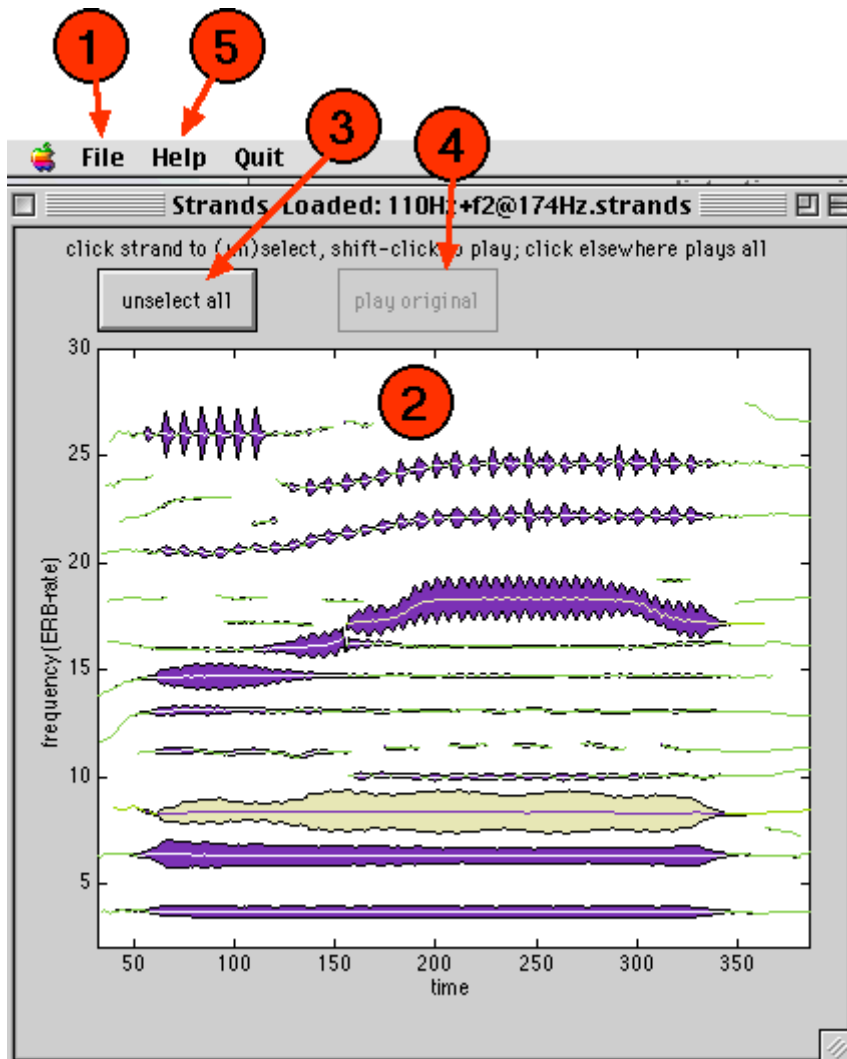


ΝΗΜΑΤΑ : ΑΝΑΦΟΡΑ ΣΤΗΝ ΑΚΟΥΣΤΙΚΗ ΧΡΟΝΟ-ΣΥΧΝΟΤΗΤΑ

ΕΙΣΑΓΩΓΗ

Τα νήματα (Cooke, 1991) είναι περιγραφές με βάση την χρονο-συχνότητα που επιχειρούν να συλλάβουν τα σημαντικά χαρακτηριστικά των σημάτων ομιλίας με μια μορφή συμβατή με την εφαρμογή ακουστικών αρχών ομαδοποίησης. (Darwin & Carlyon, 1995. Bregman, 1990). Ατομικά νήματα επιχειρούν να ανιχνεύσουν τα ίχνη ατομικών φασματικών θέσεων ισχύος μέσα στο χρόνο. Τέτοιες θέσεις ισχύος μπορεί ν' αντιστοιχούν σε λελυμένες αρμονικές, σε formants, ή σε άλλα συστατικά των σημάτων σχετικώς στενής ζώνης συχνοτήτων. Τα νήματα μπορούν να μετατραπούν σε ήχο («ανασύνθεση») ατομικά ή συλλογικά.

Η Παρουσίαση



Πληκτρολογήστε «strands» για να ξεκινήσει η παρουσίαση. Όταν το παράθυρο εμφανίζεται, χρησιμοποιήστε το μενού φόρτωσης (1) για να φορτώσετε ένα αρχείο νημάτων. Αν υπάρχει το σχετικό αρχείο ήχου, τότε θα ενεργοποιηθεί το κουμπί «play original» (4). Άμα φορτωθούν τα νήματα, μπορούν να ανοιχτούν ή να σβηστούν αν κάνετε κλικ με το ποντίκι.

Αν κάνετε κλικ αλλού στο display panel (2) τότε παίζεται ο ήχος που προέρχεται από την πρόσθεση της επιλεγμένης ανασύνθεσης νημάτων. Το κουμπί (3) κάνει ευκολότερα επιλογή ή καταργεί την επιλογή όλων των εκτιθεμένων νημάτων. Το να κάνετε κλικ ολίσθησης (στο Mac τουλάχιστον δοκιμάστε τα δεξιά κουμπιά σε άλλες πλατφόρμες) , σε ατομικά νήματα τα κάνει να παιχθούν.

Θέματα προς έρευνα

1. Το σημείο εκκίνησης για την παραγωγή νημάτων είναι ένα μοντέλο ακουστικής περιφέρειας βασισμένο σε φίλτρα gammatone (Patterson & Holdsworth, 1990). Μια συνέπεια των εξαρτώμενων από την συχνότητα εύρους ζώνης, του διαμορφωμένου ακουστικού συνόλου ζωνοπερατών φίλτρων, είναι η λύση των αρμονικών συστατικών στις χαμηλές-μεσαίες συχνότητες (low-mid) και η έλλειψη λύσης στις μεσαίες-υψηλές συχνότητες (που οδηγούν στην εμφάνιση formants). Φορτώστε οποιοδήποτε αρχείο νημάτων για να παρατηρήσετε αυτή την διάκριση.
2. Ένα σετ από παραδείγματα νημάτων που συνοδεύουν αυτά τα φυλλάδια προέρχεται από τα «ru» - «li» ερεθίσματα που δημιουργήθηκαν από τον Chris Darwin (Darwin, 1981, Gardner & al, 1989) και εξετάστηκαν με υπολογισμούς από τον Cooke (1991). Αυτές είναι συνθετικές συλλαβές με τέσσερα formants. Υπό μια προϋπόθεση, τα formants συντίθενται στο ίδιο F0 (110 Hz). Υπό άλλες συνθήκες, το δεύτερο Formant έχει μια διαφορετική βασική συχνότητα από τους άλλους τρεις (που παραμένουν στα 110 Hz). Το αποτέλεσμα αυτής της F0 διαφοράς μπορεί να φανεί φορτώνοντας τα συντιθέμενα αρχεία νημάτων, τα οποία έχουν ονόματα όπως ru112 (που σηματοδοτούν ένα F2 στα 112 Hz). Μετακινώντας με το χέρι τα νήματα F2, μπορείτε ν' ακούσετε «li» αντί για «ru».
3. Ένα υποσύνολο του υλικού που χρησιμοποιήθηκε για να εκτιμηθεί ο διαχωρισμός της ομιλίας στον Cooke (1991) είναι διαθέσιμο στην τυποποιημένη κατανομή. Σχετικά αρχεία έχουν ονόματα που αρχίζουν από v0. Το αρχείο vo είναι μιας μοναδικής πηγής έκφραση, ενώ το v0n* σημαίνει v0 με προστιθέμενο θόρυβο. Επιχειρώντας χειρωνακτικά να ομαδοποιήσετε νήματα ομιλίας, μπορείτε να πάρετε μια ιδέα του τι απόδοση μπορεί να είναι δυνατή από ένα σύστημα για ομαδοποίηση βασισμένη σε νήματα. [Υπάρχει ελπίδα να εφαρμοστεί κάτι όμοιο με του Cooke το πρωτότυπο σύστημα σαν παρουσίαση στην επόμενη έκδοση.

Περαιτέρω μελέτη

- Για άλλες απεικονήσεις σημάτων ομιλίας, δες την εκδοθείσα συλλογή Cooke, Beet & Crawford (1993). Visual Representations of Speech Signals. Wiley .

