών για την αυτόματη μετατροπή οποιουδήποτε επισημειωμένου εγγράφου σε ODF και αντίστροφα, π.χ. ODFCONVERT για την HTML [193], ODT2DTBOOK για το DAISY/NISO [194], το Sun ODF Plugin for Microsoft Office για το DOC/DOCX [195]. Το έγγραφο ODF είναι μια συλλογή από διάφορα XML αρχεία (υπο-έγγραφα - sub-documents), το καθένα με συγκεκριμένο ρόλο/λειτουργικότητα. Ένα αρχείο τύπου .odt περιλαμβάνει τα εξής αρχεία:

1. content.xml: το περιεχόμενο του εγγράφου,
2. styles.xml: πληροφορίες για την παρουσίαση του περιεχομένου (π.χ. τυπογραφία),
3. meta.xml: μετα-πληροφορίες (meta-information) σχετικά με το περιεχόμενο του εγγράφου, π.χ. το όνομα του συγγραφέα, τελευταία αλλαγή στο έγγραφο κ.α.,
4. αρχεία που χρησιμοποιούνται για επιπλέον πληροφορίες, όπως εικόνες,
5. Mimetype: μία γραμμή κειμένου με τον τύπο MIME του εγγράφου,
6. settings.xml: πληροφορίες για το μέγεθος/θέση του παραθύρου, ρυθμίσεις εκτυπωτή κ.α.,
7. manifest.xml: λίστα με τα αρχεία που περιλαμβάνονται στο αρχείο .odt.

Σύμφωνα με το ODF:

- το αρχείο content.xml μπορεί να περιλαμβάνει πληροφορίες σχετικά με τον τρόπο παρουσίασης του περιεχομένου (σε τοπικό επίπεδο) και
- τα τυπογραφικά στοιχεία περιλαμβάνονται στο αρχείο styles.xml και/ή στο content.xml



Σχήμα 22: Το άρθρωμα της προ-επεξεργασίας εγγράφου του DocEmoX.

Ο Πίνακας 10 παρουσιάζει τον τρόπο επισημείωσης κατά ODF των τυπογραφικών στοιχείων που χρησιμοποιούνται στη διατριβή. Τα δομικά στοιχεία των εγγράφων, όπως επικεφαλίδα (heading), λεζάντα (caption), υποσημείωση (footnote), περιγράφονται στο αρχείο styles.xml χρησιμοποιώντας τα τυπογραφικά στοιχεία. Για παράδειγμα, το Σχήμα 23 παρουσιάζει τον τρόπο που η επικεφαλίδα (δομικό στοιχείο) περιγράφεται στο αρχείο styles.xml.

Πίνακας 10: Ο τρόπος επισημείωσης του μεγέθους γραμματοσειράς, του τύπου γραμματοσειράς και του τρόπου γραφής κατά ODF.

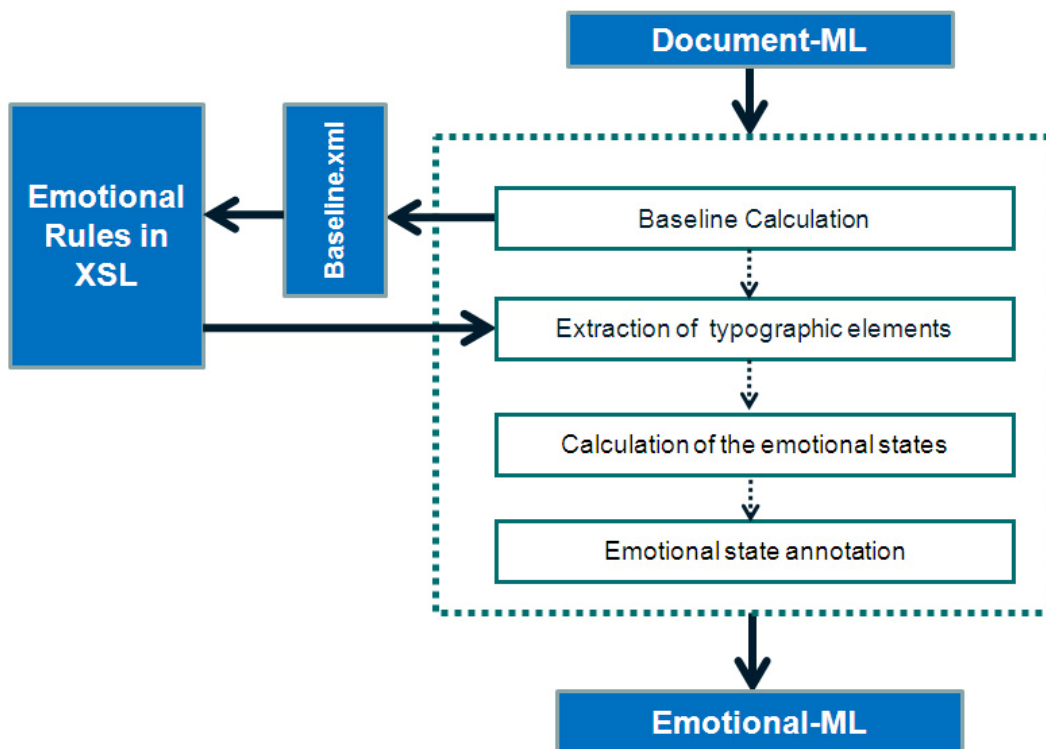
Τυπογραφικά στοιχεία	Περιγραφή ML
Μέγεθος Γραμματοσειράς	fo:font-size, e.g. fo:font-size="10pt"
Τύπος Γραμματοσειράς	style:font-name
Arial	style:font-name="Arial"
Times New Roman	style:font-name="Times New Roman"
Τρόπος Γραφής	
Απλό κείμενο	fo:font-weight="normal" and fo:font-style="normal"
Έντονη γραφή	fo:font-weight="bold"
Πλάγια γραφή	fo:font-style="italic"
Έντονη-πλάγια γραφή	fo:font-weight="bold" and fo:font-style="italic"

```
<style:style style:name="Heading_20_2" style:display-name="Heading 2" style:family="paragraph" style:parent-style-name="Heading" style:next-style-name="Text_20_body" style:class="text" style:default-outline-level="2">  
<style:text-properties fo:font-size="14pt" fo:font-style="italic" fo:font-weight="bold" style:font-size-asian="14pt" style:font-style-asian="italic" style:font-weight-asian="bold" style:font-size-complex="14pt" style:font-style-complex="italic" style:font-weight-complex="bold"/>  
</style:style>
```

Σχήμα 23: Το δομικό στοιχείο «Heading», όπως περιγράφεται στο αρχείο styles.xml

3.4.3 Εξαγωγή της επαγόμενης συναισθηματικής κατάσταση του αναγνώστη και επισημείωση των εγγράφων με την αντίστοιχη πληροφορία

Το άρθρωμα Εξαγωγής και Επισημείωσης Συναισθηματικής Κατάστασης (Σχήμα 24) λαμβάνει ως είσοδο το αρχείο Document-ML (σε μορφή ODF), εφαρμόζει τους συναισθηματικούς κανόνες (Emotional Rules) και παράγει το επισημειωμένο αρχείο Emotional-ML χρησιμοποιώντας την Emotional Markup Language.



Σχήμα 24: Το άρθρωμα του DocEmoX για την εξαγωγή και επισημείωση της συναισθηματικής κατάστασης.

Η Emotion Markup Language (EmotionML), προτείνεται από τη W3C's Incubator Group [196]:

1. επιτρέπει σε ένα τεχνολογικό συστατικό (technological component) να απεικονίσει (represent) και να επεξεργαστεί (process) δεδομένα, και
2. παρέχει τη δυνατότητα λειτουργίας μεταξύ διαφορετικών τεχνολογικών συστατικών για την επεξεργασία των δεδομένων (interoperability).

Είναι πολύ δύσκολο να τυποποιηθεί η περιγραφή των συναισθημάτων χρησιμοποιώντας συγκεκριμένες περιγραφικές μεθόδους (descriptors) και μεθοδολογίες. Για αυτό, η EmotionML προτείνει και υποστηρίζει πολλαπλούς τρόπους αναπαράστασης των συναισθημάτων χρησιμοποιώντας επισημείωση, π.χ. χρησιμοποιώντας συγκεκριμένα συναισθήματα, συναισθηματικές καταστάσεις και/ή περιγραφή αξιολόγησης (appraisal) των συναισθημάτων. Ακολουθώντας τη διαστατική προσέγγιση των συναισθημάτων, οι τρεις καταστάσεις μπορούν να περιγραφούν από την EmotionML χρησιμοποιώντας είτε διπολικές (bipolar) τιμές από το -1 έως το 1 είτε διπολικές διακριτές τιμές (bipolar discrete values) (παράδειγμα στην Σχήμα 25).

```
<emotion>
  <dimensions set="valenceArousalPotency" >
    <valence value="0.9"/>
    <arousal value="-0.4"/>
    <potency value="-0.1"/>
  </dimensions>
</emotion>
```

(α)

```
<emotion>
  <dimensions set="valenceArousalPotency" >
    <valence value="very high"/>
    <arousal value="very negative"/>
    <potency value="slightly negative"/>
  </dimensions>
</emotion>
```

(β)

Σχήμα 25: Η περιγραφή της συναισθηματικής κατάστασης σύμφωνα με την EmotionML χρησιμοποιώντας: (α) διπολικές τιμές από το -1 έως το 1, και (β) διπολικές διακριτές τιμές.

Για παράδειγμα έστω ότι έχουμε τον παρακάτω συναισθηματικό κανόνα:

Η μετάβαση από το απλό κείμενο σε πλάγια-έντονη γραφή, διατηρώντας το τύπο της γραμματοσειράς σε «Times New Roman» και το μέγεθος της γραμματοσειράς «12pt», μειώνει την «Ευχαρίστηση» κατά 21% και οι διαστάσεις «Διέγερση» και «Επικράτηση» δε μεταβάλλονται.

Το Σχήμα 27 δείχνει την περιγραφή σε XSL του παραπάνω κανόνα. Οι συναισθηματικοί κανόνες δεν είναι πάντα τόσο απλοί όπως στο παραπάνω παράδειγμα. Μπορεί να περιλαμβάνουν κανόνες οι οποίοι περιγράφονται από μία εξίσωση δευτέρου ή τρίτου βαθμού (π.χ. μέγεθος γραμματοσειράς), όπως και εναλλαγές στον τρόπο γραφής, το μέγεθος και τον τύπο γραμματοσειράς (π.χ. εναλλαγή από απλό κείμενο – «Times New Roman» χρησιμοποιώντας μέγεθος γραμματοσειράς «12pt» σε έντονη γραφή – «Arial» χρησιμοποιώντας μέγεθος γραμματοσειράς «20pt»).

Η λειτουργικότητα του αρθρώματος της εξαγωγής της συναισθηματικής κατάστασης και η επισημείωση με αυτή την πληροφορία βασίζεται στην εφαρμογή της ομάδας των συναισθηματικών κανόνων στα αρχεία content.xml και styles.xml. Η επισημείωση του εγγράφου έπεται των υπολογισμών των συναισθηματικών καταστάσεων για κάθε τυπογραφική εναλλαγή. Το σύστημα παράγει-δημιουργεί επισημείωση και στα δύο αρχεία ακολουθώντας τις συστάσεις που παρέχονται από την EmotionML. Το αρχείο εξόδου .odt είναι εμπλουτισμένο με την επισημείωση συναισθηματικής κατάστασης, αλλά διατηρεί τη λειτουργικότητα του κατά τη χρήση του στο OpenOffice. Το Σχήμα 28 παρουσιάζει το αρχείο εξόδου Emotional-ML για το παράδειγμα που χρησιμοποιήθηκε στο Σχήμα 27.

```
<font>
  <size>10pt</size>
  <name>Times New Roman</name>
  <typesetting>
    <weight>bold</weight>
    <style>normal</style>
  </typesetting>
</font>
```

Σχήμα 26: Ένα παράδειγμα της δομής του αρχείου baseline.xml με τα βασικά τυπογραφικά στοιχεία

```
<xsl:when test="style:text-properties[@fo:font-weight='bold' and
@fo:font-style='italic']">
  <xsl:copy>
    <xsl:apply-templates select="node()|@*" />
    <xsl:element name="emotion">
      <xsl:element name="dimensions">
        <xsl:attribute name="set">valenceArousalPotency</xsl:attribute>
        <xsl:element name="valence">
          <xsl:attribute name="value">-0.21</xsl:attribute>
        </xsl:element>
        <xsl:element name="arousal">
          <xsl:attribute name="value">0.0</xsl:attribute>
        </xsl:element>
        <xsl:element name="potency">
          <xsl:attribute name="value">0.0</xsl:attribute>
        </xsl:element>
      </xsl:element>
    </xsl:element>
  </xsl:copy>
</xsl:when>
```

Σχήμα 27: Παράδειγμα XSL υλοποίησης ενός συναισθηματικού κανόνα

```
<style:style style:name="Heading_20_2" style:display-
name="Heading 2" style:family="paragraph" style:parent-
style-name="Heading" style:next-style-name="Text_20_body"
style:class="text" style:default-outline-level="2">
<style:text-properties fo:font-size="12pt" fo:font-
style="italic" fo:font-weight="bold" style:font-size-
asian="12pt" style:font-style-asian="italic" style:font-
weight-asian="bold" style:font-size-complex="12pt"
style:font-style-complex="italic" style:font-weight-
complex="bold"/>
  <emotion>
    <dimensions set="valenceArousalPotency" >
      <valence value="-0.21"/>
      <arousal value="0.0"/>
      <potency value="0.0"/>
    </dimensions>
  </emotion>
</style:style>
```

Σχήμα 28: Τα συναισθηματικά επισημειωμένα δομικά στοιχεία στο αρχείο styles.xml

3.4.4 Υλοποίηση και αξιολόγηση του συστήματος

Η αρχιτεκτονική του DocEmoX που περιγράφηκε παραπάνω υλοποιήθηκε σε MS Windows XP και Linux (Ubuntu 8.10) χρησιμοποιώντας:

- Java SE Development Kit (JDK 1.6.0_11) και Java SE Runtime Environment (JRE 6)
- Xerces2 java Parser Ver. 2.9.1

- Apache Ant scripts (apache-ant-antunit-1.1)
- OpenOffice suite 3.0.0

Για την αξιολόγηση του συστήματος χρησιμοποιήσαμε ένα αντιπροσωπευτικό δοκιμαστικό αρχείο .odt που περιείχε όλους τους δυνατούς συνδυασμούς των τυπογραφικών στοιχείων που μας ενδιαφέρουν σε αυτή την εργασία (τύπος γραμματοσειράς «Arial» και «Times New Roman», τρόπος γραφής - απλό κείμενο, έντονη, πλάγια και έντονη/πλάγια γραφή, καθώς και διαφορετικά μεγέθη γραμματοσειράς από 10pt έως 17pt). Δοκιμάσαμε όλες τις ομάδες κανόνων, από τους πιο απλούς, χρησιμοποιώντας σενάρια όπως «αλλαγή μεγέθους γραμματοσειράς» ή «αλλαγή τύπου γραμματοσειράς», έως πιο σύνθετους όπως «αλλαγή του μεγέθους γραμματοσειράς και ταυτόχρονα αλλαγή του τύπου γραμματοσειράς και τρόπου γραφής». Να αναφερθεί ότι υπολογίστηκαν χειροκίνητα οι τιμές των συναισθηματικών καταστάσεων. Αυτό έγινε με σκοπό να συγκρίνουμε τα αποτελέσματα με εκείνα που προκύπτουν από το σύστημα DocEmoX και για να ελεγχθεί η ορθότητα λειτουργίας του.

4. ΑΚΟΥΣΤΙΚΟΠΟΙΗΣΗ ΤΩΝ ΤΥΠΟΓΡΑΦΙΚΩΝ ΣΤΟΙΧΕΙΩΝ ΜΕ ΤΗΝ ΧΡΗΣΗ ΕΚΦΡΑΣΤΙΚΗΣ ΣΥΝΘΕΤΙΚΗΣ ΟΜΙΛΙΑΣ

4.1 Εισαγωγή

Έχοντας πλέον τη μαθηματική περιγραφή των αντιστοιχήσεων α) τυπογραφικών εναλλαγών σε μεταβολές της επαγόμενης συναισθηματικής κατάστασης και β) μεταβολών συναισθηματικής κατάστασης σε προσωδιακές μεταβολές, δίνεται η δυνατότητα περαιτέρω μελέτης του συνδυασμού τους για την ανάδειξη των κανόνων ακουστικής απόδοσης των τυπογραφικών στοιχείων, σύμφωνα με την προτεινόμενη μεθοδολογία. Τα αποτελέσματα του εδαφίου 3.3.5, δείχνουν τον τρόπο με τον οποίο μεταβάλλονται οι συναισθηματικές διαστάσεις που επηρεάζονται από τις τυπογραφικές εναλλαγές, και κατ' επέκταση τα προσωδιακά χαρακτηριστικά μέσω της ακουστικής αντιστοίχησης.

Οι ακουστικές μεταβολές που προκύπτουν είναι μικρές, στις περισσότερες περιπτώσεις στα όρια ή ακόμα και παρακάτω από την ακουστική αντίληψη του ακροατή. Για να επιλυθεί αυτό το πρόβλημα, εισάγεται η γραμμική κβάντιση αυτών των μεταβολών και η αντιστοίχηση των επιπέδων μεταβολής που προκύπτουν σε ευδιάκριτες μεταβολές.

Μέσα από αυτόν τον συλλογισμό προκύπτουν πολλαπλές προσωδιακές μεταβολές για κάθε μία τυπογραφική εναλλαγή, που αν συνδυαστούν π.χ. σε ένα έγγραφο, τότε μπορούν να προκαλέσουν σύγχυση ακόμα και ενόχληση στον ακροατή. Αυτό οφείλεται στους διάφορους συνδυασμούς εναλλαγής τυπογραφικών στοιχείων. Για αυτό το λόγο προτείνεται μια πειραματική διαδικασία με σκοπό την επιλογή των βέλτιστων προσωδιακών μεταβολών.

4.2 Το προσωδιακό μοντέλο

Ακολουθώντας την προτεινόμενη μεθοδολογία και χρησιμοποιώντας το μοντέλο επαγόμενης συναισθηματικής κατάστασης, που παρουσιάζεται στο Κεφάλαιο 3, υπολογίζονται οι μεταβολές της συναισθηματικής κατάστασης από τη συναισθηματική βάση αναφοράς (εξισώσεις 17 και 18). Η συναισθηματική κατάσταση που πηγάζει από την αντίστοιχη τυπογραφική βάση αναφοράς (typographic baseline) ορίζεται ως «συναισθηματική βάση αναφοράς». Ως τυπογραφική βάση αναφοράς ορίζεται: *οι πιο συχνά παρατηρούμενες τιμές των τυπογραφικών στοιχείων σε όλο το έγγραφο.*

Τα πιο συνήθη μεγέθη που περιγράφουν τη βάση αναφοράς των εγγράφων είναι 10pt, 12pt και 14pt (σε όλα τα βιβλία και εφημερίδες) [94]. Επίσης, παρατηρείται ότι η αύξηση του μεγέθους γραμματοσειράς (για παράδειγμα στον τίτλο, υπότιτλο) μπορεί να έχει

τιμές 12pt, 14pt, 16pt, 18pt, 20pt, 22pt και η μείωση 8pt, 9pt, 10pt, 11pt, 12pt, 13pt (για παράδειγμα στην υποσημείωση), που εξαρτώνται από τη βάση αναφοράς του μεγέθους. Μέσω της μαθηματικής περιγραφής ο παραπάνω συλλογισμός μπορεί να γίνει εύκολα κατανοητός ως εξής:

$$\Delta E = E_{\text{final}} - E_{\text{baseline}} \quad (22)$$

όπου:

$$E = \begin{bmatrix} P \\ A \\ D \end{bmatrix}$$

P, A και D: οι διαστάσεις «Ευχαρίστηση», «Διέγερση» και «Επικράτηση» αντίστοιχα.

E_{final} , η ποσοστιαία μεταβολή των τριών διαστάσεων από την ουδέτερη κατάσταση για το αντίστοιχο τυπογραφικό στοιχείο.

E_{baseline} , η ποσοστιαία μεταβολή των τριών διαστάσεων από την ουδέτερη κατάσταση για την αντίστοιχη τυπογραφική βάση αναφοράς.

Μετρώντας το μέγεθος της γραμματοσειράς σε εικονοστοιχεία (pixels), είναι δυνατός ο χειρισμός του μεγέθους σύμφωνα με τις προδιαγραφές της οθόνης-συσκευής και ο πλήρης έλεγχος του μετρούμενου ύψους. Τα σημεία (points) είναι για το χαρτί και τα εικονοστοιχεία (pixels) για την οθόνη. Το φυσικό μήκος των 72 points μεγέθους γραμματοσειράς είναι 2.54cm (1 ίντσα) [197]. Σύμφωνα με τις προδιαγραφές της πειραματικής διαδικασίας (ανάλυση οθόνης 1024x768 σε 17 ίντσες διαγώνια διάσταση οθόνης), 72 σημεία είναι ίσα με 75.3 εικονοστοιχεία, εισάγοντας ένα παράγοντα μετατροπής $\varphi = \frac{pt}{px} = 0.956$. Έτσι, 1 σημείο ισούται με $d_{po} = \frac{1}{72} \cdot 2.54 = 0.035 \text{ cm}$ και ένα εικονοστοιχείο ισούται με $d_{pi} = d_{po} * \varphi = 0.033 \text{ cm}$ (φυσικό ύψος γραμματοσειράς). Για παράδειγμα, το μέγεθος γραμματοσειράς 9 σημείων έχει φυσικό ύψος $9 \cdot d_{po} = 0.315 \text{ cm}$. Το μέγεθος γραμματοσειράς σε εικονοστοιχεία προκύπτει $\frac{0,315}{d_{pi}} = 9.5 \text{ px}$, στρογγυλοποιείται σε 10 px (η τιμή των εικονοστοιχείων πρέπει να είναι ακέραιος αριθμός σύμφωνα με την εξίσωση 14). Ένα επιπλέον παράδειγμα, το μέγεθος γραμματοσειράς 18 σημείων μπορεί εύκολα να μετατραπεί $18 \cdot d_{po} = 0.635 \text{ cm}$ και $\frac{0,635}{d_{pi}} = 19.1 \text{ px}$, στρογγυλοποιημένο 19 px. Στον Πίνακα 11 παρουσιάζονται τα μεγέθη

γραμματοσειράς σε σημεία και εικονοστοιχεία (σε παρένθεση η στρογγυλοποιημένη τιμή). Παρατηρώντας τον πίνακα, η μετατροπή μπορεί να πραγματοποιηθεί με τη χρήση της γενικής εξίσωσης:

$$[\text{μέγεθος γραμματοσειράς σε εικονοστοιχεία}] = [\text{μέγεθος γραμματοσειράς σε σημεία}] + 1$$

Από τις εξισώσεις 14 και 22, υπολογίζονται οι διαφορές σε ποσοστιαία κλίμακα των διαστάσεων «Ευχαρίστηση», «Διέγερση» και «Επικράτηση» σύμφωνα με τις μεταβολές του μεγέθους γραμματοσειράς για διαφορετικές τιμές της βάσης αναφοράς. Οι τιμές που προκύπτουν από τους υπολογισμούς παρουσιάζονται στον Πίνακα 12.

Παρόμοια προκύπτει για τις μεταβολές της συναισθηματικής κατάστασης κατά τη μεταβολή του τρόπου γραφής, όπως προσδιορίζονται στην παράγραφο 3.3.4. Η μόνη διάσταση που επηρεάζεται είναι η «Ευχαρίστηση». Συνοψίζοντας:

1. Στην περίπτωση τύπου γραμματοσειράς «Arial»:

- Εναλλαγή από απλό κείμενο σε έντονη γραφή \Rightarrow
 \Rightarrow μείωση της «Ευχαρίστησης» (-3%)
- Εναλλαγή από απλό κείμενο σε πλάγια γραφή \Rightarrow
 \Rightarrow αύξηση της «Ευχαρίστησης» (+20%)
- Εναλλαγή από απλό κείμενο σε έντονη-πλάγια γραφή \Rightarrow
 \Rightarrow αύξηση της «Ευχαρίστησης» (+13%)

2. Στην περίπτωση τύπου γραμματοσειράς «Times New Roman»:

- Εναλλαγή από απλό κείμενο σε έντονη γραφή \Rightarrow
 \Rightarrow μείωση της «Ευχαρίστησης» (-31%)
- Εναλλαγή από απλό κείμενο σε πλάγια γραφή \Rightarrow
 \Rightarrow μείωση της «Ευχαρίστησης» (-19%)
- Εναλλαγή από απλό κείμενο σε έντονη-πλάγια γραφή \Rightarrow
 \Rightarrow μείωση της «Ευχαρίστησης» (-21%)

Πίνακας 11: Μεγέθη γραμματοσειράς σε σημεία και η μετατροπή τους σε εικονοστοιχεία.

Μέγεθος γραμματοσειράς σε σημεία	Μέγεθος γραμματοσειράς σε εικονοστοιχεία (στρογγυλοποιημένες τιμές)
8	8.5 (9)
9	9.5 (10)
10	10.6 (11)
11	11.7 (12)
12	12.7 (13)
13	13.8 (14)
14	14.8 (15)
16	17 (17)
18	19.1 (19)
20	21.2 (21)
22	23.3 (23)

Πίνακας 12: Οι μεταβολές των συναισθηματικών διαστάσεων από τη βάση αναφοράς σε ποσοστιαία κλίμακα.

	Εναλλαγές μεγέθους γραμματοσειράς (pt)	ΔΕ (Μεταβολές συναισθηματικής κατάστασης)		
		Ευχαρίστηση (%)	Διέγερση (%)	Επικράτηση (%)
+2pt	10→12	24.78	-17.95	-21.55
	12→14	20.60	-8.35	-14.74
	14→16	16.41	-0.79	-8.88
+4pt	10→14	45.38	-26.29	-36.29
	12→16	37.01	-9.14	-23.62
	14→18	28.64	3.91	-12.88
+6pt	10→16	61.79	-27.09	-45.17
	12→18	49.24	-4.43	-27.61
	14→20	36.68	12.07	-12.93
+8pt	10→18	74.02	-22.38	-49.16
	12→20	57.28	3.73	-27.67
	14→22	40.54	21.64	-10.02
-1pt	10→9	-13.96	13.22	13.63
	12→11	-11.87	7.65	9.86
	14→13	-9.77	3.10	6.58
-2pt	10→8	-28.96	29.60	29.32
	12→10	-24.78	17.95	21.55
	14→12	-20.60	8.35	14.74

Η αντιστοίχιση των τυπογραφικών εναλλαγών σε προσωδιακές μεταβολές μπορεί να επιτευχθεί συνδυάζοντας τα παραπάνω αποτελέσματα με το μοντέλο της εκφραστικής συνθετικής ομιλίας που παρουσιάζεται στην [142] και χρησιμοποιείται από το σύστημα OpenMary [143]. Οι τιμές που προκύπτουν για το μέγεθος και τον τύπο γραμματοσειράς – τρόπο γραφής παρουσιάζονται στους Πίνακες 13α και 13β, αντίστοιχα. Οι μεταβολές του τονικού ύψους εκφράζονται σε Hz και αντιστοιχούν στην αύξηση ή μείωση από τη βάση αναφοράς τονικού ύψους (για παράδειγμα 110Hz). Οι μεταβολές ρυθμού ομιλίας υπολογίζονται από την ποσοστιαία μεταβολή από τη βάση αναφοράς (συνήθως 180 λέξεις το λεπτό - wpm) και για την ένταση, η αύξηση ή η μείωση σε κλίμακα 0-100 με βάση αναφοράς την τιμή 50 (για παράδειγμα -4% αντιστοιχεί στην τιμή 46 της απόλυτης κλίμακας).

Πίνακας 13: Οι προσωδιακές μεταβολές για α) εναλλαγές του μεγέθους γραμματοσειράς για τρεις διαφορετικές βάσεις αναφοράς, 10pt, 12pt και 14 pt και β) εναλλαγές τρόπου γραφής από το απλό κείμενο, για κάθε τύπο γραμματοσειράς.

(α)

Εναλλαγές μεγέθους γραμματοσειράς		Προσωδιακές Μεταβολές			Εναλλαγές μεγέθους γραμματοσειράς		Προσωδιακές Μεταβολές		
		Τονικό ύψος (Hz)	Ρυθμός (%)	Ένταση (%)			Τονικό ύψος (Hz)	Ρυθμός (%)	Ένταση (%)
+2pt	10→12	-0.8	-4.0	-5.92	+8pt	10→18	5.6	3.6	-7.39
	12→14	1.0	-0.1	-2.75		12→20	9.6	13.3	1.23
	14→16	2.3	2.9	-0.26		14→22	11.5	18.9	7.14
+4pt	10→14	0.3	-4.1	-8.68	-1pt	10→9	1.2	3.8	4.36
	12→16	3.3	2.8	-3.02		12→11	0.1	1.4	2.52
	14→18	5.3	7.7	1.29		14→13	-0.7	-0.4	1.02
+6pt	10→16	2.6	-1.2	-8.94	-2pt	10→8	3.1	9.0	9.77
	12→18	6.4	7.6	-1.46		12→10	0.8	4.0	5.92
	14→20	8.6	13.4	3.98		14→12	-1.0	0.1	2.75

(β)

	Απλή → Έντονη	Απλή → Πλάγια	Απλή → Έντονη-Πλάγια
Times New Roman			
Τονικό ύψος (Hz)	-3.1	-1.9	-2.1
Ρυθμός (%)	-6.2	-3.8	-4.2
Ένταση (%)	0.0	0.0	0.0
Arial			
Τονικό ύψος (Hz)	-0.3	2.0	1.3
Ρυθμός (%)	-0.6	4.0	2.6
Ένταση (%)	0.0	0.0	0.0

4.2.1 Ελάχιστες προσωδιακές μεταβολές της ομιλίας που είναι αντιληπτές από τον ακροατή

Από τα δεδομένα των Πινάκων 13α και 13β, παρατηρούμε ότι οι προσωδιακές μεταβολές είναι πολύ μικρές, επειδή είναι μικρές και οι μεταβολές των συναισθηματικών καταστάσεων που πηγάζουν από τις εναλλαγές του μεγέθους γραμματοσειράς. Αυτές οι μεταβολές δεν είναι ευδιάκριτες από τον ακροατή (όπως περιγράφεται παρακάτω), επομένως, χρειάζεται να «επταυξηθούν» ώστε να γίνουν αντιληπτές και κατανοητές. Για αυτό, προτείνεται η κανονικοποίηση των τιμών χρησιμοποιώντας τη γραμμική κβάντιση (εδάφιο 4.2.2). Πριν προχωρήσουμε όμως στην κβάντιση πρέπει να οριστούν οι ελάχιστες τιμές αυτών των επιπέδων κβάντισης.

Τονικό ύψος

Οι μεταβολές του τονικού ύψους για περισσότερο από τρία ημιτόνια μπορεί να είναι αξιόπιστα διακριτές [198]. Αλλά, όπως προτείνεται στην [199], τονικές διαφορές 1.5 ημιτονίου δημιουργούν διαφορές στην αντίληψη της κορύφωσης (prominence). Οι Truillet et al. [151] στην εργασία τους προτείνουν 13% αύξηση του προεπιλεγμένου τονικού ύψους. Όμοια, οι Xydias et al. [36] προτείνουν 15% ελάχιστη μεταβολή χρησιμοποιώντας ως βάση αναφοράς τη συχνότητα 110Hz ανδρικής φωνής. Οι Kallinen et al. [150] μελέτησαν τη μείωση κατά δύο ημιτόνια, ώστε να γίνεται αντιληπτή η διάκριση συγκεκριμένων πληροφοριών. Οι Braga et al. [200] απέδειξαν ότι οι ακροατές προτιμούν γυναικείες φωνές με μέσο F0 που κυμαίνεται από 186Hz και 206Hz. Οι Chan et al. [201] στα πειράματά τους χρησιμοποιούν 4 επίπεδα F0 (με βήμα 20Hz - 200, 220, 240 και 260Hz) χρησιμοποιώντας γυναικεία φωνή.

Με διαφορά 2 ημιτόνιων (~12%) μπορεί να διακρίνονται δύο γειτονικά στοιχεία (components), ακόμα και όταν είναι παρόμοιας έντασης. Δεν υπάρχει σημαντική βελτίωση για μεταβολή από μηδέν σε ένα ημιτόνιο. Έχουμε όμως το αντίθετο αποτέλεσμα για μεταβολή από μηδέν σε δύο ημιτόνια. Επίσης, η αναγνώριση βελτιώνεται σημαντικά από 5 σε 10 ημιτόνια [202].

Ρυθμός ομιλίας

Οι Xydias et al. [36] προτείνουν μεταβολή ρυθμού κατά 15% (χρησιμοποιώντας ως βάση αναφοράς 140wpm). Έχοντας ως φυσιολογικό ρυθμό ομιλίας 125–225 wpm, και αυξάνοντας το ρυθμό, δεν παρατηρείται σημαντική συσχέτιση με τη μειωμένη ανάκτηση (recall) για γηγενείς ακροατές [203]. Υπερβαίνοντας το όριο των 225 wpm, υπάρχει απόκλιση στην κατανόηση [203]. Στη μελέτη αυτή χρησιμοποιούν 2 επίπεδα ρυθμού i) 155wpm (ως φυσιολογικό) και ii) 178wpm (ως fast-normal).

Ένταση

Είναι πειραματικά αποδεδειγμένα ότι η γραμμική αύξηση του πλάτους του σήματος φωνής κατά 1.3 φορές στις τονισμένες συλλαβές (stressed syllables) συνεισφέρει στη θετική αξιολόγηση από τους ακροατές της παραγόμενης συνθετικής ομιλίας (αύξηση του επιπέδου της έντασης κατά 1 - 3 dB) [204]. Στην εργασία [36] αυξάνεται/μειώνεται η ένταση κατά 30% (χρησιμοποιείται μόνο ένα επίπεδο μεταβολής).

Παύσεις

Η ομιλία υψηλής ποιότητας χωρίς γραμματικές παύσεις (grammatical speech pauses) μέσα στην πρόταση μπορεί να είναι πολύ κατανοητή και αποδεκτή. Όμως, όσο η ποιότητα της ομιλίας είναι λιγότερο φυσιολογική ή η ομιλία ακούγεται σε θορυβώδες περιβάλλον, η εισαγωγή γραμματικών παύσεων μπορεί να βοηθήσει στη διατήρηση της καταληπτότητας [205]. Οι παύσεις αποτελούν δείκτες της δομής της πληροφορίας [206]. Η ύπαρξη υποτονικών φράσεων (Intonational phrases) – που χαρακτηρίζονται από σιγή και συνήθως από χαρακτηριστικές προσωδιακές μεταβολές – είναι απαραίτητη μεταξύ των προτάσεων και συνήθως κοντά στα όρια στίξης [207]. Φωνολογικές φράσεις – που χαρακτηρίζονται από σιγή παρά από τη χαρακτηριστική μεταβολή του τονικού ύψους μπορεί να αξιολογηθούν με δυσκολία [207]. Οι ακροατές φαίνεται να προτιμούν φωνές με υψηλό ρυθμό ομιλίας συνδυασμένο με μεγάλες παύσεις στην έκφραση [200].

4.2.2 Γραμμική κβάντιση των προσωδιακών μεταβολών

Τα επίπεδα της γραμμικής κβάντισης προσδιορίζονται, υπολογίζοντας τη διαφορά μεταξύ ελάχιστων και μέγιστων τιμών για κάθε προσωδιακό στοιχείο και διαιρώντας το

με τα έξι επίπεδα αύξησης ή μείωσης (Πίνακες 14α και 14β). Ο αριθμός των επιπέδων θεωρείται ως ο βέλτιστος και προκύπτει από παρατηρήσεις, ώστε να έχουμε όσο το δυνατόν τις λιγότερες διπλότυπες περιπτώσεις στις μεταβολές των προσωδιακών χαρακτηριστικών (που παρουσιάζεται στις περιπτώσεις για μικρότερο ή/και μεγαλύτερο αριθμό).

Πίνακας 14: Τα έξι επίπεδα γραμμικής κβάντισης των προσωδιακών μεταβολών για: α) το μέγεθος και β) τον τύπο γραμματοσειράς – τρόπο γραφής.

(α)

	Τονικό ύψος (Hz)	Ρυθμός (%)	Ένταση (%)
<i>Ελάχιστες τιμές</i>	-1.03	-4.07	-8.94
<i>Μέγιστες τιμές</i>	11.55	18.93	9.77
<i>Εύρος τιμών</i>	12.58	23.00	18.71
Επίπεδα μεταβολών	Όρια κάθε επιπέδου		
± 1	± 2.10	± 3.83	± 3.12
± 2	± 4.20	± 7.66	± 6.24
± 3	± 6.30	± 11.49	± 9.36
± 4	± 8.40	± 15.32	± 12.48
± 5	± 10.50	± 19.15	± 15.60
± 6	± 12.60	± 22.98	± 18.72

(β)

	Τονικό ύψος (Hz)	Ρυθμός (%)
<i>Ελάχιστες τιμές</i>	-3.1	-6.2
<i>Μέγιστες τιμές</i>	2.0	4.0
<i>Εύρος τιμών</i>	5.1	10.2
Επίπεδα μεταβολών	Όρια κάθε επιπέδου	
± 1	± 1.00	± 2.00
± 2	± 2.00	± 4.00
± 3	± 3.00	± 6.00
± 4	± 4.00	± 8.00
± 5	± 5.00	± 10.00
± 6	± 6.00	± 12.00

Έτσι προκύπτει, ότι: Οι διαφορές των κορυφών στις προσωδιακές μεταβολές κατά την εναλλαγή του μεγέθους γραμματοσειράς είναι: τονικό ύψος 12.58Hz, ρυθμός 23% και

ένταση 18.71% (Πίνακας 14α). Διαιρώντας αυτές τις τιμές με 6, η μεταβολή ενός επιπέδου αντιστοιχεί σε τονικό ύψος 2.1Hz, ρυθμό 3.83% και ένταση 3.12%. Για αύξηση του τονικού ύψους από 0 Hz σε 2.1Hz (ή μείωση κατά -2.1Hz) αντιστοιχεί σε επίπεδο μεταβολής +1 (ή -1 επίπεδο αντίστοιχα), 2.1Hz σε 4.2Hz (ή -2.1Hz σε -4.2Hz) σε +2 επίπεδα (ή -2 επίπεδα) κτλ. Στον Πίνακα 15α προσδιορίζονται οι κβαντισμένες μεταβολές για κάθε προσωδιακό στοιχείο. Υπάρχουν έξι περιπτώσεις μεταβολών (αύξηση κατά +2pt, +4pt, +6pt, +8pt και μείωση -1pt και -2pt) για διαφορετικές βάσεις αναφοράς (10pt, 12pt και 14pt).

Για την περίπτωση του τρόπου γραφής και του τύπου γραμματοσειράς, λόγω του μικρότερου αριθμού συνδυασμών στις προσωδιακές μεταβολές, από ότι σε αυτές που παρουσιάζονται για την περίπτωση του μεγέθους, δεν περιοριζόμαστε στο να αναθέσουμε για κάθε επίπεδο την τιμή που προκύπτει, διαιρώντας το εύρος ελάχιστων-μέγιστων τιμών με το 6. Για αυτό χρησιμοποιήσουμε για το τονικό ύψος 1Hz και το ρυθμό 2% (Πίνακες 14β και 15β).

4.2.3 Προσωδιακές μεταβολές μετά τη γραμμική κβάντιση

Ακολουθώντας τα αποτελέσματα των μελετών του εδαφίου 4.2.1, για τις ελάχιστα διακριτές τιμές των προσωδιακών μεταβολών που είναι αντιληπτές από τους ακροατές, οι ποσοστιαίες τιμές κάθε επιπέδου αύξησης/μείωσης για κάθε προσωδιακό στοιχείο είναι:

- 12% για τον τόνο (αντιστοιχεί σε αύξηση/μείωση περίπου 2 ημιτόνια). Οι Schröder και Grice στην εργασία τους [208] χρησιμοποιούν ως κανονικό τόνο συχνότητες 74 - 115Hz και ως αυξημένο 103 - 161Hz. Έτσι, στην παρούσα μελέτη, τα 6 επίπεδα προσεγγίζουν τα 189Hz, που είναι λίγο υψηλότερες από αυτές που χρησιμοποιούν και χαρακτηρίζουν ως αυξημένο τόνο οι Schröder και Grice χωρίς να ξεπερνιέται το όριο των 202Hz, που είναι πολύ μεγάλη και ακούγεται μη-φυσική και ακραία.
- 10% - 14wpm για το ρυθμό (6 επίπεδα αύξησης αντιστοιχούν σε μέγιστο ρυθμό των 224wpm χωρίς να υπερβαίνει το ανώτατο όριο των 225wpm, όπως προσδιορίστηκε στο εδάφιο 4.2.1)
- 15% για την ένταση αντιστοιχεί σε περισσότερο από >1dB αύξηση/μείωση.
- Πριν και μετά από κάθε προσωδιακή μεταβολή, υπάρχει μία παύση διάρκειας 150 ms.

Πίνακας 15: Τα επίπεδα γραμμικής κβάντισης των προσωδιακών εναλλαγών όπως προκύπτουν για: α) το μέγεθος και β) τον τύπο γραμματοσειράς – τρόπο γραφής.

(α)

Εναλλαγές μεγέθους γραμματοσειράς		Μεταβολές		
		Τονικό ύψος (επίπεδο)	Ρυθμός (επίπεδο)	Ένταση (επίπεδο)
+2pt	10→12	-1	-2	-2
	12→14	+1	-1	-1
	14→16	+2	+1	-1
+4pt	10→14	+1	-2	-3
	12→16	+2	+1	-1
	14→18	+3	+3	+1
+6pt	10→16	+2	-1	-3
	12→18	+4	+2	-1
	14→20	+5	+4	+2
+8pt	10→18	+3	+1	-3
	12→20	+5	+4	+1
	14→22	+6	+5	+3
-1pt	10→9	+1	+1	+2
	12→11	+1	+1	+1
	14→13	-1	-1	+1
-2pt	10→8	+2	+3	+4
	12→10	+1	+2	+2
	14→12	-1	+1	+1

(β)

Εναλλαγές από την Απλή γραφή		Έντονη	Πλάγια	Έντονη-Πλάγια
Times New Roman	Τονικό ύψος (Επίπεδο)	-4	-2	-3
	Ρυθμός (Επίπεδο)	-4	-2	-3
	Ένταση (Επίπεδο)	-	-	-
Arial	Τονικό ύψος (Επίπεδο)	-1	+3	+2
	Ρυθμός (Επίπεδο)	-1	+3	+2
	Ένταση (Επίπεδο)	-	-	-

Χρησιμοποιώντας 6 επίπεδα αύξησης και μείωσης προκύπτουν τα ελάχιστα και μέγιστα όρια μεταβολών (Πίνακας 16). Οι Πίνακες 17 και 18 παρουσιάζουν τις ποσοστιαίες αυξήσεις/μειώσεις και τις πραγματικές τιμές για κάθε προσωδιακή μεταβολή χρησιμοποιώντας 6 επίπεδα κβάντισης των προσωδιακών στοιχείων για το μέγεθος γραμματοσειράς και τον τρόπο γραφής. Η προσωδιακή βάση αναφοράς προσδιορίζεται ως Τονικό ύψος=110Hz, Ρυθμός=140wpm και Ένταση=0dB.

Πίνακας 16: Ελάχιστες και μέγιστες τιμές των προσωδιακών μεταβολών

Προσωδιακό στοιχείο (μονάδα μέτρησης)	Ελάχιστες – Μέγιστες Τιμές
Τονικό ύψος (Hz)	97 – 189
Ρυθμός (wpm)	112 – 210
Ένταση (dB)	-5.19 έως +4.08

Πίνακας 17: Οι προσωδιακές μεταβολές μετά την κβάντιση (ποσοστιαίες μεταβολές) και οι αντίστοιχες πραγματικές τιμές για την εναλλαγή του μεγέθους γραμματοσειράς

Εναλλαγές μεγέθους γραμματοσειράς		Ποσοστιαία Μεταβολή			Πραγματικές Τιμές		
		Τονικό ύψος (%)	Ρυθμός (%)	Ένταση (%)	Τονικό ύψος (Hz)	Ρυθμός (wpm)	Ένταση (dB)
+2pt	10→12	-12	-20	-30	96.8	112	-3.10
	12→14	12	-10	-15	123.2	126	-1.41
	14→16	24	10	-15	136.4	154	-1.41
+4pt	10→14	12	-20	-45	123.2	112	-5.19
	12→16	24	10	-15	136.4	154	-1.41
	14→18	36	30	15	149.6	182	1.21
+6pt	10→16	24	-10	-45	136.4	126	-5.19
	12→18	48	20	-15	162.8	168	-1.41
	14→20	60	40	30	176	196	2.28
+8pt	10→18	36	10	-45	149.6	154	-5.19
	12→20	60	40	15	176	196	1.21
	14→22	72	50	45	189.2	210	3.23
-1pt	10→9	12	10	30	123.2	154	2.28
	12→11	12	10	15	123.2	154	1.21
	14→13	-12	-10	15	96.8	126	1.21
-2pt	10→8	24	30	60	136.4	182	4.08
	12→10	12	20	30	123.2	168	2.28
	14→12	-12	10	15	96.8	154	1.21

Πίνακας 18: Οι προσωδιακές μεταβολές μετά τη κβάντιση (ποσοστιαίες μεταβολές) και οι αντίστοιχες πραγματικές τιμές για την εναλλαγή του τρόπου γραφής.

<i>Εναλλαγές από την απλή γραφή</i>	Έντονη	Πλάγια	Έντονη-Πλάγια
Times New Roman			
<i>Τονικό ύψος (%)</i>	-48	-24	-36
<i>Ρυθμός (%)</i>	-40	-20	-30
<i>Ένταση (%)</i>	-	-	-
<i>Τονικό ύψος (Hz)</i>	57	84	70
<i>Ρυθμός (wpm)</i>	84	112	98
<i>Ένταση (dB)</i>	-	-	-
Arial			
<i>Τονικό ύψος (%)</i>	-12	36	24
<i>Ρυθμός (%)</i>	-10	30	20
<i>Ένταση (%)</i>	-	-	-
<i>Τονικό ύψος (Hz)</i>	97	150	136
<i>Ρυθμός (wpm)</i>	126	182	168
<i>Ένταση (dB)</i>	-	-	-

4.3 Επιλογή βέλτιστης ακουστικής απόδοσης των τυπογραφικών μεταβολών

Το μοντέλο που περιγράφεται στην προηγούμενη παράγραφο, εισάγει πολλαπλούς τρόπους ακουστικοποίησης κάθε τυπογραφικής εναλλαγής (3 περιπτώσεις για κάθε μέγεθος γραμματοσειράς και 2 περιπτώσεις για κάθε εναλλαγή του τρόπου γραφής ανάλογα με τον χρησιμοποιούμενο τύπο γραμματοσειράς). Πολλές προσωδιακές μεταβολές στην ομιλία ή η υλοποίηση με διαφορετικούς τρόπους του ίδιου τυπογραφικού στοιχείου, μπορεί να μπερδέψει ακόμα και να ενοχλήσει τον ακροατή. Επίσης, στη μελέτη μας θέλουμε να δημιουργήσουμε ένα μοντέλο που μπορεί να εφαρμοσθεί σε κάθε έγγραφο, μελετώντας μόνο τις τυπογραφικές εναλλαγές και όχι τις απόλυτες τιμές τους.

Για αυτό, σχεδιάστηκε και αναπτύχθηκε μία πειραματική διαδικασία για την επιλογή της βέλτιστης ακουστικής απόδοσης των τυπογραφικών εναλλαγών (για το μέγεθος της γραμματοσειράς και τον τρόπο γραφής). Αυτή η διαδικασία θα αναδείξει τον προτιμότερο τρόπο ακουστικοποίησης των τυπογραφικών εναλλαγών, σύμφωνα με τις υλοποιήσεις που εισάγει το προσωδιακό μοντέλο.

4.3.1 Η πειραματική διαδικασία

Συμμετέχοντες

Στην πειραματική διαδικασία συμμετείχαν 28 άτομα, ηλικίας από 19 έως 33 ετών (μέση Ηλικία = 27.7 έτη, τυπική απόκλιση = 3.2). Ήταν προπτυχιακοί και μεταπτυχιακοί φοιτητές από το τμήμα Πληροφορικής και Τηλεπικοινωνιών του Πανεπιστημίου Αθηνών, 18 άνδρες και 10 γυναίκες και η μητρική τους γλώσσα ήταν ελληνική. Δεν είχαν κανένα πρόβλημα ακοής και η οπτική τους οξύτητα ήταν φυσιολογική ή διορθωμένη στο φυσιολογικό.

Ερεθίσματα

Δημιουργήθηκαν 24 ερεθίσματα στην Ελληνική γλώσσα. Κάθε ερέθισμα ήταν μια σύντομη πρόταση (διάρκειας εκφώνησης περίπου 5 δευτερολέπτων), όπου υπήρχε μία φράση στην οποία εφαρμοζόταν οι τυπογραφικές εναλλαγές δηλαδή: από απλή σε έντονη, πλάγια, έντονη-πλάγια γραφή και μέγεθος γραμματοσειράς αυξημένο κατά 2pt, 4pt, 6pt, 8pt ή μείωση κατά 1pt ή 2pt. Οι προτάσεις μετατράπηκαν σε ομιλία με τη χρήση της πλατφόρμας Document-to-Audio που βασίζεται στο σύστημα μετατροπής Κειμένου-σε-Ομιλία DEMOSTHeNES [36].

Κάθε ερέθισμα παρουσιάστηκε με οπτικό και ακουστικό τρόπο ταυτόχρονα. Πριν ξεκινήσει η πειραματική διαδικασία, οι συμμετέχοντες μπορούσαν να ακούσουν μερικά παραδείγματα με σκοπό να εξοικειωθούν με τις προσωδιακές μεταβολές. Κατά τη διάρκεια του κύριου μέρους του πειράματος, μπορούσαν να ακούσουν το κάθε ερέθισμα μόνο δύο φορές μέσα από υψηλής πιστότητας ακουστικά (AKG K171 MKII). Στη συνέχεια ερωτούνταν, π.χ. «Συμφωνείτε ότι οι λέξεις “δοκιμαστικό κείμενο” είναι σε [τυπογραφικό στοιχείο];» (όπου [τυπογραφικό στοιχείο] αναφερόταν η υλοποιημένη τυπογραφική εναλλαγή, π.χ. έντονη γραφή) έχοντας τη δυνατότητα να δώσουν την απάντησή τους σε μία κλίμακα 5 σημείων: Το 1 αντιστοιχεί στο «ΟΧΙ», το 3 στο «Δε γνωρίζω» και το 5 στο «ΝΑΙ». Οι ενδιάμεσες απαντήσεις αντιστοιχούν στο «Πιθανόν ΟΧΙ» και «Πιθανόν ΝΑΙ» αντίστοιχα. Η μέση τιμή των χρόνων απαντήσεων ήταν περίπου 18 λεπτά.

4.3.2 Αποτελέσματα

Χρησιμοποιήθηκε την ανάλυση t-test για να μελετηθεί εάν οι μέσες τιμές των απαντήσεων είναι στατιστικώς σημαντικά μεγαλύτερες από την τιμή 3 (που αντιστοιχεί στο «Δε γνωρίζω» και $p < 0.05$). Στον Πίνακα 19 παρουσιάζονται τα αποτελέσματα της στατιστικής ανάλυσης. Η στήλη με πιθανότητα p υποδηλώνει τη σημαντικότητα των αποτελεσμάτων. Οι τιμές σε έντονη γραφή αντιστοιχούν σε στατιστικώς σημαντικά αποτελέσματα.

4.3.3 Σύνοψη των αποτελεσμάτων

Από τον Πίνακα 19, παρατηρούμε ότι οι εναλλαγές του τρόπου γραφής και του μεγέθους γραμματοσειράς με αύξηση κατά 2pt ή με μείωση κατά 1pt είναι στατιστικώς σημαντικά.

Επομένως προκύπτει ότι οι ακροατές προτίμησαν:

- την έντονη γραφή για την περίπτωση του τύπου γραμματοσειράς «Arial»,
- την πλάγια γραφή για την περίπτωση του τύπου γραμματοσειράς «Arial»,
- την έντονη-πλάγια γραφή για την περίπτωση του τύπου γραμματοσειράς «Times New Roman»,
- την αύξηση του μεγέθους γραμματοσειράς κατά 2pt με τυπογραφική βάση αναφοράς τα 12pt,
- τη μείωση του μεγέθους γραμματοσειράς κατά 1pt με τυπογραφική βάση αναφοράς τα 10pt.

Ο Πίνακας 20 συνοψίζει τα παραπάνω αποτελέσματα παρουσιάζοντας τις προσωδιακές μεταβολές σε επίπεδα, ποσοστά και απόλυτες τιμές.

Πίνακας 19: Τα στατιστικά αποτελέσματα της ανάλυσης των αποτελεσμάτων από την πειραματική διαδικασία. Στις παρενθέσεις αναφέρεται ο τύπος γραμματοσειράς (TNR – Times New Roman και Arial).

Εναλλαγή Τρόπου Γραφής		Μέση τιμή	Τυπικό Σφάλμα	p
Έντονη γραφή	ερέθισμα 1 (TNR)	2.96429	1.52709	$p>0.05$
	ερέθισμα 2 (Arial)	3.5	1.47824	$p<0.05$
Πλάγια γραφή	ερέθισμα 1 (TNR)	3.07143	1.30323	$p>0.05$
	ερέθισμα 2 (Arial)	3.53571	1.40059	$p<0.05$
Έντονη-πλάγια γραφή	ερέθισμα 1 (TNR)	3.57143	1.50132	$p<0.05$
	ερέθισμα 2 (Arial)	2.67857	1.46701	$p>0.05$
Μέγεθος Γραμματοσειράς				
Μέγεθος +2	ερέθισμα 1	3.32143	1.36228	$p>0.05$
	ερέθισμα 2	3.92857	1.01575	$p<0.05$
	ερέθισμα 3	3.35714	1.22366	$p>0.05$
Μέγεθος +4	ερέθισμα 1	2.85714	1.1455	$p>0.05$
	ερέθισμα 2	2.75	1.43049	$p>0.05$
	ερέθισμα 3	2.39286	1.42307	$p>0.05$
Μέγεθος +6	ερέθισμα 1	2.35714	1.33927	$p>0.05$
	ερέθισμα 2	2.35714	1.47106	$p>0.05$
	ερέθισμα 3	2.57143	1.47645	$p>0.05$
Μέγεθος +8	ερέθισμα 1	1.89286	1.31485	$p>0.05$
	ερέθισμα 2	2.14286	1.45842	$p>0.05$
	ερέθισμα 3	2.21429	1.54817	$p>0.05$
Μέγεθος -1	ερέθισμα 1	3.46429	1.23175	$p<0.05$
	ερέθισμα 2	2.89286	1.19689	$p>0.05$
	ερέθισμα 3	2.5	1.55158	$p>0.05$
Μέγεθος -2	ερέθισμα 1	3	1.33333	$p>0.05$
	ερέθισμα 2	3.21429	1.25778	$p>0.05$
	ερέθισμα 3	3.35714	1.19301	$p>0.05$

Πίνακας 20: Τα στατιστικώς σημαντικά αποτελέσματα μαζί με τα επίπεδα αύξησης/μείωσης, με τις αντίστοιχες ποσοστιαίες μεταβολές και τις απόλυτες τιμές των προσωδιακών χαρακτηριστικών

	Εναλλαγή από απλό τρόπο γραφής			Αύξηση Μεγέθους Γραμματοσειράς	
	Έντονη (Arial)	Πλάγια (Arial)	Έντονη-Πλάγια (Times New Roman)	Μέγεθος +2 (12→14)	Μέγεθος -1 (10→9)
Τονικό ύψος (επίπεδο)	-1	+3	-3	+1	+1
Ρυθμός (επίπεδο)	-1	+3	-3	-1	+1
Ένταση (επίπεδο)	-	-	-	-1	+2
Τονικό ύψος (%)	-12	36	-36	12	12
Ρυθμός (%)	-10	30	-30	-10	10
Ένταση (%)	-	-	-	-15	30
Τονικό ύψος (Hz)	97	150	70	123	123
Ρυθμός (wpm)	126	182	98	126	154
Ένταση (db)	-	-	-	-1.41	2.28

4.4 Συμπεράσματα

Χρησιμοποιώντας το μοντέλο της μεταβολής της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη ως προς τις τυπογραφικές εναλλαγές, που παρουσιάζεται στο Κεφάλαιο 3, υπολογίζονται οι μεταβολές της επαγόμενης συναισθηματικής κατάστασης. Συνδυάζοντας αυτά τα αποτελέσματα με το προσωδιακό μοντέλο εκφραστικής συνθετικής ομιλίας του Schröder [142], δίνεται η δυνατότητα της ακουστικής αναπαράστασης (τονικό ύψος, ρυθμός και ένταση ομιλίας) των τυπογραφικών εναλλαγών. Οι μεταβολές που προκύπτουν είναι κάτω από το κατώφλι της ελάχιστης ακουστικής αντίληψης. Για αυτό το λόγο εισάγεται η κανονικοποίησή τους μέσω της γραμμικής κβάντισης των προσωδιακών μεταβολών και η αντιστοίχηση σε κάθε επίπεδο της ελάχιστης αντιληπτής τιμής για κάθε προσωδιακό χαρακτηριστικό.

Το μοντέλο που προκύπτει θέτει πολλαπλούς τρόπους ακουστικής απόδοσης του τρόπου γραφής και του μεγέθους γραμματοσειράς. Για αυτό το λόγο διεξήχθη ένα πείραμα για την επιλογή της βέλτιστης ακουστικής απόδοσης της εναλλαγής του τρόπου γραφής και μεγέθους γραμματοσειράς. Οι ακροατές επέλεξαν τα εξής:

- Η χρήση έντονης γραφής σε γραμματοσειρά «Arial» μειώνει το τονικό ύψος κατά 12% και το ρυθμό κατά 10%,

- Η πλάγια γραφή σε «Arial» αυξάνει το τονικό ύψος και το ρυθμό κατά 36% και 30% αντίστοιχα.
- Η έντονη-πλάγια γραφή σε «Times New Roman» μειώνει το τονικό ύψος και το ρυθμό κατά 36% και 30%, αντίστοιχα.
- Η μείωση του μεγέθους κατά 2pt αυξάνει το τονικό ύψος κατά 12% και μειώνει το ρυθμό και την ένταση κατά 10% και 15% αντίστοιχα.
- Η μείωση του μεγέθους κατά 1pt αυξάνει και τα τρία προσωδιακά χαρακτηριστικά κατά 12%, 10% και 30% (τονικό ύψος, ρυθμό και ένταση αντίστοιχα).

Παρατηρείται ότι μολονότι το αρχικό μοντέλο καταδεικνύει διάφορους τρόπους υλοποίησης εναλλαγών του μεγέθους γραμματοσειράς, μόνο δύο μεταβολές αξιολογήθηκαν ως αποδεκτές. Αυτό πιθανόν οφείλεται στις οριακά διακριτές μεταβολές της έντασης.

5. ΑΞΙΟΛΟΓΗΣΗ ΤΟΥ ΑΚΟΥΣΤΙΚΟΥ ΜΟΝΤΕΛΟΥ

5.1 Εισαγωγή

Η αξιολόγηση του προσωδιακού μοντέλου πραγματοποιήθηκε μέσω δύο πειραματικών διαδικασιών. Στην πρώτη εξετάστηκε το κατά πόσο οι ακροατές είναι σε θέση, μέσω των προσωδιακών μεταβολών, να αναγνωρίσουν τις τυπογραφικές εναλλαγές (από απλό κείμενο σε έντονη, σε πλάγια και σε έντονη-πλάγια γραφή καθώς επίσης και την αύξηση ή μείωση του μεγέθους γραμματοσειράς) και στη δεύτερη διαδικασία το κατά πόσο η χρήση του μοντέλου που αναπτύχθηκε στο Κεφάλαιο 4 μπορεί να βελτιώσει την απόδοση των τυφλών και βλεπόντων μαθητών κατά τη διδακτική διαδικασία. Η δεύτερη μελέτη διεξήχθη στο πλαίσιο της διδακτορικής διατριβής της Καλλιόπης Εικοσπεντάκη και κατά τη συνεργασία και συμμετοχή στο έργο ΠΕΝΕΔ-2003 (στο οποίο συμμετέχει και η παρούσα διατριβή) [209].

5.2 Ακουστική αναγνώριση των τυπογραφικών στοιχείων

Σκοπός της παρούσας πειραματικής διαδικασίας είναι να αξιολογηθεί κατά πόσο οι ακροατές μπορούν να αναγνωρίσουν τις τυπογραφικές εναλλαγές σε μία πρόταση, χρησιμοποιώντας το προτεινόμενο προσωδιακό μοντέλο που περιγράφεται στο προηγούμενο κεφάλαιο. Αυτή θα δείξει κατά πόσο μπορεί η προτεινόμενη μεθοδολογία να αντιστοιχήσει τις τυπογραφικές μεταβολές από το οπτικό στο ακουστικό επίπεδο. Να τονιστεί σε αυτό το σημείο ότι θέλουμε να μελετηθεί κατά πόσο αντιλαμβάνονται οι ακροατές τις εναλλαγές αυτές, *χωρίς προηγουμένως να έχουν εκπαιδευτεί και εξοικειωθεί σε αυτές.*

5.2.1 Η μεθοδολογία

Συμμετέχοντες

Στη διαδικασία συμμετείχαν 11 άνδρες και 8 γυναίκες (συνολικά 19 συμμετέχοντες) – ηλικίας από 20 έως 32 ετών (μέση ηλικία = 27.8 έτη, τυπική απόκλιση = 3.2). Ήταν προπτυχιακοί και μεταπτυχιακοί φοιτητές από το Τμήμα Πληροφορικής και Τηλεπικοινωνιών του Πανεπιστημίου Αθηνών, η μητρική τους γλώσσα είναι η Ελληνική και δεν έχουν καμία εμπλοκή σε προηγούμενο πείραμα που να σχετίζεται με τη μελέτη μας.

Ερεθίσματα

Συνολικά, χρησιμοποιήθηκαν 20 ερεθίσματα στην Ελληνική γλώσσα, 12 για τον τρόπο γραφής και 8 για το μέγεθος της γραμματοσειράς. Κάθε ερέθισμα περιείχε μία πρόταση (διάρκειας περίπου 5 δευτερολέπτων) ή μία φράση της πρότασης (με 1, 2 ή 3 λέξεις) η οποία είχε αποδοθεί με τη χρήση συγκεκριμένου τυπογραφικού στοιχείου (στον Πίνακα 21, για κάθε υλοποίηση των τυπογραφικών εναλλαγών έχει οριστεί το όνομα «ερέθισμα 1», «ερέθισμα 2» κτλ). Οι προτάσεις μετατράπηκαν σε ομιλία χρησιμοποιώντας την πλατφόρμα Document-to-Audio [36].

5.2.2 Η πειραματική διαδικασία

Δεν υπήρχε δοκιμαστική διαδικασία ώστε οι συμμετέχοντες να εξοικειωθούν με τις προσωδιακές μεταβολές. Ο καθένας συμμετείχε σε δύο μέρη: αξιολόγηση της ακουστικής απόδοσης της εναλλαγής α) του τρόπου γραφής και β) του μεγέθους γραμματοσειράς, αντίστοιχα.

Παράλληλα με τα ακουστικά ερεθίσματα, γινόταν προβολή και της πρότασης χωρίς τα τυπογραφικά στοιχεία (η πρόταση ή η φράση ήταν υπογραμμισμένη με σκοπό να υποδεικνύεται το μέρος της πρότασης ή ολόκληρη η πρόταση στους συμμετέχοντες). Ρωτήθηκαν ποιο τυπογραφικό στοιχείο πιστεύουν ότι συσχετίζεται με την πρόταση ή φράση (τρόπος γραφής: έντονη, πλάγια, έντονη-πλάγια ή καθόλου / μέγεθος γραμματοσειράς: αύξηση - μείωση μεγέθους, ή καθόλου).

Η Ευρώπη θα εγκαταστήσει βάση στη Σελήνη πριν από τη ΝΑΣΑ

Σχήμα 29: Παράδειγμα ερεθίσματος, όπως παρουσιαζόταν οπτικά κατά την πειραματική διαδικασία

5.2.3 Τα αποτελέσματα

Τα αποτελέσματα (Πίνακας 21 και 22, η οπτική τους απεικόνιση μέσω διαγραμμάτων στα Σχήματα 30 και 31, αντίστοιχα) υποδεικνύουν ότι οι συμμετέχοντες αναγνώρισαν

εύκολα της εναλλαγές του τρόπου γραφής και του μεγέθους γραμματοσειράς. Στους Πίνακες 21 και 22, με έντονη γραφή φαίνονται οι μέγιστες τιμές για κάθε μία περίπτωση. Οι εναλλαγές του μεγέθους έγιναν αντιληπτές καλύτερα από ότι ο τρόπος γραφής, δηλαδή:

- Η έντονη-πλάγια γραφή δεν αναγνωρίστηκε από τους συμμετέχοντες και συγχέεται με την έντονη (σχεδόν όλες οι περιπτώσεις αξιολογήθηκαν ως έντονη γραφή από την πλειοψηφία των συμμετεχόντων). Η πλάγια γραφή αναγνωρίστηκε επιτυχώς.
- Οι εναλλαγές του μεγέθους γραμματοσειράς αναγνωρίστηκαν επιτυχώς. Στην πρώτη περίπτωση για την αύξηση του μεγέθους, οι περισσότερες απαντήσεις (42.1%) δήλωναν μειούμενο μέγεθος (οι σωστές απαντήσεις ήταν οριακά χαμηλότερες 36.8%). Εκτιμάται ότι οι λάθος απαντήσεις οφείλονταν στην υλοποίηση του ερεθίσματος (για παράδειγμα, η ποιότητα της συνθετικής ομιλίας για την συγκεκριμένη πρόταση).

Βασισμένη στη μετρική «precision and recall», οι Πίνακες 21 και 22 παρουσιάζουν τα αποτελέσματά της για τις μέσες τιμές, των εναλλαγών του τρόπου γραφής και της αύξησης/μείωσης του μεγέθους γραμματοσειράς.

Η ακρίβεια (Precision) προκύπτει αθροίζοντας τις σωστές απαντήσεις για το συγκεκριμένο ερέθισμα προς το άθροισμα των σωστών και των λάθος απαντήσεων για αυτό. Η ανάκληση (Recall) προκύπτει αθροίζοντας τις σωστές απαντήσεις του ερεθίσματος προς το άθροισμα των απαντήσεων που αξιολογήθηκαν ως το αναμενόμενο ερέθισμα.

Η F-measure υπολογίζεται από τη σχέση:

$$F - measure = \frac{2 * Precision * Recall}{Precision + Recall}$$

Για παράδειγμα:

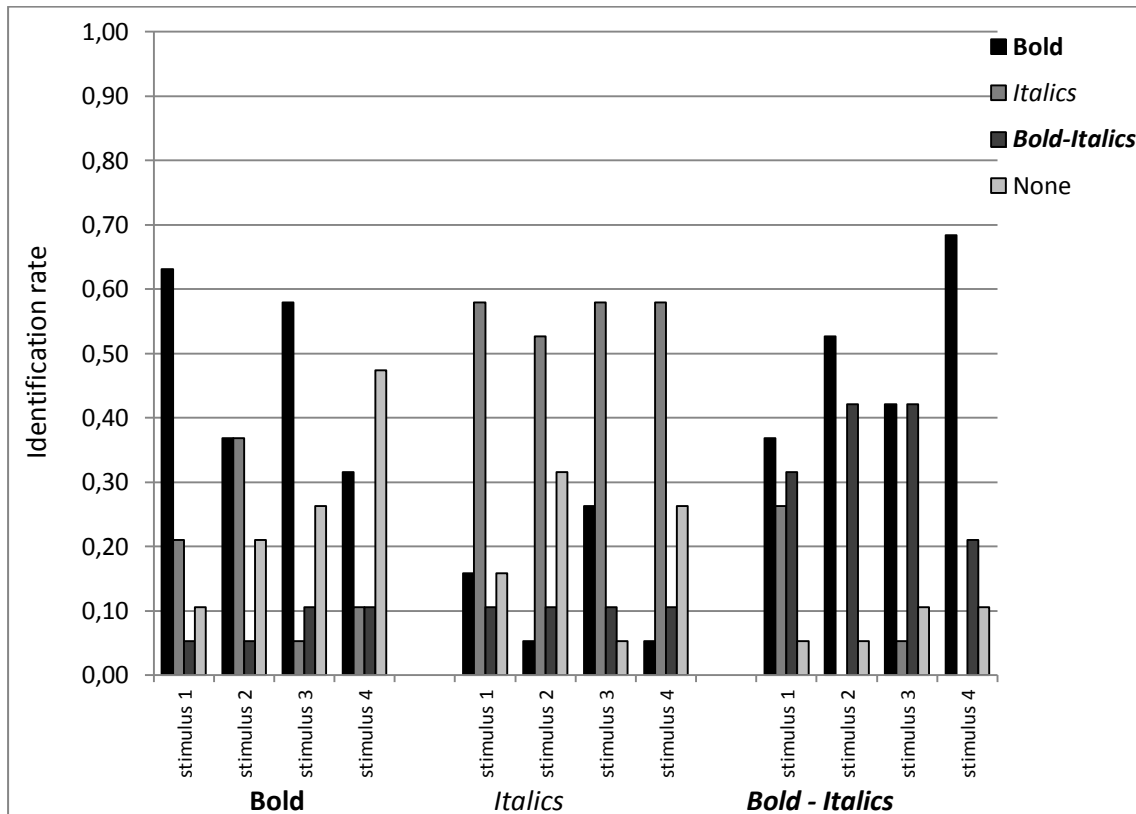
Από τον Πίνακα 21 η ακρίβεια για την πλάγια γραφή προκύπτει $P = \frac{2,263}{4} = 0,566$, η ανάκληση από $R = \frac{2,263}{0,737 + 2,263 + 0,316} = 0,683$ και f-measure $f = \frac{2 * 0,566 * 0,683}{0,566 + 0,683} = 0,619$.

Πίνακας 21: Αποτελέσματα για την ακουστικοποίηση της εναλλαγής του τρόπου γραφής.

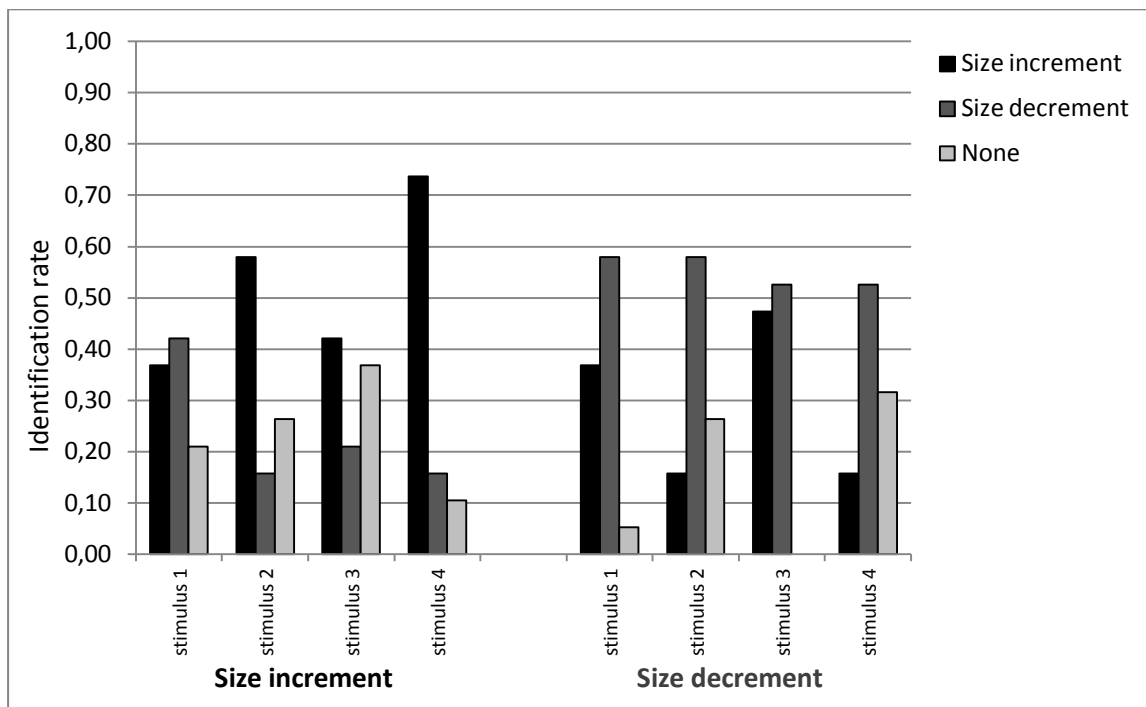
Ερεθίσματα (εναλλαγή από απλή γραφή)		Πιθανές απαντήσεις				P	R	f
		Έντονη Γραφή	Πλάγια Γραφή	Έντονη-Πλάγια Γραφή	Κανένα			
Έντονη γραφή	ερέθισμα 1	0.632	0.211	0.053	0.105	0.474	0.429	0.450
	ερέθισμα 2	0.368	0.368	0.053	0.211			
	ερέθισμα 3	0.579	0.053	0.105	0.263			
	ερέθισμα 4	0.316	0.105	0.105	0.474			
Πλάγια γραφή	ερέθισμα 1	0.158	0.579	0.105	0.158	0.566	0.683	0.619
	ερέθισμα 2	0.053	0.526	0.105	0.316			
	ερέθισμα 3	0.263	0.579	0.105	0.053			
	ερέθισμα 4	0.053	0.579	0.105	0.263			
Έντονη -Πλάγια γραφή	ερέθισμα 1	0.368	0.263	0.316	0.053	0.342	0.650	0.448
	ερέθισμα 2	0.526	0.000	0.421	0.053			
	ερέθισμα 3	0.421	0.053	0.421	0.105			
	ερέθισμα 4	0.684	0.000	0.211	0.105			

Πίνακας 22: Αποτελέσματα για την ακουστικοποίηση της εναλλαγής του μεγέθους της γραμματοσειράς.

Ερεθίσματα		Πιθανές απαντήσεις			P	R	f
		Αύξηση μεγέθους	Μείωση μεγέθους	Κανένα			
Αύξηση μεγέθους	ερέθισμα 1	0.368	0.421	0.211	0.526	0.645	0.580
	ερέθισμα 2	0.579	0.158	0.263			
	ερέθισμα 3	0.421	0.211	0.368			
	ερέθισμα 4	0.737	0.158	0.105			
Μείωση μεγέθους	ερέθισμα 1	0.368	0.579	0.053	0.553	0.700	0.618
	ερέθισμα 2	0.158	0.579	0.263			
	ερέθισμα 3	0.474	0.526	0.000			
	ερέθισμα 4	0.158	0.526	0.316			



Σχήμα 30: Βαθμός αναγνώρισης για την εναλλαγή του τρόπου γραφής.



Σχήμα 31: Βαθμός αναγνώρισης για τη μείωση/αύξηση του μεγέθους γραμματοσειράς.

5.3 Εφαρμογή του ακουστικού μοντέλου στη διδακτική διαδικασία για τυφλούς και βλέποντες μαθητές

Τα παιδιά με απώλεια όρασης αντιμετωπίζουν πολύ περισσότερες δυσκολίες στην κατανόηση επιστημονικών εννοιών από τα βλέποντα παιδιά που φοιτούν στις ίδιες σχολικές τάξεις [210]. Τα συγκεκριμένα ευρήματα υποδηλώνουν ότι παρά την πλήρη έλλειψη οπτικών εμπειριών, τα παιδιά με απώλεια όρασης προσλαμβάνουν αρκετές πληροφορίες μέσω των άλλων αισθήσεων τους, καθώς και από τις σχετικές πολιτισμικές πληροφορίες, ώστε να σχηματίσουν κάποιου είδους νοητικές αναπαραστάσεις του φυσικού κόσμου. Οι νοητικές αυτές αναπαραστάσεις αποτελούν το έναυσμα για τη δημιουργία παρανοήσεων που στέκονται εμπόδιο στην κατανόηση των επιστημονικών εννοιών, όπως ακριβώς συμβαίνει και με τα βλέποντα παιδιά [211] [212].

Είναι γεγονός ότι η παραδοσιακή διδασκαλία δε βοηθά τα παιδιά (βλέποντα και με απώλεια όρασης) να κατανοήσουν τις επιστημονικές έννοιες με σαφήνεια [210]. Αυτό συμβαίνει κυρίως επειδή αγνοείται η προοπτική των παιδιών, με αποτέλεσμα οι νέες επιστημονικές πληροφορίες που προστίθενται στις ήδη υπάρχουσες γνώσεις, να τα επηρεάζουν τόσο ώστε να κατασκευάζουν τελικά συνθετικές έννοιες. Αυτές, όμως, στέκονται τελικά εμπόδιο στην πλήρη κατανόηση και μεταγνώση των επιστημονικών εξηγήσεων των φυσικών φαινομένων.

5.3.1 Μεθοδολογία

Με βάση τα παραπάνω, γίνεται η υπόθεση ότι η υλοποίηση μιας διαφορετικής διδακτικής παρέμβασης που θα περιελάμβανε ακουστικοποίηση των τυπογραφικών μεταδεδομένων ενός εγγράφου θα βοηθούσε τα παιδιά να κατανοήσουν με μεγαλύτερη ευκολία τις επιστημονικές εξηγήσεις π.χ. των φυσικών φαινομένων. Κατά το σχεδιασμό της θα πρέπει να ληφθούν υπόψη οι δυσκολίες και οι ιδιαιτερότητες των παιδιών, καθώς και η νοητική διαδικασία που τίθεται σε εφαρμογή για την κατανόηση των επιστημονικών εννοιών. Για να επιβεβαιωθεί ή διαψευθεί η υπόθεση, υλοποιήθηκε ένα σενάριο διδασκαλίας υπό μορφή διδακτικής παρέμβασης, στην οποία συμμετείχαν και οι δύο ομάδες μαθητών – βλέποντες και εκ γενετής τυφλοί.

Συμμετέχοντες

Οι ομάδες αποτελούνταν από δέκα εκ γενετής τυφλούς και δέκα βλέποντες μαθητές της Στ' τάξης του Δημοτικού σχολείου, με μέσο όρο ηλικίας τα 11 έτη και 6 μήνες. Στο σημείο αυτό πρέπει να σημειωθεί ότι οι συγκεκριμένοι μαθητές είχαν συμμετάσχει

αρκετό καιρό πριν τη διδακτική παρέμβαση σε μια ερευνητική διαδικασία. Είχε πραγματοποιηθεί προσωπική συνέντευξη με κάθε μαθητή, χρησιμοποιώντας ένα ερωτηματολόγιο που περιελάμβανε διερευνητικές ερωτήσεις. Αφορούσαν τις γνώσεις τους και την κατανόηση των επιστημονικών εννοιών της Παρατηρησιακής Αστρονομίας, εφόσον τις είχαν διδαχθεί στα σχολεία τους με την παραδοσιακή διδασκαλία. Η συγκεκριμένη ερευνητική διαδικασία χρησιμοποιήθηκε ως pre-test.

Ερεθίσματα

Η παρέμβαση πραγματοποιήθηκε με τη χρήση ενός ηλεκτρονικού διαδραστικού πίνακα (Smart Board) [213], μέσω του οποίου γινόταν η προβολή των οπτικών ερεθισμάτων με τη χρήση διαφανειών του MS PowerPoint. Επίσης, το περιεχόμενο κάθε διαφάνειας μετατράπηκε σε ομιλία με τη χρήση του ακουστικού μοντέλου που αναπτύχθηκε στο προηγούμενο κεφάλαιο και της πλατφόρμας μετατροπής Εγγράφου-σε-Ακουστική μορφή που βασίζεται στο σύστημα μετατροπής Κειμένου-σε-Ομιλία DEMOSTHeNES [36]. Τα τυπογραφικά στοιχεία που χρησιμοποιήθηκαν ήταν η έντονη γραφή (για τις επιστημονικές έννοιες) και η αύξηση μεγέθους γραμματοσειράς κατά 2pt (για τους τίτλους). Το περιεχόμενο ήταν ένα ολοκληρωμένο διδακτέο κείμενο του μαθήματος της Παρατηρησιακής Αστρονομίας που διδάσκεται στο Δημοτικό Σχολείο.

Σε πραγματικό χρόνο η ακουστική αναπαράσταση των τυπογραφικών στοιχείων του κειμένου συνδέθηκε με την παρουσίαση του ανάλογου οπτικού υλικού (εικόνες). Σε αυτό είχαν πρόσβαση οι βλέποντες μαθητές και γινόταν λεπτομερής περιγραφή στους εκ γενετής τυφλούς.

5.3.2 Πειραματική διαδικασία

Η διδακτική παρέμβαση διήρκησε περίπου όσο και μια διδακτική ώρα (45-50 λεπτά). Στη συνέχεια ακολούθησε προσωπική συνέντευξη με κάθε μαθητή, εκ γενετής τυφλό και βλέποντα, βασισμένη σε ένα ερωτηματολόγιο με ερωτήσεις κατάλληλες, ώστε να μπορεί να διερευνηθεί το επίπεδο κατανόησης των επιστημονικών εννοιών. Η προσωπική συνέντευξη διήρκησε περίπου 15 λεπτά. Μέσω της διαδικασίας των συνεντεύξεων καταγράφηκε ο βαθμός ακουστικής αντίληψης των διαφορετικών μεταπληροφοριών, καθώς και ο βαθμός κατανόησης των εννοιολογικών και σημασιολογικών πληροφοριών.

5.3.3 Αποτελέσματα

Πραγματοποιήθηκαν στατιστικές αναλύσεις σε δύο άξονες για να αξιολογηθούν οι απαντήσεις των παιδιών:

1. Ο πρώτος άξονας αφορά στη μετατροπή της οπτικής πληροφορίας σε ακουστική.
2. Ο δεύτερος άξονας αφορά στη μετατροπή της οπτικής πληροφορίας σε ακουστική σε συνδυασμό με την οπτική πληροφορία που μπορεί να αποκτήσει ο μαθητής (με τη χρήση του διαδραστικού πίνακα).

Από τα αποτελέσματα [209] προκύπτει το συμπέρασμα ότι μια τέτοιου είδους πολυπαραγοντική διδακτική παρέμβαση που περιλαμβάνει ακουστικοποίηση των τυπογραφικών στοιχείων ενός εγγράφου βοήθησε τα παιδιά να κατανοήσουν τις επιστημονικές έννοιες της Παρατηρησιακής Αστρονομίας και τις επιστημονικές εξηγήσεις κάποιων φαινομένων. Οι επιδόσεις τους ήταν σαφώς καλύτερες σε σχέση με αυτές στη πρώτη ερευνητική διαδικασία [209] [210].

5.4 Συμπεράσματα

Σε αυτό το κεφάλαιο παρουσιάστηκε η αξιολόγηση του ακουστικού μοντέλου που προτείνεται στο Κεφάλαιο 4. Γενικά παρατηρούμε ότι τα αποτελέσματα είναι ενθαρρυντικά. Η εναλλαγή των τυπογραφικών στοιχείων και του μεγέθους γραμματοσειράς αναγνωρίστηκε από τους ακροατές όταν:

1. η εναλλαγή από απλό κείμενο σε έντονη γραφή μεταβάλλει το τονικό ύψος (-12%) και το ρυθμό ομιλίας (-10%),
2. η εναλλαγή από απλό κείμενο σε πλάγια γραφή μεταβάλλει το τονικό ύψος και το ρυθμό (αυξάνοντάς τα κατά 36% και 30% αντίστοιχα),
3. αυξάνοντας το μέγεθος γραμματοσειράς (+2pt) αυξάνεται το τονικό ύψος (+12%), μειώνεται ο ρυθμός (-10%) και η ένταση (-15%),
4. μειώνοντας το μέγεθος γραμματοσειράς (-1pt) αυξάνεται το τονικό ύψος (+12%), ο ρυθμός (+10%) και η ένταση (+30%).

Τα αποτελέσματα έδειξαν την έλλειψη ορθής αντίληψης της ακουστικής απεικόνισης για την έντονη-πλάγια γραφή. Αυτό προφανώς οφείλεται στην ομοιότητα των προσωδιακών μεταβολών και για τις δύο περιπτώσεις, έντονης και έντονης-πλάγιας γραφής (μείωση του τονικού ύψους και του ρυθμού). Να τονίσουμε πάλι σε αυτό το σημείο ότι τα αποτελέσματα αποτελούν τη μελέτη της αντίληψης των εναλλαγών των τυπογραφικών

χαρακτηριστικών με ακουστικό τρόπο χωρίς οι συμμετέχοντες να έχουν εξοικειωθεί ή εκπαιδευτεί σε αυτές. Τέλος, στο δεύτερο πείραμα (διδακτική διαδικασία για τυφλούς και βλέποντες μαθητές) παρουσιάζεται σημαντική βελτίωση της απόδοσης των μαθητών σε σχέση με την παραδοσιακή διδασκαλία όταν χρησιμοποιείται ακουστικοποίηση των τυπογραφικών στοιχείων στα έγγραφα.

6. ΣΥΜΠΕΡΑΣΜΑΤΑ – ΜΕΛΛΟΝΤΙΚΗ ΕΡΕΥΝΑ

6.1 Συμπεράσματα

Σε αυτή τη διατριβή εξετάζεται το πρόβλημα της ακουστικοποίησης για τυφλούς και βλέποντες των τυπογραφικών στοιχείων των δομημένων εγγράφων κατά τη μετατροπή τους σε ομιλία. Η προσέγγιση του προβλήματος χωρίζεται σε δύο μέρη: α) την αυτόματη εξαγωγή των επαγόμενων συναισθημάτων – συναισθηματικής κατάστασης του αναγνώστη από τα τυπογραφικά στοιχεία του εγγράφου και β) την ακουστική απόδοσή τους χρησιμοποιώντας εκφραστική συναισθηματική ομιλία.

Καταρχήν, διατυπώνεται το γενικότερο πρόβλημα που πραγματεύεται η διατριβή. Οι άνθρωποι με εντυποαναπηρία, αλλά και οι μετακινούμενοι χρήστες ή οι ηλικιωμένοι, πολλές φορές προτιμούν να έχουν πρόσβαση στα έντυπα και ηλεκτρονικά έγγραφα με εναλλακτικό τρόπο, όπως π.χ. απτικά με χρήση της γραφής Braille, ακουστικά μέσω μετατροπής κειμένου σε ομιλία, και μέσω μεγεθυμένης εκτύπωσης ή μεγεθυμένης προβολής τους. Τα συστήματα μετατροπής Κειμένου-σε-Ομιλία (Text-to-Speech) μετατρέπουν σε πραγματικό χρόνο κάθε ηλεκτρονικό κείμενο σε ομιλία. Όμως, σχεδόν όλα δεν υποστηρίζουν την αποτελεσματική ακουστική απόδοση πληροφοριών για τα οπτικά (π.χ. τρόπος γραφής και μέγεθος γραμματοσειράς) και μη-οπτικά στοιχεία (π.χ. έμφαση). Έτσι, αυτή η επιπλέον πληροφορία «χάνεται» κατά την επεξεργασία του εγγράφου και την ακουστική απόδοσή του. Οι προσπάθειες που έχουν γίνει μέχρι σήμερα προς αυτή την κατεύθυνση, βασίζονται στην εφαρμογή συγκεκριμένων κανόνων ακουστικοποίησης της μεταπληροφορίας των εγγράφων και είναι δύσκολη η υλοποίηση ενός αυτοματοποιημένου συστήματος - ανεξάρτητου γλώσσας, περιεχομένου ή/και πολιτισμού - για την ακουστική απόδοση των τυπογραφικών στοιχείων οποιουδήποτε εγγράφου.

Παρουσιάζεται η θεωρητική ανάλυση της έννοιας του εγγράφου και προτείνεται η διάκριση των στοιχείων που το απαρτίζουν, μέσω μιας αρχιτεκτονικής τριών επιπέδων: λογικό, διάταξης και τυπογραφικό. Η θεωρητική ανάλυση οδηγεί στην επιλογή των τυπογραφικών στοιχείων που θα μελετηθούν. Παρουσιάζονται βασικές έννοιες σχετικές με την προσβασιμότητα των εγγράφων και οι διεθνείς τυποποιήσεις για ηλεκτρονικά προσβάσιμα έγγραφα. Γίνεται θεωρητική επισκόπηση του επιστημονικού πεδίου των συναισθημάτων με σκοπό να καταστεί κατανοητός ο τρόπος μελέτης τους στα έγγραφα, στη συνθετική ομιλία, καθώς και στην προτεινόμενη μεθοδολογία.

Προτείνεται η αρχιτεκτονική ενός πρωτότυπου συστήματος βασισμένο στην τεχνολογία XML για την πολυτροπική παραγωγή, παρουσίαση και πλοήγηση σε πραγματικό χρόνο σε προσβάσιμα έγγραφα ακολουθώντας τις οδηγίες και τα πρότυπα που συζητήθηκαν στη θεωρητική ανάλυση. Αυτή η προσέγγιση περιλαμβάνει μία ενιαία μεθοδολογία για την πολυτροπική απόδοση των τριών επιπέδων των εγγράφων μέσω: α) σημασιολογικού ή β) συναισθηματικού επιπέδου. Το σημασιολογικό επίπεδο στοχεύει στην καταγραφή της κατανόησης του εγγράφου από τον αναγνώστη. Το συναισθηματικό επίπεδο κωδικοποιεί τη συναισθηματική κατάσταση του ατόμου κατά τη διαδικασία ανάγνωσης. Τα δεδομένα εξόδου αυτών των δύο επιπέδων μπορούν να μεταφερθούν με κάθε αισθητηριακό τρόπο, αλλά το ενδιαφέρον επικεντρώνεται στον ακουστικό.

Για την ακουστική απόδοση των τυπογραφικών εναλλαγών, εισάγεται ένας έμμεσος τρόπος αντιστοίχισής τους. Αυτό επιτυγχάνεται με το συνδυασμό της μοντελοποίησης της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη και την εκφραστική συνθετική ομιλία. Προτείνεται ένα γενικό μοντέλο, βασισμένο στη μαθηματική μοντελοποίηση της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη από τις εναλλαγές των τυπογραφικών στοιχείων. Έτσι, μέσω της εκφραστικής συνθετικής ομιλίας, που στηρίζεται στη διαστατική φύση των συναισθημάτων, μπορεί να επιτευχθεί η αντιστοίχιση των τυπογραφικών εναλλαγών σε προσωδιακές μεταβολές της ομιλίας.

Σε πρώτο στάδιο μέσω θεωρητικής, πειραματικής και στατιστικής μελέτης - ανάλυσης, προτείνεται μια μαθηματική περιγραφή του τρόπου με τον οποίο μεταβάλλεται η επαγόμενη συναισθηματική κατάσταση του αναγνώστη κατά τις τυπογραφικές εναλλαγές. Σκοπός είναι η δημιουργία ενός αυτόματου συστήματος εξαγωγής της επαγόμενης συναισθηματικής κατάστασης και της κατάλληλης επισημείωσης των εγγράφων με αυτή την πληροφορία. Πιο συγκεκριμένα, περιγράφεται η επίδραση των μεταβολών του μεγέθους της γραμματοσειράς στις διαστάσεις «Ευχαρίστηση», «Διέγερση» και «Επικράτηση», χρησιμοποιώντας εξισώσεις δευτέρου και τρίτου βαθμού. Για την περίπτωση, του τρόπου γραφής, δημιουργείται ένα σύνολο κανόνων για την αντιστοίχιση των τυπογραφικών μεταβολών από το απλό κείμενο. Η πρωτοτυπία αυτών των αποτελεσμάτων είναι η ποσοτική περιγραφή των εναλλαγών, σε αντίθεση με τις υπάρχουσες μελέτες που απλώς παρέχουν οδηγίες για τη βέλτιστη χρήση των τυπογραφικών στοιχείων.

Πιο συγκεκριμένα, καθώς αυξάνεται το μέγεθος γραμματοσειράς μέχρι 26px, αυξάνεται η συναισθηματική διάσταση «Ευχαρίστηση». Από αυτό το μέγεθος και πάνω υπάρχει

μείωση. Αντίθετα, οι διαστάσεις «Διέγερση» και «Επικράτηση» μειώνονται μέχρι το μέγεθος 15px και 18px, αντίστοιχα. Πέρα από αυτά τα μεγέθη αυξάνονται και οι δύο διαστάσεις. Τα αποτελέσματα επίσης δείχνουν ότι μόνο η διάσταση «Ευχαρίστηση» επηρεάζεται από τον τύπο γραμματοσειράς. Ο τύπος γραμματοσειράς «Times New Roman» θεωρείται ως πιο ευχάριστη γραμματοσειρά από ότι ο «Arial» (23% αύξηση της «Ευχαρίστηση» χρησιμοποιώντας απλό κείμενο). Τέλος, μόνο η διάσταση «Ευχαρίστηση» επηρεάζεται από τον τρόπο γραφής.

Στη συνέχεια, παρουσιάζεται η υλοποίηση του πρωτότυπου συστήματος λογισμικού DocEmoX (Document Emotion eXtractor), το οποίο, ακολουθώντας διεθνή πρότυπα τυποποίησης:

- υλοποιεί τους παραπάνω κανόνες,
- εξάγει τη συναισθηματική κατάσταση που προκαλείται από τα τυπογραφικά στοιχεία έντυπων ή ηλεκτρονικών εγγράφων και
- επισημειώνει τα έγγραφα με τις σχετικές συναισθηματικές ετικέτες (tags).

Στο δεύτερο μέρος, μελετάται η αντιστοίχιση των τυπογραφικών εναλλαγών σε προσωδιακές μεταβολές, η οποία επιτυγχάνεται συνδυάζοντας τα αποτελέσματα που προκύπτουν από το μοντέλο επαγόμενης συναισθηματικής κατάστασης του αναγνώστη και τα υπάρχοντα μοντέλα εκφραστικής συνθετικής ομιλίας. Παρουσιάζεται η κανονικοποίηση των προσωδιακών μεταβολών (μέσω της γραμμικής κβάντισης) και η επαύξησή τους, ως λύση στο πρόβλημα που ανακύπτει για τις δυσδιάκριτες μεταβολές της ομιλίας από τον ακροατή.

Το μοντέλο αντιστοίχισης τυπογραφικών εναλλαγών σε προσωδιακές μεταβολές που προτείνεται, εισάγει πολλαπλές περιπτώσεις ακουστικοποίησης για κάθε τυπογραφική εναλλαγή. Η χρήση πολυάριθμων ακουστικών αποδόσεων (1:N) θα προκαλούσε σύγχυση και ενόχληση στον ακροατή. Για αυτό το λόγο διεξήχθη ένα ψυχοακουστικό πείραμα για την επιλογή της βέλτιστης απόδοσης (από απλό κείμενο σε έντονη, πλάγια, έντονη-πλάγια γραφή, με αύξηση +2pt και μείωση -1pt του μεγέθους γραμματοσειράς). Τα αποτελέσματα, μέσα από τις προτιμήσεις των ακροατών, έδειξαν ότι χρησιμοποιώντας: έντονη γραφή σε γραμματοσειρά «Arial» μειώνει το τονικό ύψος κατά 12% και το ρυθμό κατά 10%, πλάγια γραφή σε «Arial» αυξάνει το τονικό ύψος και το ρυθμό κατά 36% και 30%, αντίστοιχα. Η έντονη-πλάγια γραφή σε «Times New Roman» μειώνει το τονικό ύψος κατά 36% και το ρυθμό κατά 30%. Για το μέγεθος γραμματοσειράς η αύξηση του μεγέθους κατά 2pt αυξάνει το τονικό ύψος κατά 12% και μειώνει το ρυθμό και την ένταση κατά 10% και 15%, αντίστοιχα. Η μείωση του μεγέθους

κατά 1pt αυξάνει και τα τρία προσωδιακά χαρακτηριστικά κατά 12%, 10% και 30% για το τονικό ύψος, το ρυθμό και την ένταση, αντίστοιχα. Από τα συγκεκριμένα αποτελέσματα παρατηρούμε ότι οι εναλλαγές στον τρόπο γραφής επηρεάζουν μόνο το τονικό ύψος και το ρυθμό, ενώ στην περίπτωση της μεταβολής του μεγέθους γραμματοσειράς μεταβάλλεται και η ένταση.

Το προσωδιακό μοντέλο αξιολογήθηκε μέσω δύο ψυχοακουστικών πειραμάτων. Στο πρώτο εξετάστηκε το κατά πόσο οι ακροατές είναι σε θέση να αναγνωρίσουν και να διακρίνουν τις τυπογραφικές εναλλαγές με ακουστικό τρόπο μόνο. Παρατηρείται ότι οι ακροατές, χωρίς πρότερη εκπαίδευση, μπόρεσαν να αναγνωρίσουν τις τυπογραφικές εναλλαγές (εκτός της περίπτωσης της έντονης-πλάγιας γραφής, που την αναγνώρισαν ως έντονη γραφή λόγω των παρόμοιων προσωδιακών μεταβολών τους). Στο δεύτερο πείραμα έγινε μια συγκριτική μελέτη σε βλέποντες και τυφλούς μαθητές δημοτικού, για να διερευνηθεί κατά πόσο το προτεινόμενο μοντέλο συνεισφέρει στην απόδοσή τους κατά τη διδακτική διαδικασία της παρατηρησιακής αστρονομίας. Τα αποτελέσματα έδειξαν σημαντική βελτίωση.

6.2 Περιορισμοί

Τα αποτελέσματα της παρούσας διατριβής δείχνουν ότι είναι εφικτή η ακουστικοποίηση των τυπογραφικών εναλλαγών με τη χρήση εκφραστικής συνθετικής ομιλίας. Βέβαια, παρατηρούνται και αρκετοί περιορισμοί, λόγω της πρωτοτυπίας της μεθοδολογίας. Στο επόμενο εδάφιο (6.3) παρουσιάζεται ο τρόπος εξάλειψης των περιορισμών αυτών, μέσω προτεινόμενης μελλοντικής μελέτης.

Μερικοί περιορισμοί είναι:

Ο αριθμός των τυπογραφικών στοιχείων.

Στην παρούσα μελέτη χρησιμοποιήθηκε περιορισμένος αριθμός τυπογραφικών στοιχείων, καθώς και συνδυασμών αυτών. Δεν είναι εφικτή η μαθηματική μοντελοποίηση των χρωματικών συνδυασμών γραμματοσειράς/υποβάθρου με την χρήση του συγκεκριμένου συνόλου δεδομένων παρά μόνο το γενικό συμπέρασμα ότι επηρεάζουν και τις τρεις διαστάσεις της συναισθηματικής κατάστασης του αναγνώστη (ποιοτική περιγραφή).

Επίσης, στο ακουστικό επίπεδο είναι αναγκαία η μελέτη μεγαλύτερου αριθμού τυπογραφικών στοιχείων, ώστε να υπάρχει η δυνατότητα ακουστικοποίησης των περισσότερων στοιχείων.

Δυσκολίες στην ακουστική αντιστοίχιση ορισμένων τυπογραφικών εναλλαγών.

Από τα πειραματικά αποτελέσματα παρατηρήθηκε ότι οι ακροατές δυσκολεύτηκαν στην αντιστοίχιση ορισμένων τυπογραφικών εναλλαγών σε προσωδιακές μεταβολές. Αυτός ο περιορισμός προφανώς οφείλεται στην αρχική επιλογή των ελάχιστα διακριτών επιπέδων των προσωδιακών μεταβολών, καθώς και στον τρόπο κβάντισής τους (γραμμική κβάντιση).

6.3 Μελλοντική έρευνα

Διάφορα σημαντικά αποτελέσματα, όπως για παράδειγμα, η ανάγκη μαθηματικής μοντελοποίησης της επαγόμενης συναισθηματικής κατάστασης του αναγνώστη για τους σκοπούς της παρούσας προσέγγισης, οδήγησαν σε ποσοτικές περιγραφές-αντιστοιχίσεις, οι οποίες δεν είχαν μελετηθεί διεξοδικά και συστηματικά στο παρελθόν.

Αυτή η μεθοδολογία εισάγει νέες προκλήσεις για την περαιτέρω μελέτη των τυπογραφικών στοιχείων, και όχι μόνο, με σκοπό τη μαθηματική μοντελοποίηση της συναισθηματικής κατάστασης και γενικά των συναισθημάτων του αναγνώστη ή του χρήστη, κατά την αλληλεπίδρασή του με τον υπολογιστή. Σε μελλοντική έρευνες μπορεί να μελετηθούν και μοντελοποιηθούν επιπλέον συνδυασμοί τυπογραφικών στοιχείων (π.χ. μελέτη μεγαλύτερου αριθμού χρωματικών συνδυασμών και τρόπου γραφής/τύπου γραμματοσειράς). Επίσης, μπορεί να συμπεριληφθεί το επίπεδο διάταξης ή/και το λογικό επίπεδο των εγγράφων. Άλλο ένα σημείο είναι η μελέτη του ρόλου της τυπογραφίας στην αντίληψη του αναγνώστη, δηλαδή αν το κείμενο είναι γραμμένο π.χ. σε εφημερίδα, χάρτινη συσκευασία κ.α. (δηλαδή εξαρτώμενο από το θεματικό πεδίο). Για αυτό, πρέπει να διερευνηθεί ο τρόπος με τον οποίο επηρεάζει η τυπογραφία το κείμενο το οποίο παρουσιάζεται με διάφορους τρόπους και σε διάφορους τύπους οθόνης. Επίσης, ενδιαφέρον παρουσιάζει η μελέτη της επιρροής των συναισθημάτων κατά το συνδυασμό της τυπογραφίας και του περιεχόμενου (είτε είναι συναισθηματικό είτε όχι) χρησιμοποιώντας αυστηρούς μαθηματικούς φορμαλισμούς (ποσοτική περιγραφή).

Στο ακουστικό κομμάτι, μεγάλο ενδιαφέρον θα παρουσίαζε να εξεταστεί η προτεινόμενη μεθοδολογία σε εκπαιδευμένους-εξοικειωμένους ακροατές με τις προσωδιακές μεταβολές κατά την ακουστικοποίηση των τυπογραφικών εναλλαγών. Από τα παρόντα αποτελέσματα φαίνεται ότι μια τέτοια μελέτη θα βελτιώσει κατά πολύ την αντίληψη των τυπογραφικών εναλλαγών κατά τη μετατροπή τους σε ομιλία. Όπως παρουσιάζεται στο εδάφιο 6.2, παρατηρούμε ότι αναδείχθηκαν κάποιες δυσκολίες αντιστοίχισης των

τυπογραφικών στοιχείων σε προσωδιακές μεταβολές, καθώς και στην αξιολόγησή τους. Αυτό οφείλεται κατά πάσα πιθανότητα στην επιλογή των ελάχιστων επιπέδων που επιλέχθηκαν για την αντιστοίχιση των προσωδιακών μεταβολών ή στη χρήση της γραμμικής κβάντισης (εναλλακτικά προτείνεται η επιλογή της μη-γραμμικής κβάντισης ή η αύξηση των ελάχιστων τιμών αντιστοίχισης των προσωδιακών επιπέδων μεταβολής).

Στην παρούσα διατριβή ακολουθήθηκε η μεθοδολογία μελέτης των συναισθημάτων χρησιμοποιώντας τη διαστατική τους φύση. Θα είχε πολύ ενδιαφέρον η ίδια μεθοδολογία να εφαρμοστεί σε συγκεκριμένα συναισθήματα (διακριτά συναισθήματα). Αυτό θα επέκτεινε τη μεθοδολογία για τη χρήση των μοντέλων σε συστήματα μετατροπής Κειμένου-σε-Ομιλία, που υποστηρίζουν μόνο διακριτά συναισθήματα. Να σημειώσουμε σε αυτό το σημείο ότι το προτεινόμενο μοντέλο θα μπορεί να έχει πρακτική εφαρμογή σε ευρέως διαδεδομένες συσκευές ανάγνωσης εγγράφων - τα γνωστά e-book readers - που εν γένει υποστηρίζουν τη χρήση συνθετικής ομιλίας. Επίσης, άλλη μία εφαρμογή μπορεί να είναι η χρήση της παρούσας μεθοδολογίας για τη μελέτη της ακουστική απεικόνισης των συναισθηματικών χαρακτηριστικών που επάγονται από το περιεχόμενο ενός εγγράφου.

ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΑΣ

Ξενόγλωσσος όρος	Ελληνικός Όρος
accent	προφορά, τονισμός
accessibility	προσβασιμότητα
acoustic user interface	ακουστική διεπαφή χρήστη
affective computing	συναισθηματική υπολογιστική
alignment	ευθυγράμμιση
alteration	εναλλαγή
annotation	επισημείωση δεδομένων
appraisal theory	θεωρία αποτίμησης
arousal	διέγερση
articulatory synthesis	αρθρωτική σύνθεση
assessment	αξιολόγηση
auditory icon	ακουστικό εικονίδιο
auditory synthesizer	ακουστικός συνθέτης
augment	επαυξάνω
augmented virtual reality	επαυξημένη εικονική πραγματικότητα
automated reading device	συσκευή αυτόματης ανάγνωσης
background	υπόβαθρο
baseline	βάση αναφοράς
blog	ιστολόγιο
bold	έντονη γραφή
bold-italics	πλάγια-έντονη γραφή
break	παύση
circumplex	κυκλική διάταξη, κυκλικό σύμπλεγμα
concatenative synthesis	σύνθεση μέσω συρραφής
conceptual	εννοιολογικός
content-free	ελεύθερος - ανεξάρτητος περιεχομένου
design for all	σχεδίαση για όλους
dexterity	επιδεξιότητα άνω άκρων
dimensional	διαστατικός
diphone	δίφωνο
disability	αναπηρία
document	έγγραφο
documentation	τεκμηρίωση
document-to-audio	μετατροπή εγγράφου-σε-ακουστική μορφή
dominance	επικράτηση, υπεροχή, έλεγχος
earcon	τεχνητός ήχος μικρής διάρκειας, ηχοσήμανση
embedded user interface	ενσωματωμένη διεπαφή χρήστη
enhancement	ενίσχυση, επαύξηση
font	γραμματοσειρά
font style	τρόπος γραφής (στην οπτική παρουσίαση)
formant synthesis	σύνθεση βασισμένη σε φωνοσυντονισμούς

graphical user interface	γραφική διεπαφή χρήστη
heading	κεφαλίδα, επικεφαλίδα
impairment	ανικανότητα
intonation	επιτονισμός
italics	πλάγια γραφή
kinetic typography	κινητική τυπογραφία
layout	διάταξη
leading	απόσταση διαδοχικών γραμμών του κειμένου
legible	ευανάγνωστος
manikin	φιγούρα
markup	επισημείωση
metainformation	μεταπληροφορίες
modality	αισθητηριακός τρόπος
module	άρθρωμα
mood	διάθεση
multimodal	πολυτροπικός, πολυαισθητηριακός
navigate	πλοηγούμαι
non-textual	μη-κειμενικός
pitch	τονικό ύψος, μουσικός τόνος, τόνος
pixel	εικονοστοιχείο
plain text	απλό κείμενο
pleasure	ευχαρίστηση
print disability	εντυπο-αναπηρία
prosody	προσωδία
psycho physiological	ψυχοφυσιολογική
psychophysical	ψυχοσωματική
range-dynamics	δυναμική εύρους
readable	αναγνώσιμος
recall	ανάκληση
rendition	απόδοση
retention	μνημονικό
schema	σχήμα
sentiment	αίσθημα
session	συνεδρία
style sheet	φύλλο μορφοποίησης
temporal	χρονικός, παροδικός
text-to-speech	μετατροπή κειμένου-σε-ομιλία
typeface	είδος γραμματοσειράς
typesetting	τρόπος γραφής
underlined	υπογραμμισμένος
view	οπτική, επίπεδο
web accessibility	προσβασιμότητα ιστού
writing device	μηχανισμός γραφής

ΣΥΝΤΜΗΣΕΙΣ – ΑΡΚΤΙΚΟΛΕΞΑ – ΑΚΡΩΝΥΜΙΑ

ACSS	Aural Cascading Style Sheets
ANOVA	Analysis of Variance
ARD	Automated Reading Device
AUI	Acoustic User Interface
BrailleML	Braille Markup Language
CAD	Cluster Auditory Definition
CSS	Cascading Style Sheets
DAISY	Digital Accessible Information SYstem
DE	Discrete Emotion
DocDB	Document Database
DocEmoX	Document Emotion eXtractor
DtA	Document-to-Audio
DTB	Digital Talking Book
HMM	Hidden Markov Model
HTML	HyperText Markup Language
ED	Emotion Dimensions
EmotionalML	Emotional Markup Language
ESS	Expressive Speech Synthesis
FCRM	Forced Choice Response Measurement
FRM	Free Response Measurement
GPS	Global Positioning System
GToBI	German Tone and Break Indices
GUI	Graphical User Interface
HCI	Human-Computer Interaction
HTML	HyperText Markup Language
Hz	Hertz
IAPS	International Affective Pictorial System
InkML	Ink Markup Language
LCD	Liquid Crystal Display
MathML	Mathematic Markup Language
ML	Markup Language
OCR	Optical Character Recognition
ODF	Open Document Format
PAD	Pleasure-Arousal-Dominance
PDA	Personal Digital Assistant
rmANOVA	repeated-measures Analysis of Variance
SAM	Self Assessment Manikin
SVG	Scalar Vector Graphics
TtS	Text-to-Speech
UserML	User Markup Language
W3C	World Wide Web Consortium
WAI	Web Accessibility Initiative
wpm	words per minute
XML	Extensible Markup Language
XSL	Extensible Stylesheet Language

ΧSLT	Extensible Stylesheet Language Transformations
ΔΠΣ	Διαστατική Προσέγγιση των Συναισθημάτων
ΔΧ	Διεπαφή Χρήστη
ΕΕ	Εξερευνητής Εγγράφου
Η/Υ	Ηλεκτρονικός Υπολογιστής
ΜΕΛΑ	Μετρήσεις Ελεύθερης Απόκρισης
ΜΕΞΑ	Μετρήσεις Εξαναγκασμένης Απόκρισης
ΣΟΠΕ	Στοιχείο Οπτικής Παρουσίασης των Εγγράφων
ΠΔΣ	Προσέγγιση Διακριτών Συναισθημάτων
ΨΟΒ	Ψηφιακό Ομιλόν Βιβλίο

ΑΝΑΦΟΡΕΣ

- [1] L. Schamber, "What is a document? Rethinking the concept in uneasy times," *Journal of the American Society for Information Science*, vol. 47, no. 9, Sept. 1996, pp. 669-671.
- [2] M. K. Buckland, "What is a "document"?", *Journal of the American Society for Information Science*, vol. 48, no. 9, Sept. 1997, pp. 804-809.
- [3] D. Doermann, E. Rivlin, and A. Rosenfeld, "The function of documents," *Image and Vision Computing*, vol. 16, no. 11, Aug. 1998, pp. 799-814.
- [4] C. B. Mills and L. J. Weldon, "Reading text from computer screens," *ACM Computing Surveys*, vol. 19, no. 4, Dec. 1987, pp. 329-357.
- [5] R. H. Hall and P. Hanna, "The impact of web page text-background colour combinations on readability, retention, aesthetics and behavioural intention," *Behaviour and Information Technology*, vol. 23, no. 3, Feb. 2004, pp. 183-195.
- [6] V. Eglin and S. Bres, "Document page similarity based on layout visual saliency: application to query by example and document classification," *Proc. 7th International Conference on Document Analysis and Recognition*, 2003, pp. 1208-1212.
- [7] M. Dyson and G. J. Kipping "The legibility of screen formats: Are three columns better than one?," *Computers and Graphics*, vol. 21, no. 6, Nov.- Dec. 1997, pp. 703-712.
- [8] N. Küpper, "Recording of Visual Reading Activity: Research into Newspaper Reading Behaviour," *Analysis*, Jan. 1989, pp. 1-8.
- [9] K. Holmqvist and C. Wartenberg, "The role of local design factors for newspaper reading behaviour—an eye-tracking respective," *Lund University Cognitive Studies*, vol. 127, 2005, pp. 1–21.
- [10] N. Holmberg, "Eye movement patterns and newspaper design factors. An experimental approach," MSc thesis, Lund University Cognitive Science, Lund, Sweden, 2004.
- [11] C. Wartenberg and K. Holmqvist, "Daily Newspaper Layout—Designers' Predictions of Readers' Visual Behaviour - A Case Study," *Lund University Cognitive Studies*, vol. 126, 2005, pp. 1-11.
- [12] F. Calisir, M. Eryazici, and M. R. Lehto, "The effects of text structure and prior knowledge of the learner on computer-based learning," *Computers in Human Behavior*, vol. 24, no. 2, Mar. 2008, pp. 439-450.
- [13] A. Hill and L.V. Scharff, "Readability of screen displays with various foreground/background color combinations, font styles, and font types," *Proc. 11th National Conference on Undergraduate Research*, vol. 2, 1997, pp. 742-746.
- [14] T. Porat, R. Liss, and N. Tractinsky, "E-Stores Design: The Influence of E-Store Design and Product Type on Consumers' Emotions and Attitudes," *Lecture Notes in Computer Science*, vol. 4553, 2007, pp. 712-721.
- [15] S. J. Harrington, J. F. Naveda, R. P. Jones, P. Roetling, and N. Thakkar, "Aesthetic measures for automated document layout," *Proc. ACM symposium on Document engineering (DocEng '04)*, 2004, pp. 109-111.
- [16] J. Laarni, "Effects of color, font type and font style on user preferences," *In: Stephanidis, C. (ed.) Adjunct Proceedings of HCI International 2003*, pp. 31-32, 2003.
- [17] W3C, "XML Document Navigation Language," Mar. 2000; <http://www.w3.org/TR/xdnl/>.
- [18] W3C, "Multimodal Interaction Activity," 2011; <http://www.w3.org/2002/mmi/>.
- [19] M. Czyżowicz, "Intelligent Navigation in Documents Sets Comparative Study," *Intelligent Information Processing and Web Mining*, vol. 31, M. Kłopotek et al., eds., Springer, 2005, pp. 421-425.
- [20] A. Cockburn, C. Gutwin, and J. Alexander, "Faster document navigation with space-filling thumbnails", *Proc. SIGCHI conference on Human Factors in computing systems (CHI '06)*, 2006, pp. 1-10.
- [21] D. Tsonos, G. Xydias, and G. Kouroupetroglou: "Auditory Accessibility of Metadata in Books: A Design for All Approach," *Lecture Notes in Computer Science*, vol. 4556, 2007, pp. 436-445.

- [22] M. McLuhan and Q. Fiore, *The Medium is the Message*, Gingko Press, 2005.
- [23] R. F. Lorch, "Text-Signaling Devices and Their Effects on Reading and Memory Processes," *Educational Psychology Review*, vol. 1, no. 3, Sept. 1989, pp. 209-234.
- [24] J. H. Spyridakis, "Signaling effects: A review of the research—Part I," *Journal of technical writing and communication*, vol. 19, no. 3, 1989, pp. 227-240.
- [25] J. Lemarié, H. Eyrolle, and J.-M. Cellier, "Visual signals in text comprehension: How to restore them when oralizing a text via a speech synthesis?," *Computers in Human Behavior*, vol. 22, no. 6, Nov. 2006, pp. 1096-1115.
- [26] Z. H. Han, E. S. Park, and C. Combs, "Textual Enhancement of Input: Issues and Possibilities," *Applied Linguistics*, vol. 29, no. 4, May 2008, pp. 597-618.
- [27] G. Kouroupetroglou and D. Tsonos, "Multimodal Accessibility of Documents", chapter in the book: *Advances in Human-Computer Interaction*, S. Pinder, ed., I-Tech Education and Publishing, 2008, pp. 451-470.
- [28] NISO, "Understanding metadata", 2011; <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>
- [29] IELA, "Initiative for Equitable Library Access," 2011; <http://www.lac-bac.gc.ca/ielal/>.
- [30] RPH, "Radio for the Print Handicapped," 2008; <http://www.rph.org.au/>.
- [31] K. Fellbaum and G. Kouroupetroglou, "Principles of electronic speech processing with applications for people with disabilities," *Journal Technology and Disability*, vol. 20, no. 2, July 2008, pp. 55-85.
- [32] D. Freitas and G. Kouroupetroglou, "Speech technologies for blind and low vision persons," *Journal Technology and Disability*, vol. 20, no. 2, July 2008, pp. 135-156.
- [33] F. Fourli-Kartsouni, K. Slavakis, G. Kouroupetroglou, and S. Theodoridis, "A Bayesian network approach to semantic labelling of text formatting in XML corpora of documents," *Lecture Notes in Computer Science*, vol. 4556, 2007, pp. 299-308.
- [34] G. Xydias and G. Kouroupetroglou, "Text-to-speech scripting interface for appropriate vocalisation of e-texts," *Proc. 7th European Conference on Speech Communication and Technology (EUROSPEECH 2001)*, Sept. 2001, pp. 2247-2250.
- [35] G. Xydias, D. Spiliotopoulos, and G. Kouroupetroglou, "Modelling emphatic events from non-speech aware documents in speech based user interfaces," *Proc. 10th International Conference on Human-Computer Interaction (HCII '03)*, vol. 2, 2003, pp. 806-810.
- [36] G. Xydias, V. Argyropoulos, T. Karakosta, and G. Kouroupetroglou, "An experimental approach in recognizing synthesized auditory components in a non-visual interaction with documents," *Proc. 11th International Conference on Human-Computer Interaction (HCII '05)*, July 2005, pp. 1-10.
- [37] D. Spiliotopoulos, G. Xydias, and G. Kouroupetroglou, "Diction based prosody modeling in table-to-speech synthesis," *Lecture Notes in Artificial Intelligence*, vol. 3658, 2005, pp. 294-301.
- [38] D. Spiliotopoulos, G. Xydias, G. Kouroupetroglou, and V. Argyropoulos, "Experimentation on spoken format of tables in auditory user interfaces," *Proc. 3rd International Conference on Universal Access in Human-Computer Interaction*, July 2005, pp. 1-10
- [39] C. Stephanidis, *User Interfaces for All. Concepts, Methods, and Tools*, Lawrence Erlbaum Associates, 2001.
- [40] C. Stephanidis, *Designing for all in the Information Society: Challenges towards Universal Access in the Information Age*, ERCIM ICST Report, 1999.
- [41] C. Kouroupetroglou, M. Salampassis, and A. Manitsaris, "Browsing shortcuts as a means to improve information seeking of blind people in the WWW," *Universal Access in the Information Society*, vol. 6, no. 3, 2007, pp. 273-283.
- [42] S. Harper and Y. Yesilada, "Web authoring for accessibility (WAfA)," *Journal of Web Semantics*, vol. 5, no. 3, Sept. 2007, pp. 175-179.
- [43] X. Chen, M. Tremaine, R. Lutz, J.-w. Chung, and P. Lacsina, "AudioBrowser: a mobile browsable information access for the visually impaired," *Universal Access in the Information Society*, vol. 5, no. 1, 2006, pp. 4-22.

- [44] B. J. Rosmaita, "Accessibility first!: a new approach to web design," *Proc. 37th SIGCSE technical symposium on Computer science education (SIGCSE '06)*, ACM, 2006, pp. 270-274.
- [45] J. P. Bigham, R. S. Kaminsky, R. E. Ladner, O. M. Danielsson, and G. L. Hempton, "WebInSight: making web images accessible," *Proc. 8th international ACM SIGACCESS conference on Computers and accessibility (Assets '06)*, ACM, 2006, pp. 181-188.
- [46] S. Saito, H. Takagi, and C. Asakawa, "Transforming flash to XML for accessibility evaluations," *Proc. 8th international ACM SIGACCESS conference on Computers and accessibility (Assets '06)*, ACM, 2006, pp. 157-164.
- [47] D. Spiliotopoulos, G. Xydias, G. Kouroupetroglou, V. Argyropoulos, and K. Ikospentaki, "Auditory universal accessibility of data tables using naturally derived prosody specification," *Universal Access in the Information Society*, vol. 9, no. 2, 2010, pp. 169-183.
- [48] A. D. N. Edwards, H. McCartney, and F. Fogarolo, "Lambda: a multimodal approach to making mathematics accessible to blind students," *Proc. 8th international ACM SIGACCESS conference on Computers and accessibility (Assets '06)*, ACM, 2006, pp. 48-54.
- [49] J. M. Francioni and A. C. Smith, "Computer science accessibility for students with visual disabilities," *Proc. 33rd SIGCSE technical symposium on Computer science education (SIGCSE '02)*, ACM, 2002, pp. 91-95.
- [50] D. Tsonos and G. Kouroupetroglou, "Accessibility of board and presentations in the classroom: a design-for-all approach," *Proc. International Conference on Telehealth and Assistive Technologies*, ACTA Press, April 2008, pp. 13-18.
- [51] D. Tsonos, H. Kaccori, and G. Kouroupetroglou, "A Design-for-All Approach Towards Multimodal Accessibility of Mathematics," *Proc. Assistive Technology from Adapted Equipment to Inclusive Environments (AAATE 2009)*, P.L. Emiliani et al., eds., vol. 25, IOS Press, 2009, pp. 393-397.
- [52] B. Guillon, J.-L. Monteiro, C. Checoury, D. Archambault, and D. Burger, "Towards an Integrated Publishing Chain for Accessible Multimodal Documents," *Lecture Notes in Computer Science*, vol. 3118, 2004, pp. 514-521.
- [53] Daisy, "Daisy Consortium," 2011; www.daisy.org.
- [54] W3C, "World Wide Web Consortium," 2011; <http://www.w3.org/>.
- [55] W3C, "Web Accessibility Initiative (WAI)," 2011; <http://www.w3.org/WAI/>.
- [56] P. V. Mikhaleiko, "Multimodal interaction promises device integration, accessibility and enhanced communication services," Oct. 2008; http://articles.techrepublic.com.com/5100-10878_11-5090322.html
- [57] W3C, "Web Content Accessibility Guidelines," 2011; <http://www.w3.org/TR/WAI-WEBCONTENT/>.
- [58] OASIS, "Open Document Format for Office Applications (OpenDocument) TC," 2011; http://www.oasisopen.org/committees/tc_home.php?wg_abbrev=office.
- [59] W3C, "Extensible Markup Language (XML)," 2011; <http://www.w3.org/XML/>.
- [60] OASIS, "Organization for the Advancement of Structured Information Standards", 2011; <http://www.oasis-open.org/>.
- [61] W3C, "XSL Transformations (XSLT) Version 2.0," 2007; <http://www.w3.org/TR/xslt20/>.
- [62] OASIS, "OpenDocument - Accessibility SC," 2006; http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=office-accessibility.
- [63] W3C, "Mathematical Markup Language (MathML) Version 2.0 (Second Edition)," 2003; <http://www.w3.org/TR/2003/REC-MathML2-20031021/overview.html>.
- [64] W3C, "Math Home," 2010; <http://www.w3.org/Math/>.
- [65] N. Soiffer, "MathPlayer: web-based math accessibility," *Proc. 7th international ACM SIGACCESS conference on Computers and accessibility (Assets '05)*, ACM, 2005, pp. 204-205.
- [66] Design Science, "MathPlayer," 2011; <http://www.dessci.com/en/products/mathplayer/>.

- [67] K. Masanori, M. Takeshi, K. Eizen, and U. Kazuhito, *Proposal of BrailleML as an XML for Japanese Braille: Conversion from ODF to BrailleML*, tech. report S0532B, Institute of Electronics, Information and Communication Engineers (IEIC), 2007, vol 106, no. 485, pp. 31-36.
- [68] C. Kamei-Hannan, "Innovative Solutions for Words with Emphasis: Alternative Methods of Braille Transcription," *Journal of Visual Impairment & Blindness*, vol. 103, 2009.
- [69] W3C, "Scalable Vector Graphics (SVG) 1.1 (Second Edition)," 2011: <http://www.w3.org/TR/SVG/>.
- [70] W3C, "Accessibility Features of SVG," 2000; <http://www.w3.org/TR/SVG-access/>.
- [71] ANSI/NISO, "Specifications for the digital talking book," 2002; <http://www.niso.org/standards/resources/Z39-86-2002.html>.
- [72] W3C, "Aural Style Sheets," 2009; <http://www.w3.org/TR/CSS2/aural.html>.
- [73] Λώρη (Καλλιόπη-Ειρήνη) Μαλατέστα, "Επικοινωνία Ανθρώπου-Υπολογιστή Βασισμένη σε Ανάλυση και Σύνθεση Οπτικών Πληροφοριών", Διδακτορική Διατριβή, Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Εθνικό Μετσόβιο Πολυτεχνείο, 2009.
- [74] R. R. Cornelius, "Theoretical approaches to emotion," Proc. ISCA Workshop on Speech and Emotion, 2000, pp. 1-10.
- [75] S. Narayanan and A. Alwan, *Text to Speech Synthesis: new paradigms and advances*, Prentice Hall, 2004.
- [76] W. James, "What is an emotion?," *Mind*, vol. 19, no.34, 1884, pp. 188 - 205.
- [77] K. R. Scherer, "What are emotions? And how can they be measured?," *Social Science Information*, vol. 44, no.4, Dec. 2005, pp. 695-729.
- [78] S. Brave and C. Nass, "Emotion in human-computer interaction," *The human-computer interaction handbook: Fundamentals, Evolving Technologies and Emerging Applications*, A. Sears and J. A. Jacko, eds., Taylor and Francis Group, 2007, pp. 77-92.
- [79] R. W. Picard, "Affective computing: challenges," *International Journal of Human-Computer Studies*, vol. 59, no. 1-2, July 2003, pp. 55-64.
- [80] R. W. Picard, "Affective Computing: From Laughter to IEEE," *IEEE Trans. on Affective Computing*, vol. 1, no.1, Jan. 2010, pp. 11-17.
- [81] R. W. Picard, *Affective computing*, tech. report 321, Media Laboratory Perceptual Computing Section, M.I.T, 1995.
- [82] K. R. Scherer, "Appraisal theory," *Handbook of cognition and emotion*, T. Dalgleish & M. Power, eds., Wiley, 1999, pp. 637-663.
- [83] K. R. Scherer, *Appraisal considered as a process of multilevel sequential checking*, Appraisal processes in emotion: Theory, methods, research, K. R. Scherer, A. Schorr and T. Johnstone, eds., Oxford University Press, 2001, pp. 92-120.
- [84] P. C. Ellsworth and K. R. Scherer, *Appraisal processes in emotion*, Handbook of affective sciences, R. J. Davidson, H. Goldsmith and K. R. Scherer, eds., Oxford University Press, 2003, pp. 572-595.
- [85] P. Kuppens, "From Appraisal to Emotion," *Emotion Review*, vol. 2, no. 2, April, 2010, pp. 157-158.
- [86] P. Ekman, "An argument for basic emotions," *Cognition and Emotion*, vol. 6, no. 3, 1992, pp. 169–200.
- [87] M. M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 25, no. 1, March 1994, pp. 49-59.
- [88] M. M. Bradley and P. J. Lang, "Affective reactions to acoustic stimuli," *Psychophysiology*, vol. 37, no. 2, March 2000, pp. 204-215.
- [89] C. Dormann, "Affective experiences in the Home: measuring emotion," Proc. International Conference on Home Oriented Informatics and Telematics (HOIT2003), 2003, pp. 1-13.
- [90] J. Fontaine, K. Scherer, E. Roesch, and P. Ellsworth, "The world of emotions is not two-dimensional," *Psychological Science*, vol. 18, no. 2, 2007, pp. 1050-1057.
- [91] P. J. Lang, The cognitive psychophysiology of emotion: fear and anxiety, *Anxiety and the Anxiety Disorders*, A. H. Tuma and J. D. Maser, eds., 1985, pp. 131–170.

- [92] J. D. Morris, "Observations SAM : The Self-Assessment Manikin - An Efficient Cross-Cultural Measurement Of Emotional Response," *Journal of Advertising Research* vol. 35, np. 8, pp. 63-68, Dec. 1995.
- [93] J. A. Russell, M. Lewicka, and T. Niit, "A Cross-Cultural-Study of a Circumplex Model of Affect," *Journal of Personality and Social Psychology*, vol. 57, no. 5, Nov. 1989, pp. 848-856.
- [94] D. Tsonos, K. Ikospentaki, and G. Kouroupetrolgou, "Towards modeling of Readers' Emotional State response for the automated annotation of documents," *Proc. IEEE World Congress on Computational Intelligence (WCCI 2008)*, June 2008, pp. 3253-3260.
- [95] J. Posner, J. A. Russell, and B. S. Peterson, "The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology," *Dev Psychopathol*, vol. 17, no. 3, 2005, pp. 715-734.
- [96] T. Bänziger, V. Tran, and K. Scherer, "The Geneva Emotion Wheel: a tool for the verbal report of emotional reactions," *Poster Conference of the International Society for Research on Emotions (ISRE '05)*, 2005.
- [97] R. Cowie, E. Douglas-Cowie, S. Savvidou, E. McMahon, M. Sawey, and M. Schröder, "FEELTRACE': An instrument for recording perceived emotion in real time," *ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion*, 2000, pp. 19-24.
- [98] A. C. Boucouvalas, "Real time text-to-emotion engine for expressive Internet communications," *Emerging Communication: Studies on New Technologies and Practices in Communication*, G. Riva, F. Davide, and W. IJsselsteijn, eds., IOS Press, vol. 5, 2003, pp. 305-318.
- [99] X. Zhe, D. John, and A. C. Boucouvalas, "Text-to-Emotion Engine: tests of user preferences," *IEEE International Symposium on Consumer*, 2002.
- [100] H. Liu, H. Lieberman, and T. Selker, "A model of textual affect sensing using real-world knowledge," *Proc. 8th International Conference on Intelligent User Interfaces (IUI '03)*, ACM, 2003, pp. 125-132.
- [101] C. H. Wu, Z. J. Chuang, and Y. C. Lin, "Emotion recognition from text using semantic labels and separable mixture models," *ACM Transactions on Asian Language Information Processing (TALIP)*, vol. 5, no. 2, June 2006, pp. 165-183.
- [102] H. J. Min and J. Park, "Representing Emotions with Linguistic Acuity," *Lecture Notes in Computer Science*, vol. 4394, 2007, pp. 348-360.
- [103] Y. Zhixiao, L. Lan, and Z. Dexian, "Embodiment of text based on virtual robotic avatar," *IEEE International Conference on Robotics and Biomimetics (ROBIO '07)*, Dec. 2007, pp. 1285-1289.
- [104] A. J. Gill, D. Gergle, R. M. French, and J. Oberlander, "Emotion rating from short blog texts," *Proc. 26th SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*, ACM, 2008, pp. 1121-1124.
- [105] J. T. Hancock, C. Landrigan, and C. Silver, "Expressing emotion in text-based communication," *Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*, ACM, 2007, pp. 929-932.
- [106] C. O. Alm, "Affect in text and speech," PhD Thesis, VDM Verlag Dr. Müller, Saarbrücken, Germany, Jan. 2009.
- [107] J. Read, "Using emoticons to reduce dependency in machine learning techniques for sentiment classification," *Proc. ACL Student Research Workshop (ACLstudent '05)*, Association for Computational Linguistics, 2005, pp. 43-48.
- [108] R. Cowie, E. Douglas-Cowie, B. Appolloni, J. Taylor, A. Romano, and W. Fellenz, "What a neural net needs to know about emotion words," *Computational Intelligence and Applications*, N. Mastorakis, ed., World Scientific & Engineering Society Press, 1999, pp. 109-114.
- [109] S. Owsley, S. Sood, and K. Hammond, "Domain Specific Affective Classification of Documents," *Proc. of the AAAICAAW'06*, 2006, pp. 181-183.
- [110] K. H.-Y. Lin, C. Yang, and H.-H. Chen, "What emotions do news articles trigger in their readers?," *Proc. 30th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '07)*, ACM, 2007, pp. 733-734.
- [111] J. Borchers, O. Deussen, A. Klingert, and C. Knorz, "Layout rules for graphical Web documents," *Computers and Graphics*, vol. 20, no. 3, May-June 1996, pp. 415-426.

- [112] J. Étlier, P. Hadaya, J. Talbot, and J. Cadieux, "Interface design and emotions experienced on B2C Web sites: Empirical testing of a research model," *Computers in Human Behavior*, vol. 24, no. 16, Sept. 2008, pp. 2771-2791.
- [113] F. Birren, *Color & Human Response: Aspects of Light and Color, Bearing on the Reactions of Living Things and the Welfare of Human Beings*, John Wiley & Sons, 1984.
- [114] J. A. Sanchez, I. Kirschning, J. C. Palacio, and Y. Ostrovskaya, "Towards mood-oriented interfaces for synchronous interaction," *Proc. Latin American Conference on Human-Computer Interaction (CLIHIC '05)*, ACM, 2005, pp. 1-7.
- [115] J. A. Sanchez, N. P. Hernandez, J. C. Penagos, and Y. Ostrovskaya, "Conveying mood and emotion in instant messaging by using a two-dimensional model for affective states," *Proc. 7th Brazilian Symposium on Human Factors in Computing Systems (IHC '06)*, ACM, 2006, pp. 66-72.
- [116] T. Saari, M. Turpeinen, J. Laarni, N. Ravaja, and K. Kallinen, "Emotionally Loaded Mobile Multimedia Messaging," *Entertainment Computing – ICEC 2004*, LNCS 3166, Springer, 2004, pp. 293-339.
- [117] K. Larson, "The Technology of Text," *Spectrum, IEEE*, vol. 44, no. 5, May 2007, pp. 26-31.
- [118] A. Kalra and K. Karahalios, "TextTone: Expressing Emotion Through Text," *Lecture Notes in Computer Science*, vol. 3585, 2005, pp. 966-969.
- [119] J. Ohene-Djan and R. Shipsey, "E- Subtitles: Emotional Subtitles as a Technology to Assist the Deaf and Hearing-Impaired when Learning from Television and Film," *6th International Conference on Advanced Learning Technologies*, July 2006, pp. 464-466.
- [120] J. Ohene-Djan, J. Wright, and C.-S. K., "Emotional subtitles: a system and potential applications for deaf and hearing impaired people," *Proc. Conference and Workshop on Assistive Technologies for People with Vision and Hearing Impairments: Assistive Technology for All Ages (CVHI '07)*, 2007.
- [121] A. Yannicopoulou, "Visual Aspects of Written Texts: Preschoolers View Comics," *L1-Educational Studies in Language and Literature*, vol. 4, no. 2, Jan. 2004, pp. 169-181.
- [122] T. Rosenberger, "Prosodic font: between the spoken and the written," MSc thesis, Media Arts and Sciences, MIT, 1988.
- [123] T. Rosenberger and R. L. MacNeil, "Prosodic font: translating speech into graphics," *Proc. CHI '99 extended abstracts on Human Factors in Computing Systems (CHI EA '99)*, ACM, 1999, pp. 252-253.
- [124] D. Tsonos and G. Kouroupetroglou, "Modeling Reader's Emotional State Response on Document's Typographic Elements," *Advances in Human-Computer Interaction*, vol. 2011, 2011, pp. 1-18.
- [125] G. Kouroupetroglou, D. Tsonos, and E. Vlahos, "DocEmoX: A System for the Typography-Derived Emotional Annotation of Documents," *Lecture Notes in Computer Science*, vol. 5616, 2009, pp. 550-558.
- [126] Γ. Ν. Ξύδας, "Μοντέλα Μηχανικής Μάθησης και Μέθοδοι Επιλογής για την Ακουστική Αναπαράσταση Εγγράφων μέσω Συνθετικής Ομιλίας με Εμπλουτισμένη Προσωδία," Διδακτορική Διατριβή, Τμήμα Πληροφορικής και Τηλεπικοινωνιών, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, 2006.
- [127] Γ. Θ. Κουρουπέτρογλου, *Μαθήματα Επεξεργασίας Ομιλίας*. Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών, Αθήνα, 1998.
- [128] G. Xydias, V. Argyropoulos, T. Karakosta, and G. Kouroupetroglou, "An Open Platform for Conducting Psycho-Acoustic Experiments in the Auditory Representation of Web Documents", *Proc. Conference ACOUSTICS 2004*, Sept. 2004, pp. 157-164.
- [129] J. Launay, L. Segalen, I. Kanellos, T. Moudenc, C. Otesteanu, A. David, G. Fang, and J. Jin, "Speech expressiveness: Modeling and implementing the expressive impact of typographic and punctuation marks for textual inputs," *Proc. 3rd International Conference on Information and Communication Technologies: From Theory to Applications (ICTTA 2008)*, April 2008, pp. 1-6.
- [130] M. Tatham and K. Morton, *Expression in speech: analysis and synthesis*, Oxford University Press, 2006.
- [131] N. Campbell, "Expressive/Affective Speech Synthesis," chapter in the book: *Handbook of Speech Processing*, J. Benesty, M. M. Sondhi, and Y. Huang, Eds., ed: Springer Berlin Heidelberg, 2008, pp. 505-518.

- [132] N. Campbell, W. Hamza, H. Hoge, J. H. Tao, and G. Bailly, "Editorial - Special section on expressive speech synthesis," *IEEE Trans. on Audio Speech and Language Processing*, vol. 14, no. 4, Jul. 2006, pp. 1097-1098.
- [133] E. Eide, A. Aaron, R. Bakis, W. Hamza, M. Picheny, and J. Pitrelli, "A Corpus-Based Approach to <AHEM/> Expressive Speech Synthesis," *Proc. 5th ISCA Speech Synthesis Workshop*, 2004, pp 79-84.
- [134] C. Drioli, G. Tisato, P. Cosi, and F. Tesser, "Emotions and voice quality: experiments with sinusoidal modeling," *Proc. ISCA Tutorial and Research Workshop on Voice Quality Functions Analysis and Synthesis (VOQUAL '03)*, Aug. 2003, pp. 127-132.
- [135] J. F. Pitrelli, R. Bakis, E. M. Eide, R. Fernandez, W. Hamza, and M. A. Picheny, "The IBM expressive text-to-speech synthesis system for American English", *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 14, pp. 1099-1108, 2006.
- [136] M. Theune, K. Meijjs, D. Heylen, and R. Ordelman, "Generating expressive speech for storytelling applications," *IEEE Trans. on Audio, Speech, and Language Processing*, vol. 14, no. 4, July 2006, pp. 1137-1144.
- [137] W. L. Johnson, S. Narayanan, R. Whitney, R. Das, M. Bulut, and C. LaBore, "Limited domain synthesis of expressive military speech for animated characters," *Proc. of IEEE Workshop on Speech Synthesis*, Sept. 2002, pp. 163-166.
- [138] M. Schröder, "Emotional speech synthesis: A review," *Proc. 7th European Conference on Speech Communication and Technology*, vol. 1, 2001, pp. 561-564.
- [139] M. Schröder, "Expressive Speech Synthesis: Past, Present, and Possible Futures," *Affective Information Processing*, J. Tao and T. Tan, eds., Springer, 2009, pp. 111-126.
- [140] S. Krstulović, A. Hunecke, and M. Schröder, "An HMM-based speech synthesis system applied to german and its adaptation to a limited set of expressive football announcements," *Proc. 8th Annual Conference of the International Speech Communication Association (INTERSPEECH '07)*, 2007, pp. 1897-1900.
- [141] A. Iida, N. Campbell, F. Higuchi, and M. Yasumura, "A corpus-based speech synthesis system with emotion," *Speech Communication*, vol. 40, no. 1-2, 2003, pp. 161-187.
- [142] M. Schröder, "Expressing degree of activation in synthetic speech," *IEEE Trans. on Audio Speech and Language Processing*, vol. 14, no. 4, Jul 2006, pp. 1128-1136.
- [143] M. Schröder and J. Trouvain, "The German Text-to-Speech Synthesis System MARY: A Tool for Research, Development and Teaching," *International Journal of Speech Technology*, vol. 6, no. 4, 2003, pp. 365-377.
- [144] A. Abelin and J. Allwood, "Cross linguistic interpretation of expressions of emotions," *Proc. 8th Simposio Internacional de Comunicacion Social*, 2003.
- [145] K. R. Scherer, R. Banse, and H. G. Wallbott, "Emotion Inferences from Vocal Expression Correlate Across Languages and Cultures," *Journal of Cross-Cultural Psychology*, vol. 32, no. 1, Jan. 2001, pp. 76-92.
- [146] M. D. Pell, S. Paulmann, C. Dara, A. Alasseri, and S. A. Kotz, "Factors in the recognition of vocally expressed emotions: A comparison of four languages," *Journal of Phonetics*, vol. 37, no. 4, Oct. 2009, pp. 417-435.
- [147] F. Burkhardt, N. Audibert, L. Malatesta, O. Türk, L. Arslan, and V. Auberger, "Emotional Prosody-Does Culture Make A Difference?," *Speech Prosody, 2006*, pp. 2-5.
- [148] G. Xydias and G. Kouroupetroglou, "Augmented Auditory Representation of e-Texts for Text-to-Speech Systems Text, Speech and Dialogue," *Lecture Notes in Artificial Intelligence*, vol. 2166, 2001, pp. 134-141.
- [149] C. Gussenhoven, "Intonation and interpretation: phonetics and phonology," *Proc. of Speech Prosody*, 2002, pp. 47-57.
- [150] K. Kallinen, J. Laarni, N. Ravaja, and T. Saari, "Auditive "Boldfacing", Emotional Characteristics of Message, and Individual Differences in Memory Acquisition of Computer Mediated Business News," *Proc. 11th International Conference on Human-Computer Interaction*, 2005.
- [151] P. Truillet, B. Oriola, J. L. Nespoulous, and N. Vigoroux, "Effect of Sound Fonts in an Aural Presentation", *Proc. 6th ERCIM Workshop (UI4ALL 2000)*, 2000, pp. 135-144.

- [152] V. S. Argyropoulos, G. D. Sideridis, G. Kouroupetroglou, and G. Xydias, "Auditory discriminations of typographic attributes of documents by students with blindness", *British Journal of Visual Impairment*, vol. 27, no. 3, 2009, pp. 183-203.
- [153] K. Papadopoulos, V. Argyropoulos, and G. Kouroupetroglou, "Discrimination, perception and comprehension of synthetic speech by students with visual impairments: the case of similar acoustic patterns," *Journal of Visual Impairment and Blindness*, vol. 102, no. 7, 2008, pp. 420-429.
- [154] K. Papadopoulos, A. Koutsoklenis, E. Katemidou, and A. Okalidou, "Perception of synthetic and natural speech by adults with visual impairments," *Journal of Visual Impairment and Blindness*, vol. 103, no. 7, 2009, pp. 403-414.
- [155] K. Papadopoulos, E. Katemidou, A. Koutsoklenis, and E. Mouratidou, "Differences amongst sighted individuals and individuals with visual impairments in word intelligibility presented via synthetic and natural speech," *Augmentative and Alternative Communication*, vol. 26, no. 4, 2010, pp. 278-288.
- [156] T. V. Raman, "An Audio view of (LA)TEX Documents," *Proc. of the Annual Meeting TUGboat*, 1992.
- [157] G. Kramer, *Auditory Display: Sonification, Audification, And Auditory Interfaces*, Perseus Publishing, 1993.
- [158] S. A. Brewster, V. Rätty, and A. Kortekangas, "Earcons as a Method of Providing Navigational Cues in a Menu Hierarchy," *Proc. HCI on People and Computers XI (HCI '96)*, ACM, 1996, pp. 169-183.
- [159] E. D. Mynatt, "Designing with auditory icons: how well do we identify auditory cues?," *Proc. Conference Companion on Human Factors in Computing Systems (CHI '94)*, ACM, 1994, pp. 269-270.
- [160] P. Gorny, "Typographic semantics of Webpages Accessible for Visual Impaired Users, Mapping Layout and Interaction Objects to an Auditory Interaction Space," *Proc. International Conference on Computer Helping with Special Needs*, 2000, pp. 17-21.
- [161] S. Djennane, "3D-audio news presentation modeling for mobile environment," *Universal Access: Theoretical Perspectives, Practice, and Experience*, LNCS 2615, Springer, pp. 280-286, 2003.
- [162] C. D. Power, "Multi-Modal Exploration," PhD thesis, University of Western Ontario, Canada, 2008.
- [163] A. Conway, "Page grammars and page parsing. A syntactic approach to document layout recognition," *Proceedings of the Second International Conference on Document Analysis and Recognition*, Oct. 1993, pp. 761-764.
- [164] A. Yamashita, T. Amano, I. Takahashi, and K. Toyokawa, "A model based layout understanding method for the document recognition system," *Proc. International Conference on Document Analysis and Recognition (ICDAR91)*, 1991, pp. 130-138.
- [165] D. Derrien-Peden, "Frame-based system for macro-typographical structure analysis in scientific papers," *Proc. International Conference on Document Analysis and Recognition (ICDAR91)*, 1991, pp. 311-319.
- [166] M. Krishnamoorthy, G. Nagy, S. Seth, and M. Viswanathan, "Syntactic segmentation and labeling of digitized pages from technical journals," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, no. 7, July 1993, pp. 737-747.
- [167] B. Gatos, S. L. Mantzaris, S. J. Perantonis, and A. Tsigris, "Automatic page analysis for the creation of a digital library from newspaper archives," *International Journal on Digital Libraries (IJDL)*, vol. 3, no.1, pp. 77-84, 2000.
- [168] B. Gatos, D. Danatsas, I. Pratikakis, and S. J. Perantonis, "Automatic Table Detection in Document Images," *Proc. 3rd International Conference on Advances in Pattern Recognition (ICAPR'05)*, Aug. 2005, pp. 612-621.
- [169] S. Mao, A. Rosenfeld, and T. Kanungo, "Document structure analysis algorithms: a literature survey," *Proc. Document Recognition and Retrieval*, vol. 5010, Jan. 2003, pp. 197-207.
- [170] S. Tsujimoto and H. Asada, "Understanding multi-articled document," *Proc. 10th International Conference on Pattern Recognition*, vol. 1, 1990, pp. 551-556.
- [171] D. Tsonos, G. Xydias, and G. Kouroupetroglou, "A Methodology for Reader's Emotional State Extraction to Augment Expressions in Speech Synthesis," *Proc. 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007)*, vol 2, Oct. 2007, pp. 218-225.

- [172] J. Fink, A. Kobsa, and A. Nill, "Adaptable and adaptive information provision for all users, including disabled and elderly people," *The New Review of Hypermedia and Multimedia*, vol. 4, no. 1, Taylor & Francis, 1998, pp. 163-188.
- [173] A. Hjalmarsson, "Adaptive Spoken Dialog Systems," GSLT, Speech Technology 1 Closing Seminar, 2005.
- [174] M. F. McTear, "Spoken dialogue technology: Enabling the conversational user interface," *ACM Computing Surveys*, vol. 34, no. 1, March 2002, pp. 90-169.
- [175] M. Turunen, J. Hakulinen, K. J. Raiha, E. P. Salonen, A. Kainulainen, and P. Prusi, "An architecture and applications for speech-based accessibility systems," *IBM Systems Journal*, vol. 44, no. 3, 2005, pp. 485-504.
- [176] D. Archambault, M. Batusic, F. Berger, D. Fitzpatrick, K. Miesenberger, V. Moço, and B. Stöger, "The Universal Maths Conversion Library: an attempt to build an Open Software Library to Convert Mathematical Contents in various Formats," *Proc. of UAHCI Conf*, Las Vegas, 2005.
- [177] M. Grimm and K. Kroschel, "Evaluation of natural emotions using self assessment manikins," *IEEE Workshop on Automatic Speech Recognition and Understanding*, Nov. 2005, pp. 381-385.
- [178] Κ. Εικοσπεντακη και Δ. Τσώνος, "Ενδιάμεση Έκθεση Προόδου - ΟΜΗΡΟΣ: Μεθοδολογική προσέγγιση για τη Σχεδίαση Ακουστικής - Απτικής Αλληλεπίδρασης στη Μη-Οπτική Διεπαφή Χρήστη (User Interface) με έμφαση στην Προσβασιμότητα Ατόμων με Απώλεια Όρασης", Πρόγραμμα Ενίσχυσης του Ερευνητικού Δυναμικού (ΠΕΝΕΔ) – 2003, Γενική Γραμματεία Έρευνας και Τεχνολογίας και την Ευρωπαϊκή Ένωση-Ευρωπαϊκό Κοινωνικό Ταμείο, 2007.
- [179] PHP, "PHP 5.0," 2012; <http://www.php.net/>.
- [180] Apache HTTP Server Project, "Apache HTTP Server 2.2.21," 2011; <http://httpd.apache.org/>.
- [181] MySQL, "Open Source Database Server," 2011; <http://www.mysql.com/>.
- [182] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*, Tech. Report A-8. University of Florida, Gainesville, FL, 2005.
- [183] M. L. Bernard, B. S. Chaparro, M. M. Mills, and C. G. Halcomb, "Comparing the effects of text size and format on the readability of computer-displayed Times New Roman and Arial text," *International Journal of Human-Computer Studies*, vol. 59, no. 6, Dec. 2003, pp. 823-835.
- [184] K. E. Schmidt, M. Bauerly, Y. Liu, and S. Sridharan, "Web page aesthetics and performance: a survey and an experimental study," *Proc. 8th Annual International Conference on Industrial Engineering—Theory, Applications and Practice*, 2003, pp. 478-484.
- [185] A. Alsumait, A. Al-Osaimi, and H. AlFedaghi, "Arab Children's Reading Preference for Different Online Fonts," *Lecture Notes in Computer Science*, vol. 5613, 2009, pp. 3-11.
- [186] J. Ling and P. van Schaik, "The influence of font type and line length on visual search and information retrieval in web pages," *International Journal of Human-Computer Studies*, vol. 64, no. 5, May 2006, pp. 395-404.
- [187] A. Al-badi and P. Mayhew, "Cultural perception of user interface of a website: comparative study between liberal and prescriptive cultures (Arab vs. British)," *Proc. IADIS International Conference E-Society*, July 2004, pp. 997.
- [188] D. Cyr, M. Head, and H. Larios, "Colour appeal in website design within and across cultures: A multi-method evaluation," *International Journal of Human-Computer Studies*, vol. 68, no. 1-2, Jan. - Feb. 2010, pp. 1-21.
- [189] OriginLab, "Origin 8.x.," 2011; <http://www.originlab.com/index.aspx?go=PRODUCTS/Origin>
- [190] C. Bernardini, V. Ambrogi, G. Fardella, L. Perioli, and G. Grandolini, "How to improve the readability of the patient package leaflet: a survey on the use of colour, print size and layout," *Pharmacological Research*, vol. 43, no. 5, May 2001, pp. 437-444.
- [191] D. Tsonos and G. Kouroupetroglou, "A Methodology for the Extraction of Reader's Emotional State Triggered from Text Typography," chapter in the book: *Tools in Artificial Intelligence*, P. Fritzsche, ed., In-Tech Education and Publishing, 2008, pp. 439-454.
- [192] Ε. Βλάχος, "DocEmoX: Σύστημα Επισημείωσης Εγγράφων με τα Συναισθήματα που Απορρέουν από τα Τυπογραφικά τους Χαρακτηριστικά", Διπλωματική Εργασία, Διατμηματικό Μεταπτυχιακό

Πρόγραμμα Σπουδών Ηλεκτρονικής και Αυτοματισμού, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Αθήνα, 2009.

- [193] ArsAperta.org, "odfconvert," 2006; <https://www.arsaperta.org/odftoolsen.html>.
- [194] OpenOffice.org, "Save As DAISY (odt2dtbook) 1.1.0.," 2009; <http://extensions.services.openoffice.org/node/1839>.
- [195] OpenOffice.org, "ODF Plug-in for Microsoft Office," 2011; <http://extensions.services.openoffice.org/>.
- [196] W3C, "Emotion Markup Language Incubator Group," 2008; <http://www.w3.org/2005/Incubator/emotion/>.
- [197] W3C, "Cascading Style Sheets Level 2 Revision 1 (CSS 2.1) Specification, Chapter 4: Syntax and basic data types," 2011; <http://www.w3.org/TR/CSS2/syntax.html#values>.
- [198] S. Nootboom, "The prosody of speech: melody and rhythm," *The Handbook of Phonetic Sciences*, W. J. Hardcastle and J. Laver, eds., 1997, pp. 640-673.
- [199] A. Rietveld and C. Gussenhoven, "On the relation between pitch excursion size and prominence," *Journal of Phonetics*, vol. 13, no. 3, July 1985, pp. 299-308.
- [200] D. Braga, L. Coelho, F. G. Resende, and M. Dias, "Subjective and objective assessment of TTS voice font quality," *Proc. 12th International Conference Speech and Computer (SPECOM 2007)*, 2007.
- [201] K. M. K. Chan and E. M. L. Yiu, "A comparison of two perceptual voice evaluation training programs for naive listeners," *Journal of Voice*, vol. 20, no. 2, June 2006, pp. 229-241.
- [202] J. Bird and C. Darwin, "Effects of a difference in fundamental frequency in separating two sentences," *Psychophysical and physiological advances in hearing*, A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis, eds., Whurr Publishers, 1998, pp. 263-269.
- [203] C. Jones, L. Berry, and C. Stevens, "Synthesized speech intelligibility and persuasion: Speech rate and non-native listeners," *Computer Speech & Language*, vol. 21, no. 4, Oct. 2007, pp. 641-651.
- [204] J. Romportl and J. Kala, "Prosody modelling in Czech text-to-speech synthesis," *Proc. 6th ISCA Workshop on Speech Synthesis*, Aug. 2007, pp. 200-205.
- [205] W. J. Hardcastle, J. Laver, and F. E. Gibbon, *The handbook of phonetic sciences*, 2nd edition, Wiley-Blackwell, 2010.
- [206] M. Steedman, "Information structure and the syntax-phonology interface," *Linguistic inquiry*, vol. 31, no. 4, 2006, pp. 649-689.
- [207] X. Huang, A. Acero, H. W. Hon, and R. Reddy, *Spoken Language Processing: A Guide to Theory, Algorithm & System Development*, Prentice Hall PTR, 2001.
- [208] M. Schröder and M. Grice, "Expressing vocal effort in concatenative synthesis," *Proc. 15th International Congress of Phonetic Sciences (ICPhS-2003)*, 2003, pp. 2589-2592.
- [209] Κ. Εικοσπεντάκη, "Ακουστική-απτική αντιληπτική ικανότητα στην μη-οπτική διεπαφή χρήστη," Διδακτορική Διατριβή, Διαπανεπιστημιακό - Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών στη «Βασική και Εφαρμοσμένη Γνωσιακή Επιστήμη», Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Αθήνα, 2012.
- [210] Κ. Εικοσπεντάκη και Σ. Βοσνιάδου, "Η Ανάπτυξη των Απόψεων των Εκ Γενετής Τυφλών Παιδιών για το Σχήμα της Γης και την Εναλλαγή Ημέρας/Νύχτας," *Ψυχολογία*, τόμος 18, τεύχος 1, 2011, σελ. 20-36.
- [211] S. Vosniadou and W. F. Brewer, "Mental Models of the Earth: A Study of Conceptual Change in Childhood," *Cognitive Psychology*, vol. 24, pp. 535-585, 1992.
- [212] S. Vosniadou and W. F. Brewer, "Mental Models of the Day/Night Cycle," *Cognitive Science*, vol. 18, 1994, pp. 123-183.
- [213] Smart technologies, "SMARTBoard," 2011; <http://smarttech.com/>.

ΔΗΜΟΣΙΕΥΣΕΙΣ

Δημοσιεύσεις σε Διεθνή Περιοδικά με κρίση:

- D. Tsonos and G. Kouroupetroglou: "Modeling reader's emotional state response on document's typographic elements" *Advances in Human Computer Interaction*, vol. 2011, Article ID 206983, pp.1-18, doi:10.1155/2011/206983, 2011.
- Dimitrios Tsonos and Georgios Kouroupetroglou: "Prosodic Mapping of Typographic Alterations based on the Dimensional Theory of Emotions" *IEEE Transactions on Affective Computing* (υπό κρίση).
- K. Ikospentaki, D. Tsonos, S. Vosniadou, and G. Kouroupetroglou: "Sonification of Board Typographic Elements: Does it Enhance the Traditional Teaching Approach of Sighted and Blind Students?" *Journal of Visual Impairment and Blindness* (in preparation), 2012
- F. Katsoulis, D. Tsonos and G. Kouroupetroglou: "Analysis of Text Signaling Devices" *International Journal on Document Analysis and Recognition* , (in preparation), 2012

Δημοσιεύσεις σε Βιβλία με κρίση:

- D. Tsonos and G. Kouroupetroglou: "A Methodology for the Extraction of Reader's Emotional State Triggered from Text Typography" chapter in the book: Paula Fritzsche (ed.) *Tools in Artificial Intelligence*, In-Tech Publishing, Vienna, 2008, pp. 439-454, ISBN 978-953-7619-03-9.
- G. Kouroupetroglou and D. Tsonos: "Multimodal Accessibility of Documents" chapter in the book: Shane Pinder (ed.) *Advances in Human-Computer Interaction*, I-Tech Education and Publishing, Vienna, 2008, pp. 451-470, ISBN 978-953-7619-15-2

Δημοσιεύσεις σε Πρακτικά Διεθνών Συνεδρίων με κρίση:

- G. Kouroupetroglou, D. Tsonos, and E. Vlahos: "DocEmoX: A System for the Typography-Derived Emotional Annotation of Documents" *Lecture Notes in Computer Science (LNCS)*, Vol. 5616, 2009, pp. 550-558, Proceedings of the 5th International Conference on Universal Access in Human-Computer Interaction, 19-24 July 2009, San Diego, CA, USA.
- D. Tsonos, H. Kaccori and G. Kouroupetroglou: "A Design-for-All Approach Towards Multimodal Accessibility of Mathematics" In P.L. Emiliani et al. (Eds.) *Assistive Technology from Adapted Equipment to Inclusive Environments*, pp. 393-397, Assistive Technology Research Series, Vol. 25, IOS Press, Amsterdam. Proceedings of the 10th International Conference of the Association for the Advancement of Assistive Technology in Europe (AAATE), August 31–Sept. 2, 2009, Florence, Italy.
- D. Tsonos and G. Kouroupetroglou: "Accessibility of Board and Presentations in the classroom: a Design-for-All Approach" *Proc. of the International Conference on Telehealth and Assistive Technologies*, April 16 – 18, 2008, Baltimore, Maryland, USA, pp 13-18.
- D. Tsonos, K. Ikospentaki, and G. Kouroupetroglou: "Towards Modeling of Readers' Emotional State Response for the Automated Annotation of Documents" *Proc. of the IEEE*

World Congress on Computational Intelligence (WCCI 2008), Hong Kong, June 1-6, 2008, pp. 3252-3259.

- D. Tsonos, G. Xydas, and G. Kouroupetroglou: "Auditory Accessibility of Metadata in Books: A Design for All Approach" *Lecture Notes in Computer Science (LNCS)*, Vol. 4556, 2007, pp. 436-445, Proc. of the 4th Int. Conference on Universal Access in Human-Computer Interaction (UACHI07), Beijing, China, July 22-27, 2007.
- D. Tsonos, G. Xydas, and G. Kouroupetroglou: "A Methodology for Reader's Emotional State Extraction to Augment Expressions in Speech Synthesis" *Proc. 19th IEEE Int. Conf. on Tools with Artificial Intelligence (ICTAI 2007)*, Patras, Greece, 29-31 Oct. 2007, Vol. II, pp. 218-225.
- K. Ikospentaki, S. Vosniadou, D. Tsonos, and G. Kouroupetroglou: "HOMER: A Design for the Development of Acoustical-Haptic Representations of Document Meta-Data for Use by Persons with Vision Loss" *Proc. of the Second European Cognitive Science Conference (EuroCogSci 2007)*, Delphi, 23-27 May 2007, pp. 912.

Τεχνικές Εκθέσεις:

- Δ. Σπηλιωτόπουλος, Δ. Τσώνος, Γ. Ξύδας, Σ. Βουράκης, Μ. Μπαλαζάνη, Δ. Βάσσο, Γ. Δεβελέκος και Γ. Κουρουπέτρογλου: "Αξιολόγηση Συστήματος από Ομάδες Χρηστών," Πρόγραμμα «ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΩΝ, ΗΧΟΥ ΚΑΙ ΓΛΩΣΣΑΣ», ΓΓΕΤ, έργο «ΡΗΤΩΡ: Συστηματική Προσέγγιση διεπαφής φωνητικού διαλόγου με ακουστικοποίηση της μεταπληροφορίας και οπτικής δομής εγγράφων με σκοπό την καθολική πρόσβαση σε έντυπο περιέχομενο», 2008
- Μ. Μπαλαζάνη, Δ. Σπηλιωτόπουλος, Γ. Ξύδας, Δ. Τσώνος και Γ. Κουρουπέτρογλου: "Περιγραφή του βέλτιστου φυσικού μοντέλου ακουστικοποίησης μεταδεδομένων εγγράφων," Πρόγραμμα «ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΩΝ, ΗΧΟΥ ΚΑΙ ΓΛΩΣΣΑΣ», ΓΓΕΤ, έργο «ΡΗΤΩΡ: Συστηματική προσέγγιση διεπαφής φωνητικού διαλόγου με ακουστικοποίηση της μεταπληροφορίας και οπτικής δομής εγγράφων με σκοπό την καθολική πρόσβαση σε έντυπο περιέχομενο», 2008
- Κ. Εικοσπεντακη και Δ. Τσώνος: "Ενδιάμεση Έκθεση Προόδου," έργο ΟΜΗΡΟΣ Μεθοδολογική Προσέγγιση για τη Σχεδίαση Ακουστικής-Απτικής Αλληλεπίδρασης στη ΜΗ-Οπτική Διεπαφή Χρήστη (user interface) με Έμφαση στην Προσβασιμότητα Ατόμων με Απώλεια Όρασης», Πρόγραμμα ΠΕΝΕΔ, ΓΓΕΤ, 2007