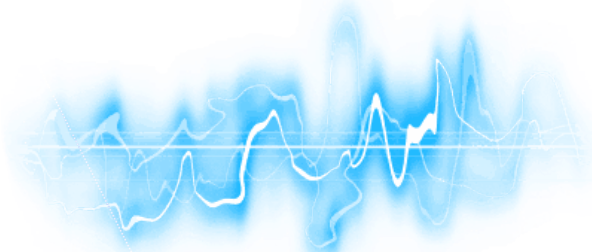


«Πλαστελίνη» φωνής



Πρόσεχε πώς μιλάς!!!



Γράψε και άκου!



Άκουσέ
με καλά!



Εργαστήριο Επεξεργασίας Σήματος Φωνής
Τμήμα Επιστήμης Υπολογιστών

Παρουσίαση: Δρ Γιώργος Καφεντζής
Επισκέπτης Καθηγητής

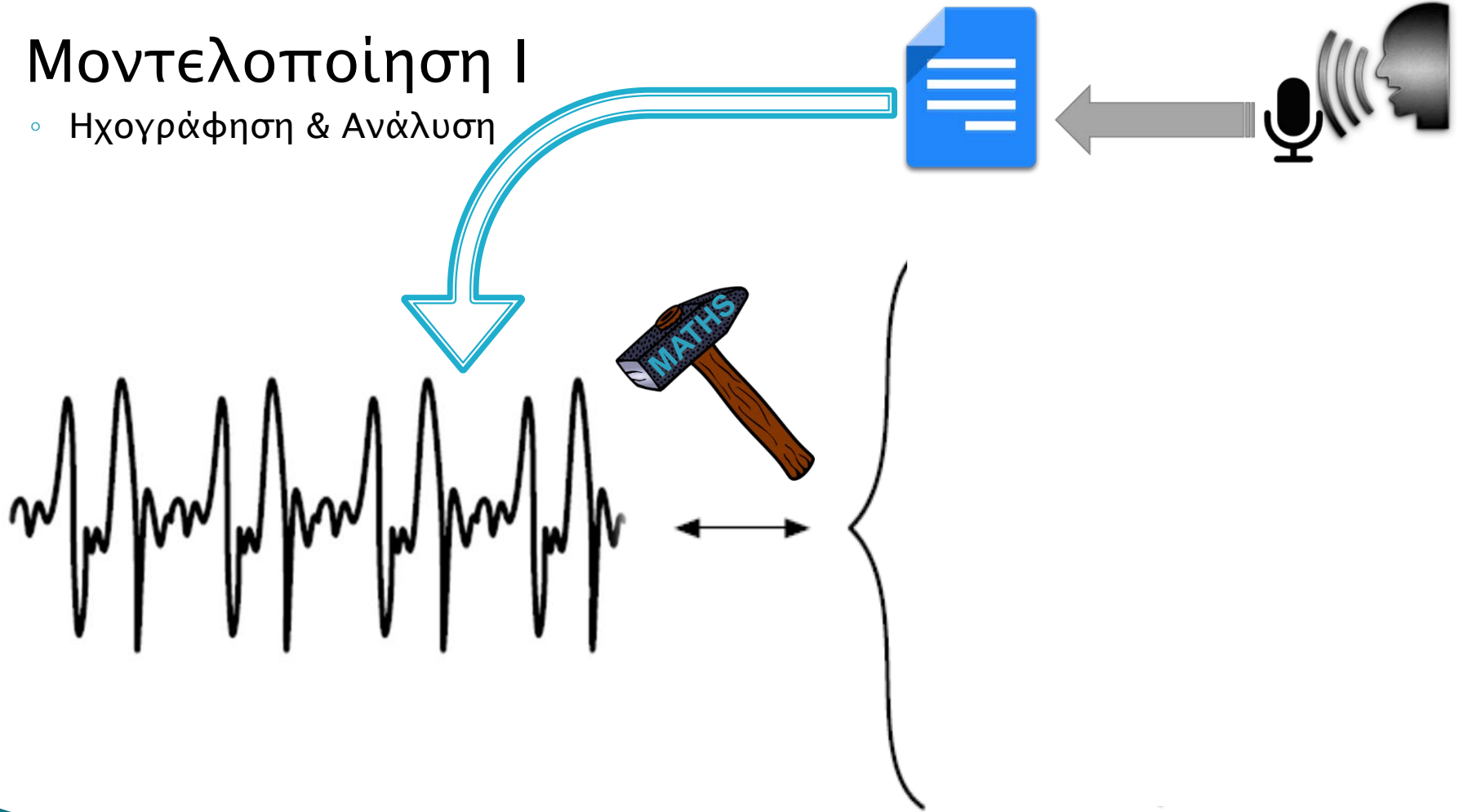
«Πλαστελίνη» φωνής



«Πλαστελίνη» Φωνής

► Μοντελοποίηση I

- Ηχογράφηση & Ανάλυση



«Πλαστελίνη» Φωνής

Τι μπορούμε να κάνουμε;

▶ Σήμα:



▶ Μετασχηματισμός I:



▶ Μετασχηματισμός II:



«Πλαστελίνη» Φωνής

► Μοντελοποίηση II

□ Βασικές Αρχές Παραγωγής Φωνής

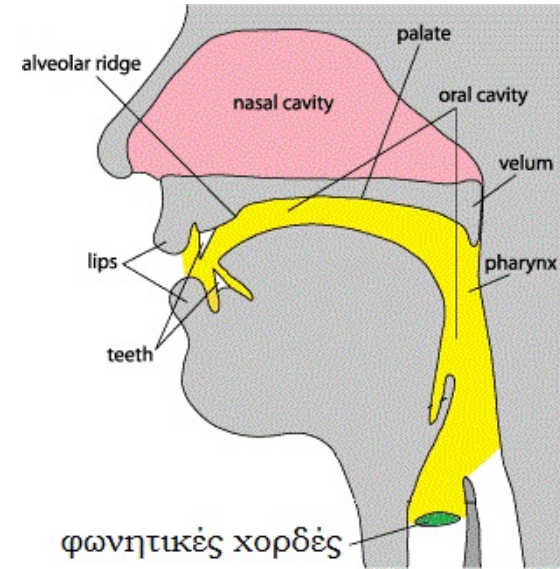
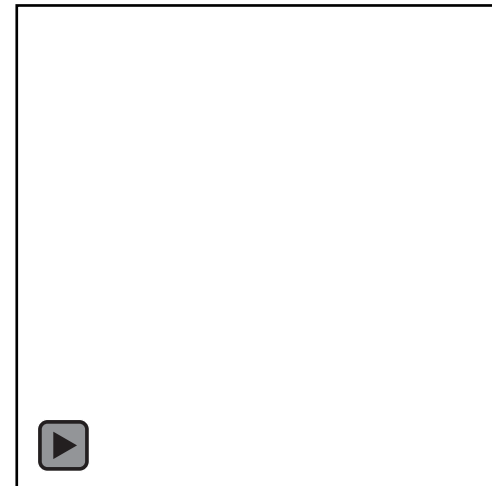
◦ Δυο στάδια:

- **Φωνητικές χορδές** – a.k.a «πηγή»
- **Φωνητική οδό** – a.k.a «φίλτρο»

- Φωνητικές χορδές

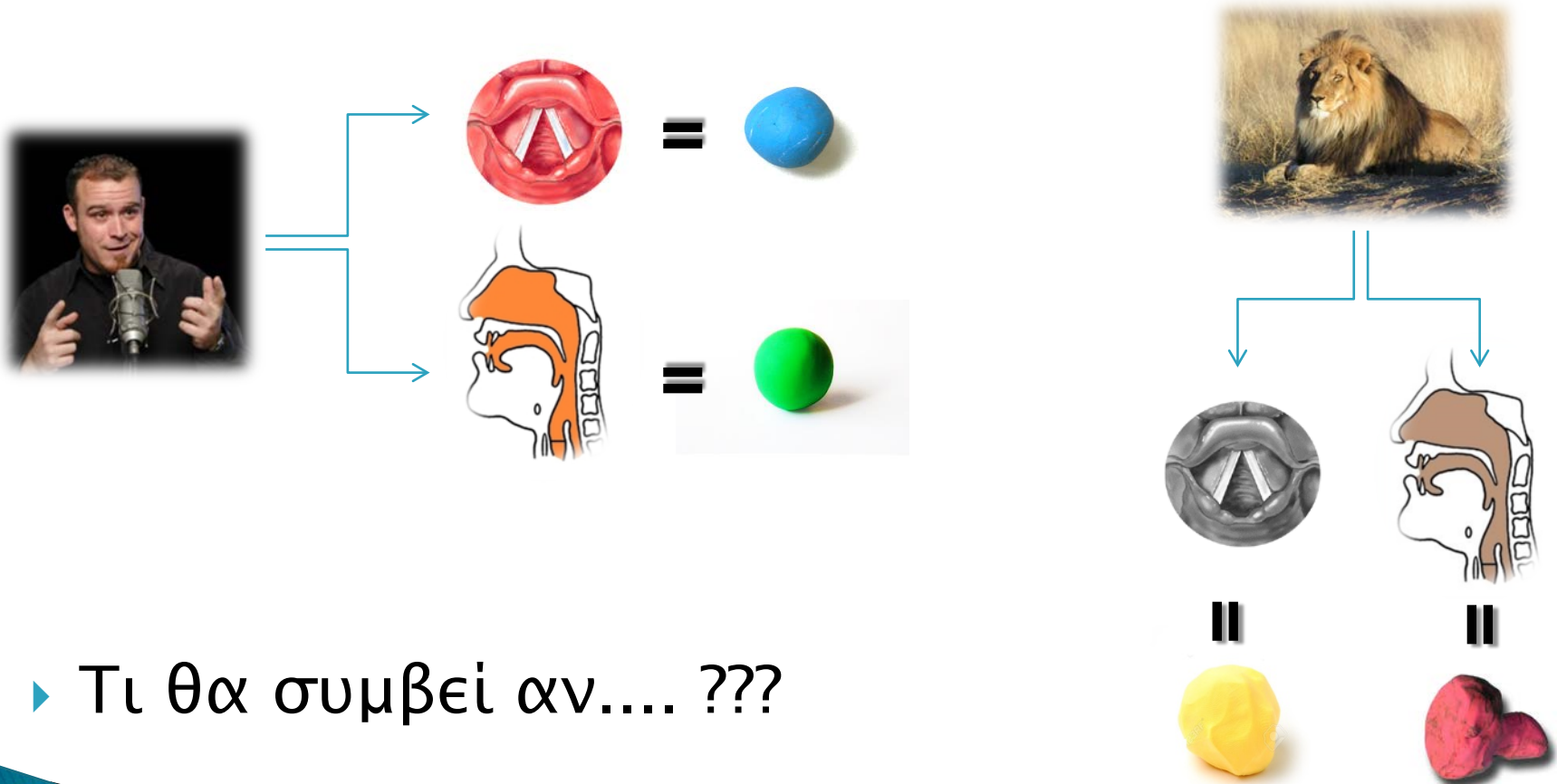


- Φωνητική Οδός (MRI)



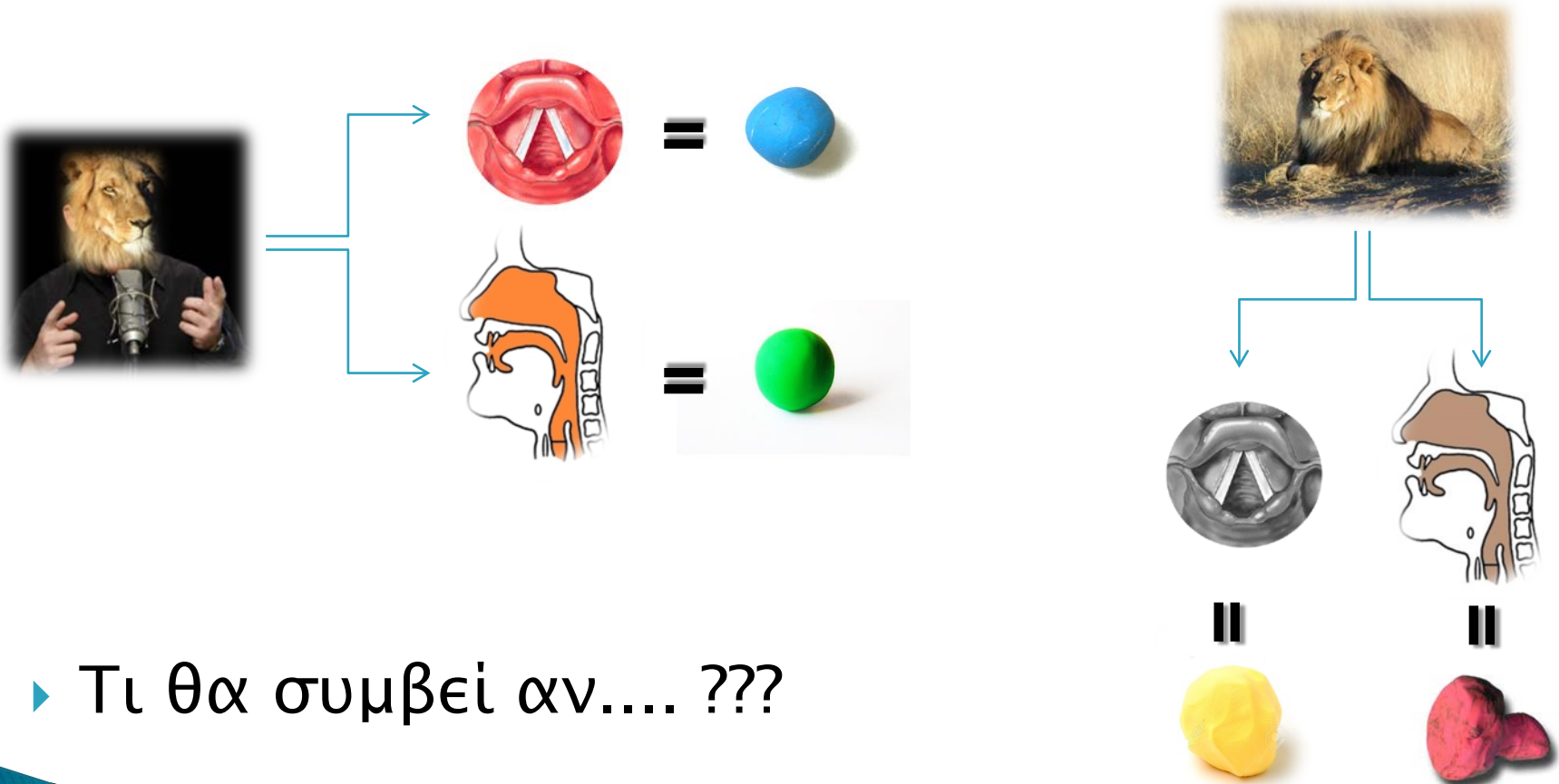
«Πλαστελίνη» Φωνής

► Μοντελοποίηση II



«Πλαστελίνη» Φωνής

► Μοντελοποίηση II



► Τι θα συμβεί αν.... ???

«Πλαστελίνη» Φωνής

Αποτέλεσμα!

▶ Φωνή:



▶ Λιοντάρι:

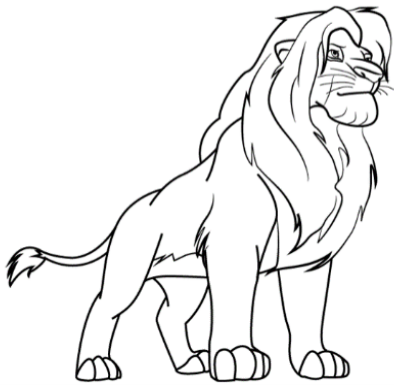
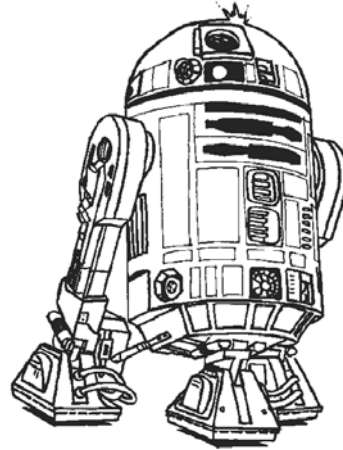
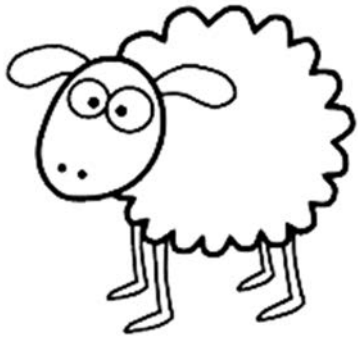


▶ Talking lion:



«Πλαστελίνη» Φωνής

Fun time! 😊



«Πλαστελίνη» Φωνής

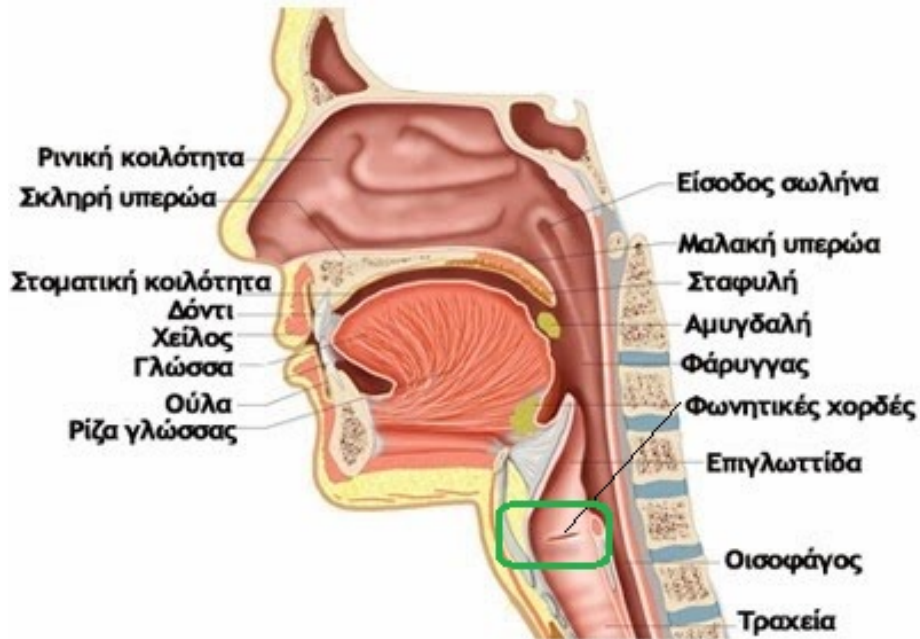




Πρόσεχε πώς μιλάς!!!



Πρόσεχε πώς μιλάς!!!



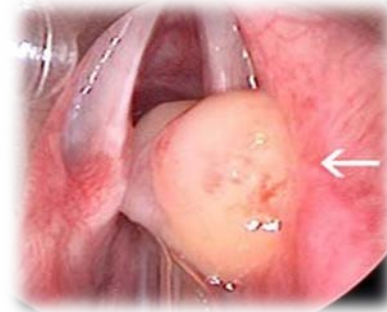
Στατιστική:

- Άνδρες: 120 Hz (περίοδοι/δευτ.)
- Γυναίκες: 210 Hz (περίοδοι/δευτ.)
- Παιδιά: 330 Hz (περίοδοι/δευτ.)

Πρόσεχε πώς μιλάς!!!

Υπερβολική χρήση της φωνής μας

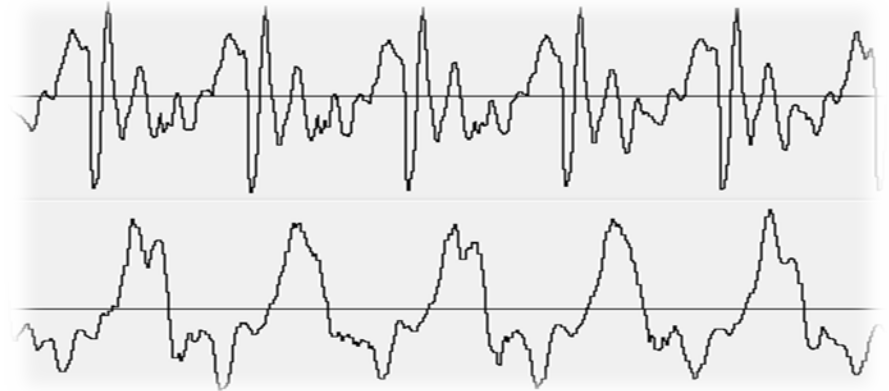
- Πολύποδες
- Κύστες
- Οζίδια



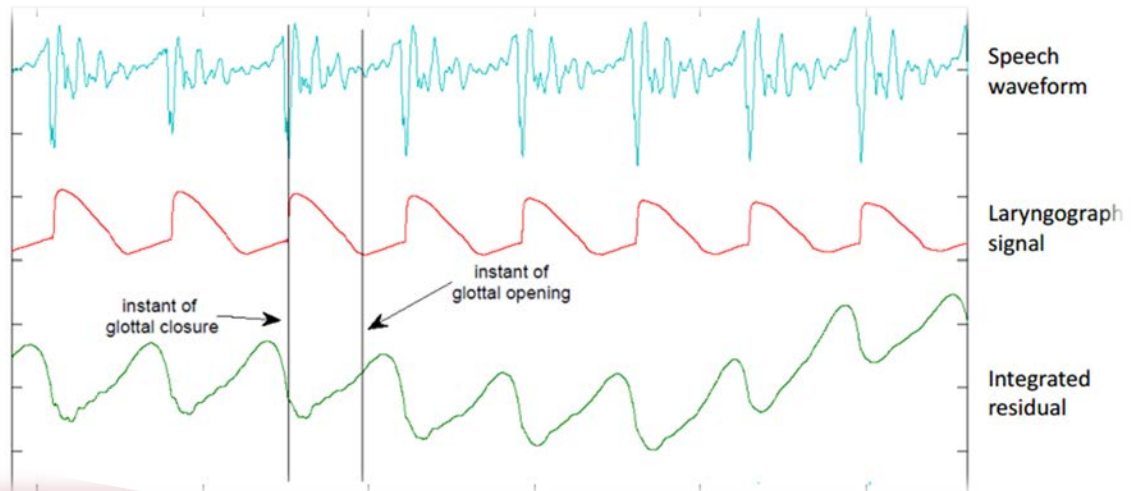
- Μείωση καπνού και του αλκοόλ
- Μην ανταγωνίζεστε τον θόρυβο
- Πίνετε νερό όταν μιλάτε για πολύ ώρα
- Μην μιλάτε όταν δεν μπορείτε



Πρόσεχε πώς μιλάς!!!



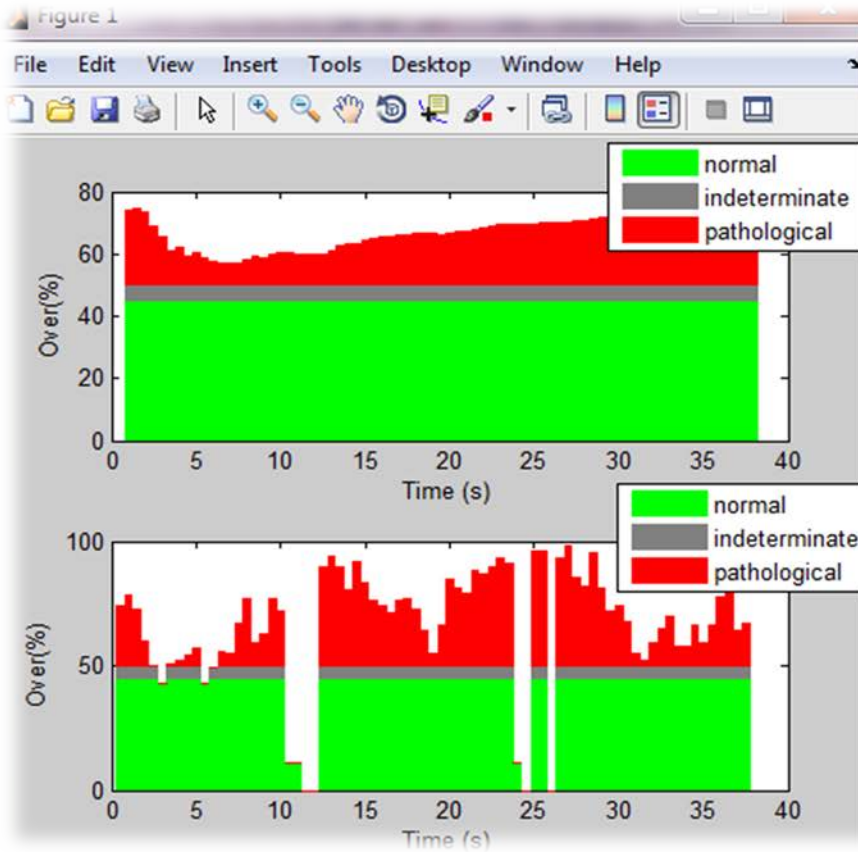
Μπορούμε να ανιχνεύσουμε παθολογίες από απλές ηχογραφήσεις ομιλίας??



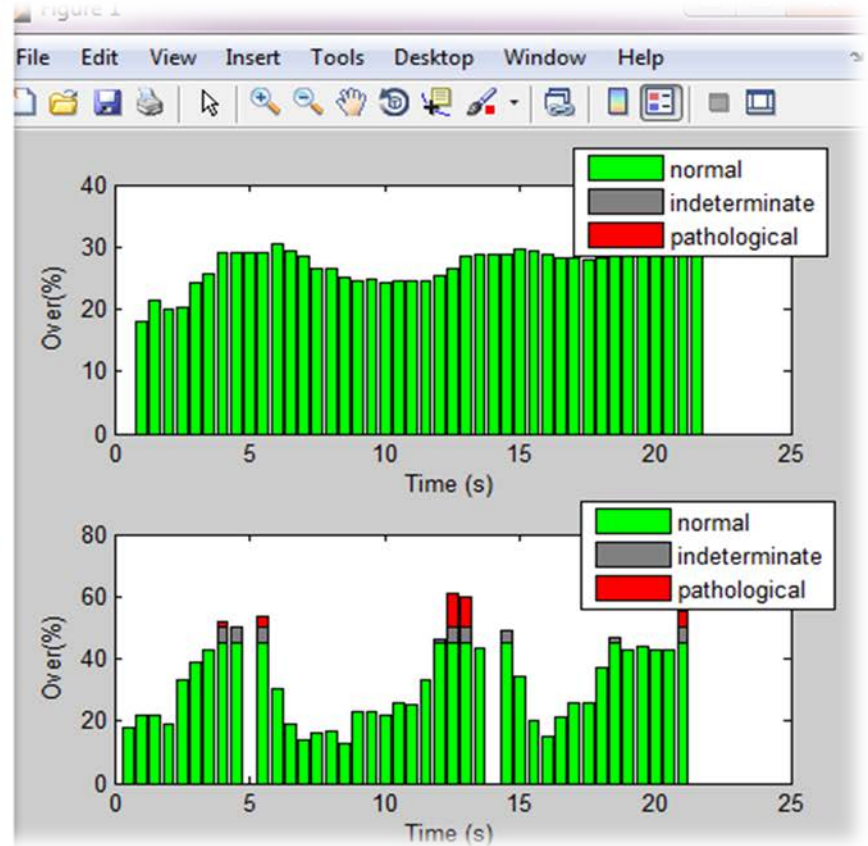
Πρόσεχε πώς μιλάς!!!



Πρόσεχε πώς μιλάς!!!



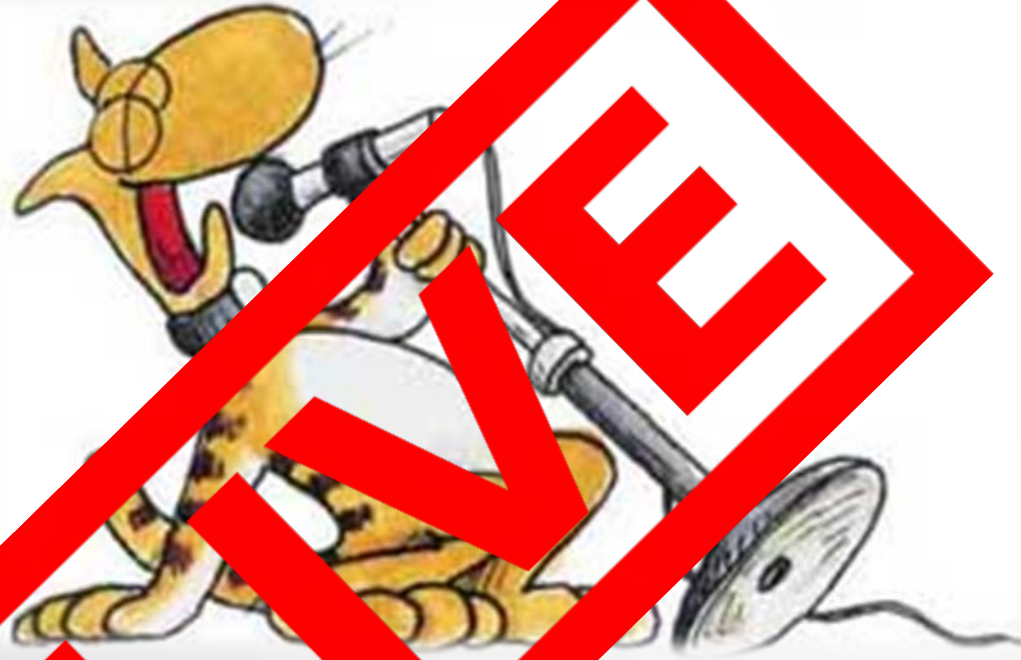
Πριν την επέμβαση:



Μετά την επέμβαση:



Πρόσεχε πώς μιλάς!!!

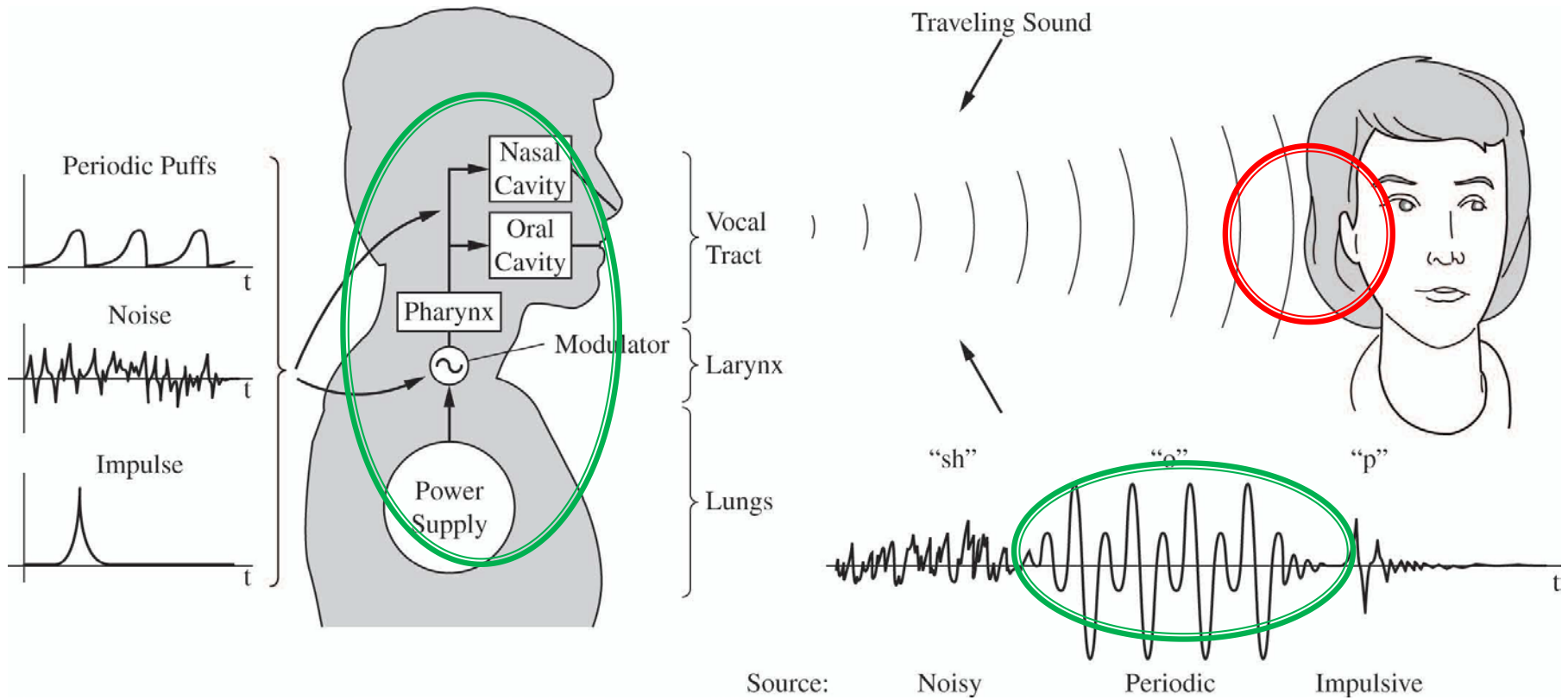




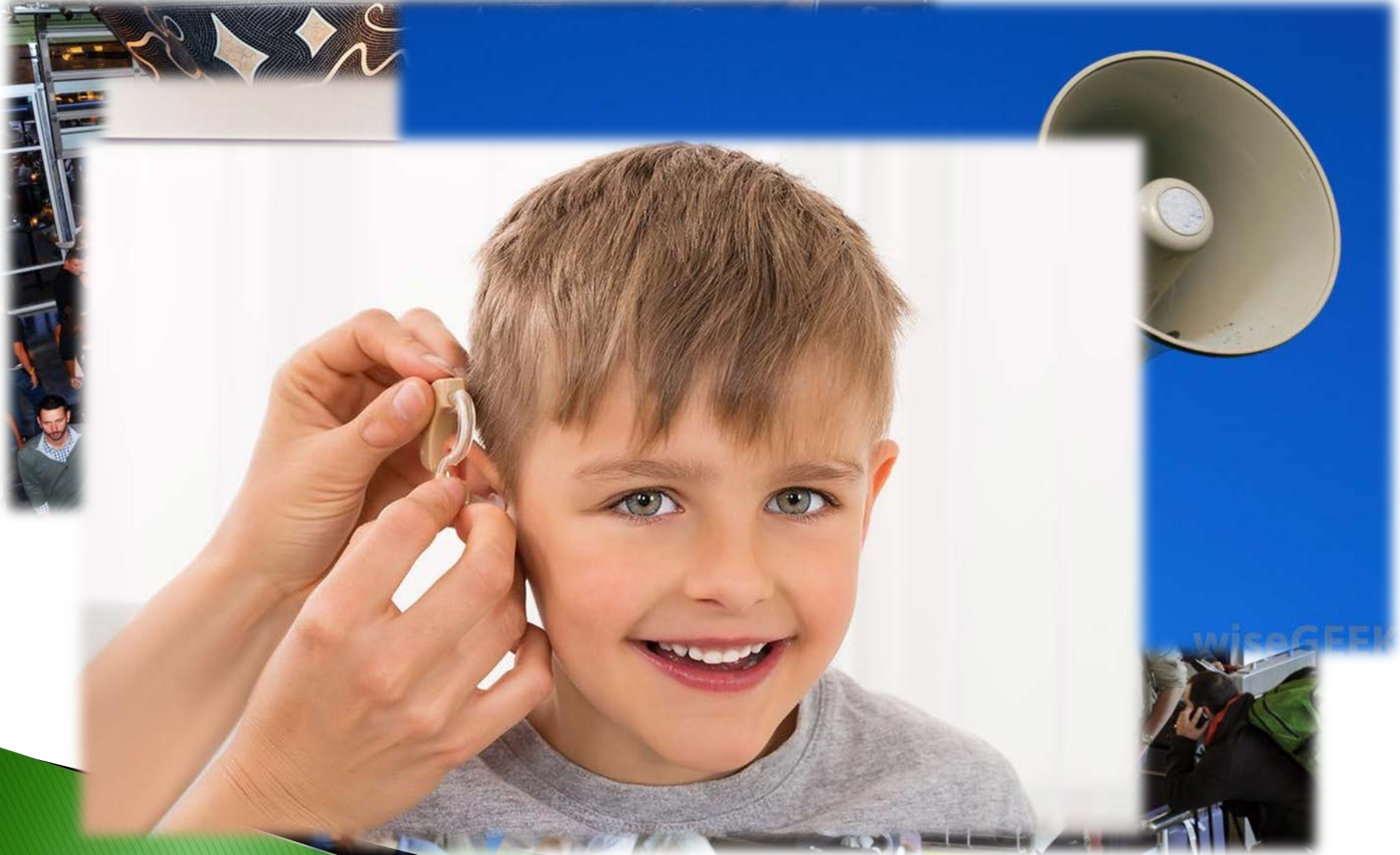
Άκουσέ με καλά!!



Άκουσέ με καλά!!

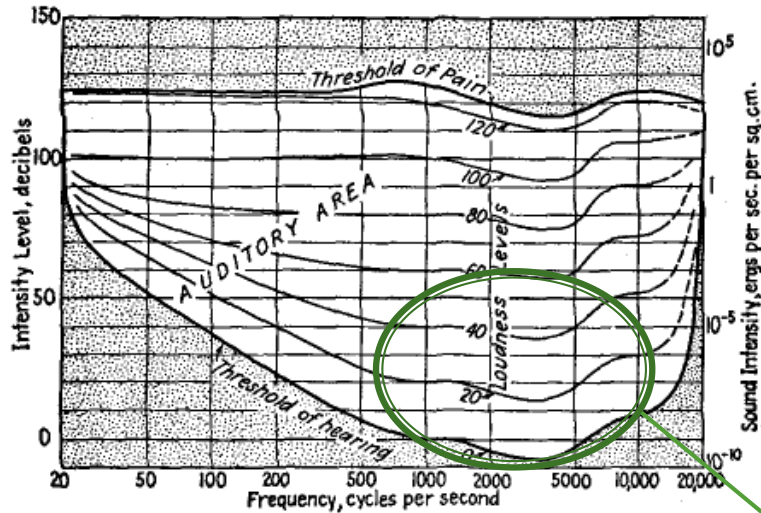


Άκουσέ με καλά!!

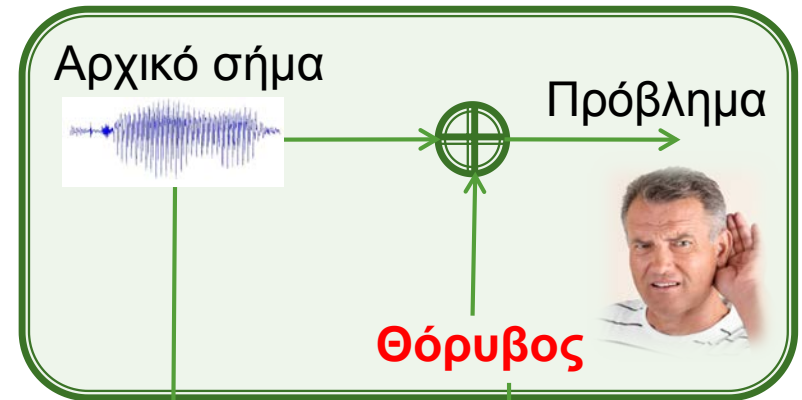
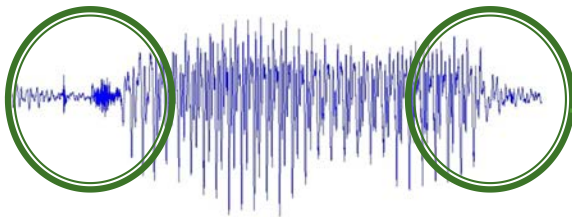


Άκουσέ με καλά!!

1. Ευαισθησία ακοής



2. Ευάλωτα-στο-θόρυβο τμήματα



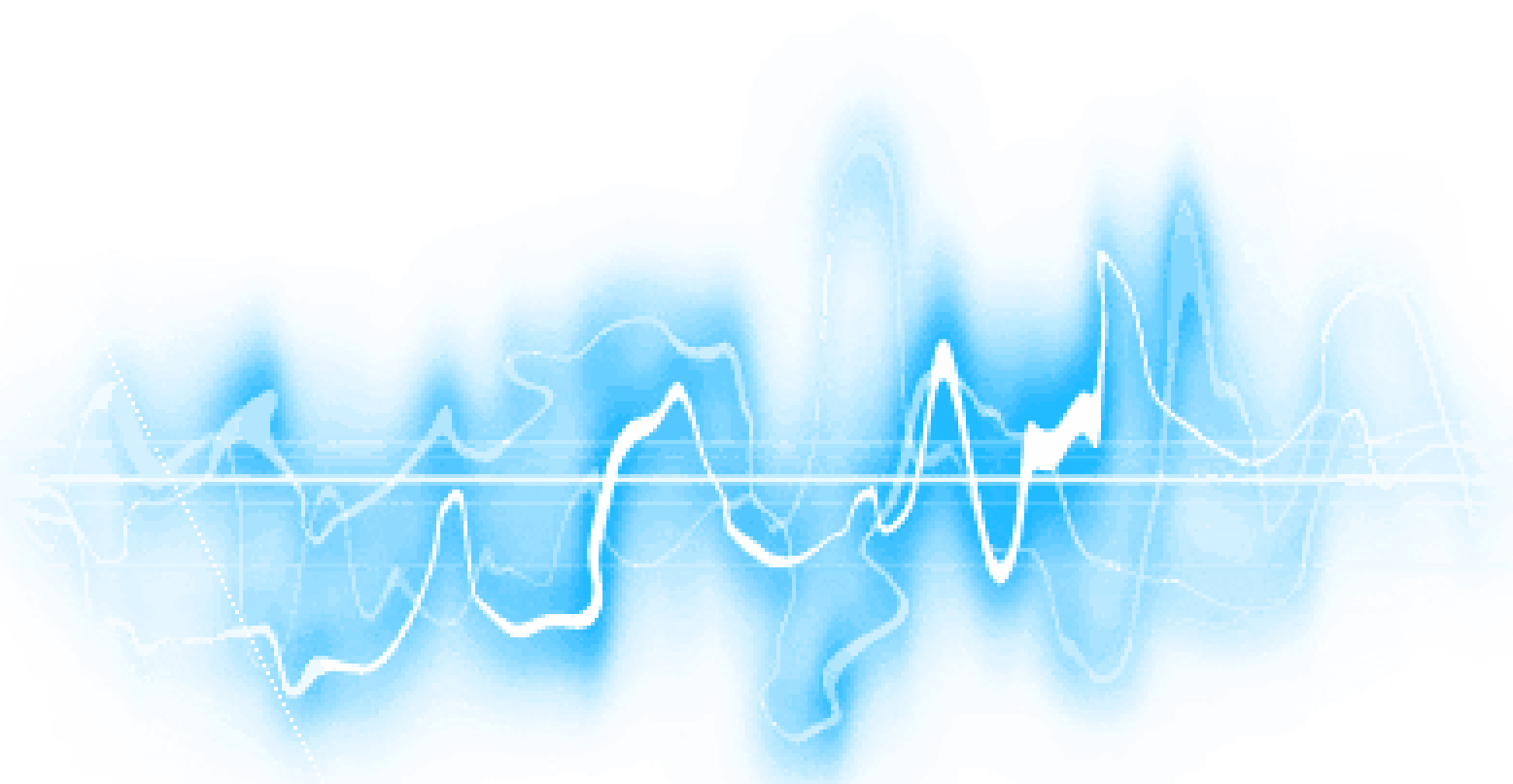
Αλγόριθμος
Βελτίωσης
Καταληπτοτητας

Βελτίωση



Άκουσέ με καλά!!





Γράψε και άκου!



Γράψε και Άκου!

Τεχνητή παραγωγή ανθρώπινης ομιλίας

Text-To-Speech Synthesis (TTS) Σύνθεση φωνής από κείμενο

Εφαρμογές

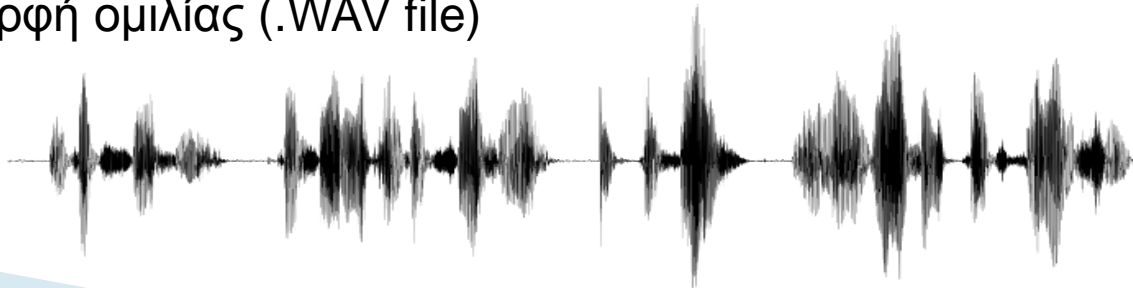
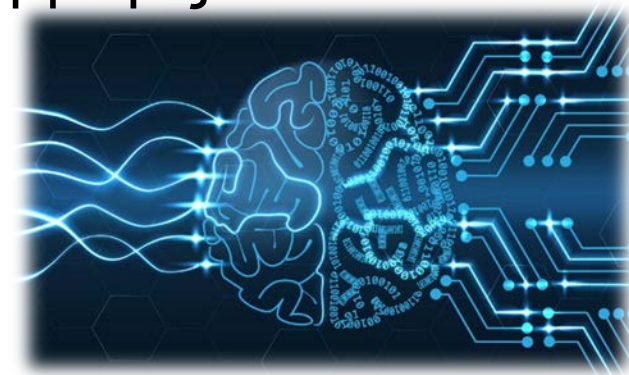
- Αναγνώστες οθόνης για άτομα με αναπηρία ή δυσλεξία
 - Stephen Hawking's voice generator
 - <https://lingojam.com/StephenHawkingVoiceGenerator>
- Παιχνίδια, animations, προϊόντα ψυχαγωγίας, κινηματογράφος
 - <https://www.animaker.com/voice>
- Εκπαιδευτικά εργαλεία για εκμάθηση ξένων γλωσσών
 - [Duolingo](#)
 - [Google Translate](#)
- Διεπαφές επεξεργασίας φυσικής γλώσσας για κινητές συσκευές
 - Cortana (Microsoft), Google Assistant (Google), Alexa (Amazon), Siri (Apple)



Γράψε και Άκου!

Τεχνητή παραγωγή ανθρώπινης ομιλίας

- ▶ Η έκρηξη της βαθιάς μάθησης (deep learning) έχει οδηγήσει σε εκπληκτικές προόδους σε τέτοιες εφαρμογές!
 - Χρειάζεται όμως **πολλά δεδομένα** (πολλές ηχογραφημένες προτάσεις)
- ▶ Δυο στάδια:
 - Εκπαίδευση (training)
 - Το σύστημα «μαθαίνει» την αντιστοίχιση κειμένου & ηχογράφησης
 - Σύνθεση (inference)
 - Το σύστημα παράγει μια νέα πρόταση βάσει κειμένου εισόδου
 - Η έξοδος είναι κυματομορφή ομιλίας (.WAV file)



Γράψε και Άκου!

Τεχνητή παραγωγή ανθρώπινης ομιλίας





Γράψε και Άκου!

Τεχνητή παραγωγή ανθρώπινης ομιλίας



English sample 1

Numbers of ladies were present although the public feeling was much against their attendance.



English sample 2

They could be easily adjusted and were still solid enough to serve as a gun rest



Ελληνικά 1



Ελληνικά 2



Ελληνικά 3



Γράψε και Άκου!

Τεχνητή παραγωγή ανθρώπινης ομιλίας



Πώς επιτυγχάνονται όλες αυτές οι τεχνολογίες που είδατε;

$Y_{(t)} = Y_0 + b \cdot K_2$
 $B = \begin{pmatrix} 3 & 0 & 1 \\ 2 & 1 & 3 \end{pmatrix}$
 $a^2 = b^2 + c^2$
 $\tan \frac{x}{2} = \frac{1 - \cos x}{\sin x} = \frac{\sin x}{1 + \cos x}$
 $\cos x$
 $\sum_{i=0}^n (p_i(x) - y_i)^2$
 $\tan 2x = \frac{2 \tan x}{1 - \tan^2 x}$
 $\tan x = \frac{\sin x}{\cos x}$
 $\lambda x - y + z = 1$
 $x + \lambda y + z = \lambda$
 $x + y + \lambda z = \lambda^2$
 $\int \int \int_M z \, dx \, dy \, dz = \int_0^{\pi/2} \int_0^1 \int_{\frac{1}{2}+r}^1 r \, dr \, d\theta \, dz$
 $\lim_{n \rightarrow \infty} \frac{\sqrt{n^2+1} + n}{\sqrt[3]{3n^2+2n-1}}$
 $2 \arctan x - x = 0, I = (1, 10)$
 $\frac{1}{\sqrt{2}} \sin^4 x \cdot \cos^3 x \, dx$
 $\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1$
 $\frac{\partial z}{\partial x} = 2; \frac{\partial z}{\partial y} = 0$
 $\vec{n} = (F_x; F_y; F_z)$
 $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0$
 $\sin 2x = 2 \sin x \cdot \cos x$
 $|z| = \sqrt{a^2 + b^2}$
 $\sqrt{2F} = 16 - x^2 + 16y^2 = 4z > 0$
 $Y = \frac{b}{\sin \alpha} = \frac{c}{\sin \gamma}$
 $y = \sqrt[3]{x+1}; x = \tan t$
 $x_y = \begin{pmatrix} \alpha + \beta + \gamma \\ \alpha \\ \beta \end{pmatrix}$
 $(1+e^x) \gamma' = e^x$
 $y(1) = 1$
 $\cos 2x = \cos^2 x - \sin^2 x$
 $\sin^2 x + \cos^2 x = 1$
 $A+B+C=8$
 $-3A-7B+2C=-10,3$
 $-18A+6B-3C=15$
 $C = \begin{pmatrix} 0,1 \\ 1,0 \end{pmatrix}$
 $f(x) = 2^{-x} + 1, \epsilon = 0,005$
 $e^{-xy} z = e; A \in [0, e, 1]$
 $\lim_{x \rightarrow 0} \frac{e^{2x} - 1}{5x} = \frac{2}{5}$
 $b \neq \pm b \neq 0; \rho \neq 0$

Μαθηματικά

Επιστήμη Η/Υ



Ερωτήσεις?

Αν θέλετε να μάθετε περισσότερα,
επισκεφτείτε το

<http://www.csd.uoc.gr/~sspl>



ή ρωτήστε ό,τι θέλετε στο

kafentz@csd.uoc.gr





Ευχαριστώ για την
προσοχή σας!