



University of Athens
Department of Informatics

ΔΗΜΟΣΘΕΝΗΣ: Μία νέα προσέγγιση για την παραγωγή βελτιωμένης συνθετικής ομιλίας από ηλεκτρονικά έγγραφα

Γεράσιμος Ξύδας

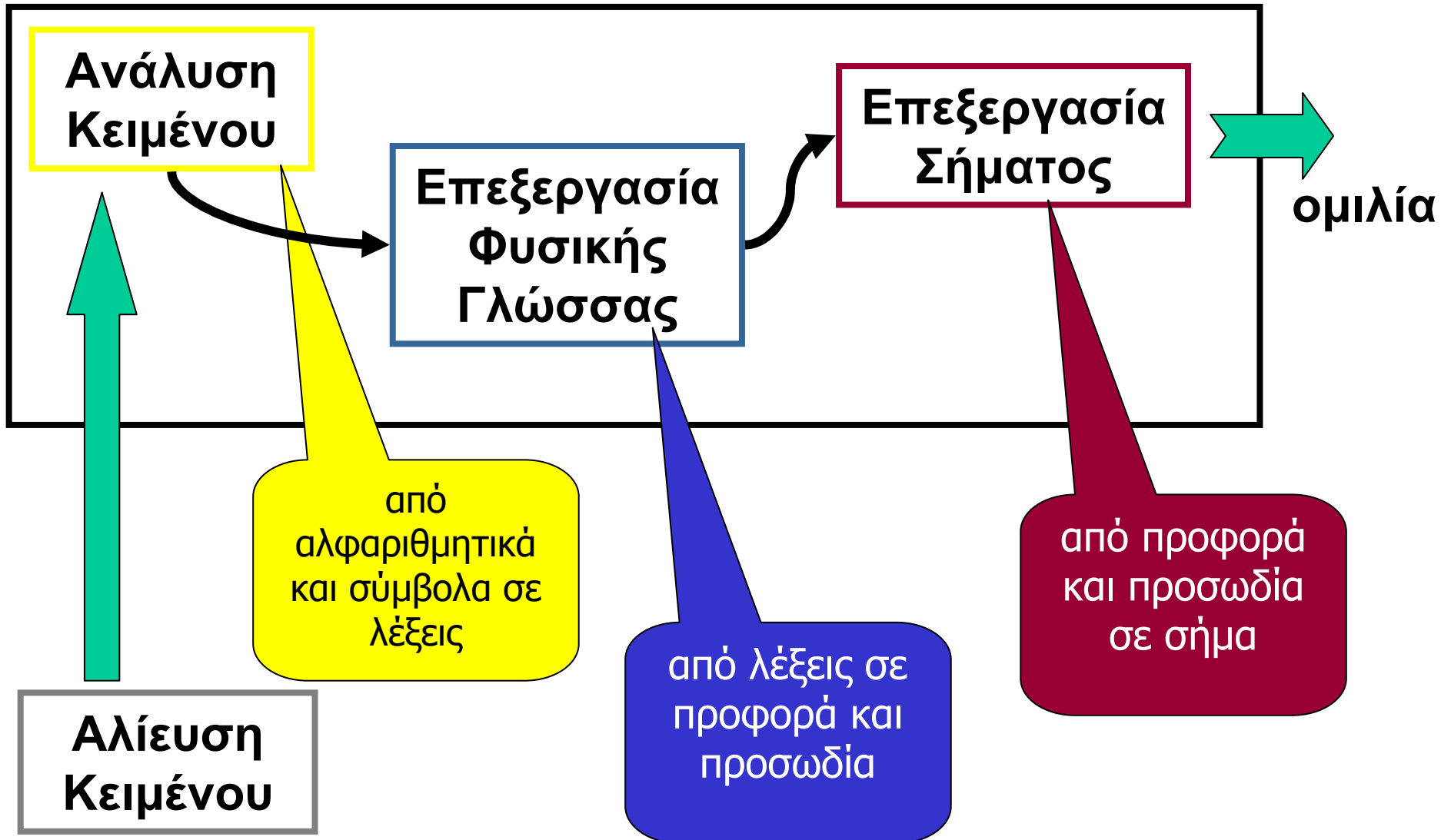
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών
Τμήμα Πληροφορικής και Τηλεπικοινωνιών

gxidas@di.uoa.gr

Δομή παρουσίασης

- Μετατροπή κειμένου σε ομιλία
- ΔΗΜΟΣΘΕΝΗΣ
- Φωνητική αναπαράσταση ηλεκτρονικών εγγράφων

Μετατροπή Κειμένου σε Ομιλία



Ανάλυση κειμένου (προ-επεξεργασία)

- Εντοπισμός και ανάπτυξη κωδικοποιημένων ή συνεπτυγμένων μη-λεξικών πληροφοριών
 - Αριθμητικά
 - «των 102 Ολυμπιακών Αγώνων»
 - Ημερομηνίες και ώρες
 - «22/2/2002», «16:45», «22-2-2002»
 - Συντμήσεις
 - «η κ. Πέτρου», «1,25 μ.»
 - Ακρώνυμα, αρκτικόλεξα
 - ΙΚΑ, Ι.Κ.Α.
 - Ειδικές περιπτώσεις
 - Greeklish, χημικά σύμβολα, ...

Ανάλυση κειμένου

- Τεχνικές
 - Λεξικό
 - 1-1 αντιστοίχιση
 - Αποτυγχάνει σε περιπτώσεις αμφισημιών:
 - «1,2 εκ.» (εκατομμύρια ή εκατοστά;)
 - «Δ.» (δυτικά, δυτικός, δύση, Διονύσης, Δημήτρης, Δημοσθένης,...)
 - «κ.» (κύριος, κυρίου, κυρία,...)
 - «12/2» (δώδεκα δεύτερα, δώδεκα δευτέρου)
 - «5-3» (πέντε τρία, πέντε τρίτου, πέντε μείον τρία, πέντε μέχρι τρία)
 - Αυτόματα Πεπερασμένων Καταστάσεων
 - Έλεγχος του περιβάλλοντος μίας οντότητας
 - Εκπαιδευόμενα ή εμπειρικά
 - Υβριδικές
 - Μοντέλα κανόνων ή εκπαιδευόμενα συνοδεύουν το λεξικό για τις περιπτώσεις αμφισημιών

Επεξεργασία Φυσικής Γλώσσας

- Εξαγωγή συντακτικής, εννοιολογικής και φωνολογικής πληροφορίας με σκοπό την φωνητική αναπαράσταση του κειμένου και την πρόβλεψη της κατάλληλης προσωδίας.
- Παράγει μία ακολουθία φωνημάτων (πληροφορία) και τρία διανύσματα: τονισμού, διάρκειας και έντασης (μέτα-πληροφορία)

Μετατροπή σε φωνήματα

- Αντιστοίχιση της γραπτής αναπαράστασης με την προφορική.
- Μετατροπή γραμμάτων σε διεθνές φωνητικό αλφάβητο (IPA, SAMPA)

«*Τα παιδιά είδαν για μιά στιγμή
τον Κώστα και την Χαρά σε μία ταβέρνα.*»

«*ta peðl̩à ìðan ja m̩ta stiymì
ton ɣɔsta ce tiŋ xaɾà se mià tavèrna* »

Προσωδία

- Η προσωδία διαμορφώνει το μήνυμα των δεδομένων της ομιλίας.
- «Δεν είναι το τι λες αλλά ο τρόπος που το λές»
⇒ Η προσωδία είναι η μετα-πληροφορία των φωνημάτων.
 - «Σήμερα περάσαμε ωραία»
 - «Σήμερα περάσαμε ωραία»
 - «Σήμερα περάσαμε ωραία»
 - «Σήμερα περάσαμε ωραία»
- 3 διανύσματα: τόνος, διάρκεια και ένταση.

ΔΗΜΟΣΘΕΝΗΣ

- Αναπτύχθηκε εξ'ολοκλήρου στο εργαστήριο της Ομάδας Φωνής του ΕΚΠΑ.
- Σύνθετο σύστημα, κατάλληλο για έρευνα σε διάφορα επίπεδα.
- Βαθμωτό: μπορεί να μικρύνει σε όγκο και απαιτήσεις CPU ανάλογα με την εφαρμογή ή να μεγαλώσει ώστε να πραγματοποιούνται πειράματα.
- Ανοικτό, πολυγλωσσικό (multilingual και polyglot) σύστημα.
- Αρθρωτή αρχιτεκτονική => Επεκτάσιμο.
- Κάθε άρθρωμα ρυθμίζεται εξωτερικά από τον σχεδιαστή ή τον τελικό χρήστη.

ΔΗΜΟΣΘΕΝΗΣ

- Πυρήνας βασισμένος σε Γράφους Ετερογεννών Σχέσεων (HRG) (Black et. al, 2000).
- Συμβατότητα με το Festival και τα εργαλεία αυτού.
- Η διαδικασία της σύνθεσης ορίζεται από Φωνητικούς Χαρακτήρες (VoiceFonts)
 - Φωνή
 - Λεξικά
 - Αλυσίδα αρθρωμάτων και παράμετροι αυτών.

Αρθρώματα - 1

- Ανάλυση κειμένου
 - Αριθμητικά, ημερομηνίες, ώρες, ακρώνυμα, συντμήσεις
 - Υποστήριξη αγγλικών (120.000 λέξεις)
- Μετατροπή σε φωνήματα
 - Μηχανή αυτομάτων (568 + 123 περιπτώσεις)
- Γραμματική ανάλυση: εύρεση μερών του λόγου
 - Αντίστροφο λεξικό (4.000 κανόνες + 3.000 εξαιρέσεις)
 - Λεξικό (1.070.000 λέξεις)



Αρθρώματα - 2

- Προσωδία

- Διάρκεια (επίπεδο φωνημάτων)

- Εκπαιδευόμενα μοντέλα 3 καταστάσεων (+ρύθμιση)
 - Εκπαιδευόμενα δέντρα παλινδρόμησης (CART), για συγκεκριμένα πεδία κειμένων.

- Επιτονισμός (πρόβλεψη επιτονικών γεγονότων)

- Μοντέλο 58-προτύπων (πρώτυπη πρόταση για τα Ελληνικά)
 - H S1 T1 S2 T2 S3 T3 T

- Εκπαιδευόμενα CART δέντρα αποφάσεων για τόνους συλλαβής και φράσεων

- Επιτονισμός (απόδοση καμπύλης F0)

- Μοντέλο Fujisaki
 - Μοντέλο γραμμικής πρόβλεψης f2b (Black & Hunt, 1996)

Αρθρώματα - 3

- Παραγωγή σήματος ομιλίας
 - Formant συνθέτης
 - Συνθέτες διφώνων
 - Αντρικές και γυναικείες φωνές (1081 δίφωνα, αλφάβητο UoA-TtS-PA)
 - Εργαλείο αναγνώρισης φωνής (HTK) για αυτοματοποίηση της διαδικασίας παραγωγής νέων φωνών.
 - LPC σύνθεση
 - PSOLA, MBROLA σύνθεση
 - 2 Ελληνικές MBROLA φωνές που αναπτύξαμε διατίθενται ελεύθερα από το site του ΔΗΜΟΣΘΕΝΗ και του MBROLA.
 - Unit Selection

ΔΗΜΟΣΘΕΝΗΣ

- Υλοποίηση
 - C/C++ from scratch, thread safe
 - Αποδοτική υλοποίηση, 200*realtime (P4, 1.6GHz)
 - Win32, αλλά γίνεται προσπάθεια μετεγγραφής σε ANSI C.
 - ActiveX control (απλές διεπαφές)
 - Server (socket based), Win32 και Linux clients

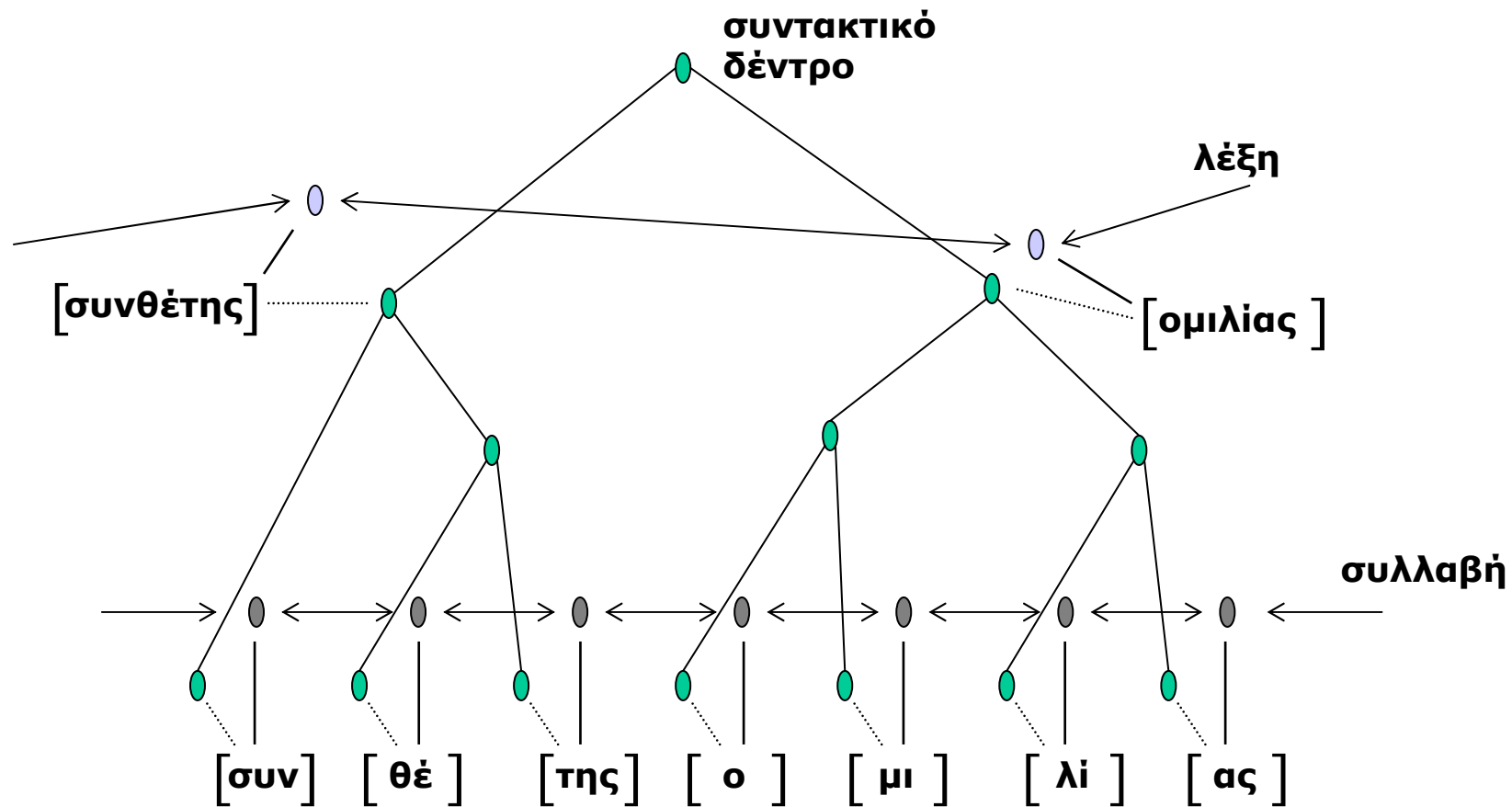
Δημοσθένης

- Διαθέσιμος από την σελίδα

<http://demosthenes.di.uoa.gr>

(προς ανανέωση...)

Γράφος Ετερογενών Σχέσεων (HRG)



HRG

- Κάθε κόμβος ανήκει σε μία ή περισσότερες σχέσεις.
- Τα περιεχόμενα του κάθε κόμβου είναι λίστες με πρωτογενείς τύπους δεδομένων, με δείκτες σε αντικείμενα οποιουδήποτε τύπου ή ακόμα και δείκτες σε συναρτήσεις.
- Είναι δυνατή η πλοήγηση σε όλες τις διαστάσεις του γράφου, ακολουθώντας σχετικά ή απόλυτα μονοπάτια.

HRG

- Όλη η πληροφορία

- Συντακτική

- Φωνολογική

- Ακουστική

- Γλωσσολογική

- Σήματα ομιλίας (π.χ. Δίφωνα)

που γεννάνται κατά την μετατροπή του κειμένου σε ομιλία αποθηκεύεται και χειρίζεται μέσω αυτών των γράφων.

Φωνητική αναπαράσταση ηλεκτρονικών εγγράφων

- Το πρόβλημα:
 - Τα ηλεκτρονικά έγγραφα φέρουν κωδικοποιημένη πληροφορία σημαντική για την σημασιολογία του περιεχόμενου κειμένου.
 - Η πληροφορία αυτή δεν αξιοποιείται από τα συστήματα μετατροπής κειμένου σε ομιλία.
 - Το αποτέλεσμα είναι τα κείμενα αυτά να αποδίδονται μερικές φορές με μειωμένη σημασιολογία.

Η φύση των ηλεκτρονικών εγγράφων

- Ενδιαφερόμαστε για «machine readable» έγγραφα (όχι κωδικοποιημένες μορφές π.χ. jpeg) που περιέχουν κείμενο+μέτα πληροφορία
- Κείμενο
 - Επεξεργάζεται από τα συστήματα κειμένου σε ομιλία.
 - Η ακουστική αναπαράσταση ποικίλει ανάλογα με τις τεχνολογίες και τους πόρους του κάθε συστήματος.
 - Δεν υπάρχει ποτέ κοινή αντιμετώπιση των κειμένων.
Διαφορετικό σύστημα = διαφορετική ποιότητα και εκφορά.
- Μέτα-πληροφορία σχετική με το κείμενο
 - Αγνοείται από τα συστήματα κειμένου σε ομιλία
 - Προβλήματα: αδυναμία σωστής απόδοσης του κειμένου (MathML, HTML, ...)

Κατηγορίες μετα-πληροφορίας

- **Οδηγίες οπτικοποίησης:** έντονα γράμματα, πλάγια, πίνακες, bullets, μέγεθος γραμμάτων κλπ (π.χ. HTML, MS-Word)
- **Οδηγίες δομής:** επικεφαλίδα, τίτλος, παράγραφος, record content (π.χ. HTML, XML, SQL), σύνθετα συνολα (π.χ. μαθηματικές πράξεις – MathML)
- **Γλωσσολογική πληροφορία:** γραμματική, σύνταξη, μορφολογία, ρητορικές σχέσεις (π.χ. SOLE-ML, plain κείμενο)
- **Δομή κειμένου:** παρενθέσεις, σημάδια, κλπ
- **Οδηγίες ομιλίας:** <prosody>, <emp>, <rate>, <pitch> κλπ (π.χ. SABLE, VoiceXML, SSML, ACSS)

Ποιά είναι η ακουστική αξία της μετα-πληροφορίας;

- Ανάλογα με τον τύπο του εγγράφου αλλάζει και ο ρόλος της μετα-πληροφορίας.
 - Σε κείμενα με οπτική διαμόρφωση βοηθάει στον εντοπισμό εστιακών (focus) σημείων.
 - Σε έγγραφα με ιεραρχική δομή βοηθάει στην απόδοση του σωστού νοήματος προς τον ακροατή.
- Η μετα-πληροφορία μπορεί να αποδοθεί με έλεγχο προσωδιακών χαρακτηριστικών, αλλαγή ομιλητή, παρεμβολή άλλων ήχων.

[Home](#)

[Document](#)

[FestVox Download](#)

[Festival Download](#)

[Voice Demos](#)

[Limited Domain](#)

[Example Databases](#)

[Mailing Lists](#)

[Search Documents](#)

[Contributed parts](#)

[Links](#)

[Contact](#)

Welcome to festvox.org

This project is part of the work at [Carnegie Mellon University's speech group](#) aimed at advancing the state of Speech Synthesis.

The Festvox project aims to make the building of new synthetic voices more systemic and better documented, making it possible for anyone to build a new voice. Specifically we offer:

- [Documentation](#), including scripts explaining the background and specifics for building new voices for speech synthesis in new and supported languages.
- Specific scripts to build new voices in supported languages, such as US and UK English.
- Aids to building synthetic voices for [limited domains](#)
- [Example speech databases](#) to help building new voices.
- Links, demos and a repository for new voices

The documentation, tools and dependent software are all free without restriction (commercial or otherwise). Licencing of voices built by these techniques are the responsibility of the builders.

This work is firmly grounded within Edinburgh University's [Festival Speech Synthesis System](#) and Carnegie Mellon University's [small footprint Flite synthesis engine](#)

Requirements for building a voice

Note the techniques and processes described here do not guarantee that you'll end up with a high quality acceptable voice, but with a little care you can likely build a new synthesis voice in a supported language in a few days, or in a new language in a few weeks (more or less depending on the complexity of the language, and the desired quality).

You will need:

- To read the [documentation](#)
- A Unix machine (e.g. Linux, FreeBSD, Solaris, etc) with working audio i/o. This may work on other

- [Home](#)
- [Document](#)
- [FestVox Download](#)
- [Festival Download](#)
- [Voice Demos](#)
- [Limited Domain](#)
- [Example Databases](#)
- [Mailing Lists](#)
- [Search Documents](#)
- [Contributed parts](#)
- [Links](#)
- [Contact](#)

Welcome to festvox.org

This project is part of the work at [Carnegie Mellon University's speech group](#) aimed at advancing the state of Speech Synthesis.

The Festvox project aims to make the building of new synthetic voices more systemic and better documented, making it possible for anyone to build a new voice. Specifically we offer:

- [Documentation](#), including scripts explaining the background and specifics for building new voices for speech synthesis in new and supported languages
- Specific scripts to build new voices in supported languages, such as US and UK English.
- Aids to building synthetic voices for [limited domains](#)
- [Example speech databases](#) to help building new voices.
- Links, demos and a repository for new voices

The documentation, tools and dependent software are all free without restriction (commercial or otherwise). Licencing of voices built by these techniques are the responsibility of the builders.

This work is firmly grounded within Edinburgh University's [Festival Speech Synthesis System](#) and Carnegie Mellon University's [small footprint Flite synthesis engine](#)

Requirements for building a voice

Note the techniques and processes described here do not guarantee that you'll end up with a high quality acceptable voice, but with a little care you can likely build a new synthesis voice in a supported language in a few days, or in a new language in a few weeks (more or less depending on the complexity of the language, and the desired quality).

You will need:

- To read the [documentation](#)
- A Unix machine (e.g. Linux, FreeBSD, Solaris, etc) with working audio i/o. This may work on other

W3C

- Οι συστάσεις της W3C (ACSS, VoiceXML) αποτελούν οδηγό για κάποιον που θέλει να συμπεριλάβει συγκεκριμένα φωνητικά χαρακτηριστικά στα έγγραφά του.
- Όμως:
 - η πλειονότητα των κειμένων είναι και εξακολουθεί να γράφεται χωρίς φωνητική αντίληψη (εξοικίωση συγγραφέων με οπτικά εφφέ).
 - Οι συστάσεις αφορούν το Web ή έγγραφα σε SGML μορφή.
 - Δεν χειρίζονται φωνητικά της συστάδες με την ιεραρχική δομή με την οποία έχουν γραφεί (ακολουθεί παράδειγμα).

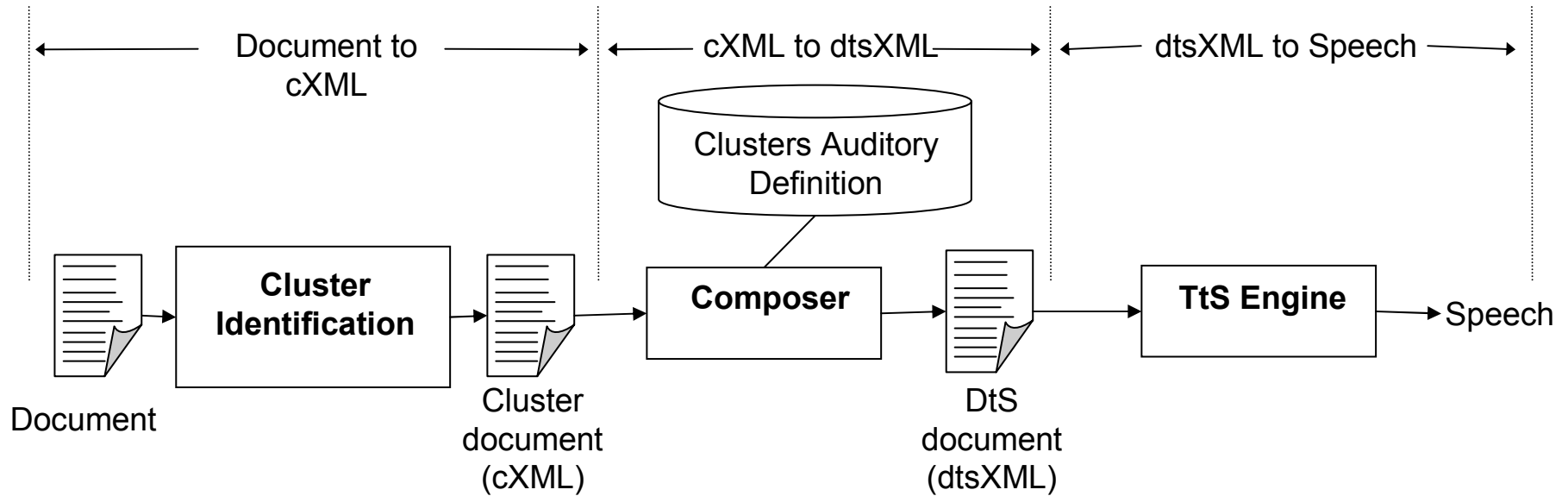
Σκοπός

- Να προσδώσουμε ελεγχόμενα προσωδιακά χαρακτηριστικά σε συστάδες μέτα-πληροφορίας, ώστε αυτές να αποκτήσουν διακριτή ακουστική αναπαράσταση κατά την μετατροπή του εγγράφου σε ομιλία.
- Να προσδώσουμε σε έγγραφα χωρίς φωνητική αντίληψη ένα φωνητικό τρόπο αναπαράστασης, συμβατό ή όχι με W3C σε περιπτώσεις που δεν γίνεται.
- Να ενισχύσουμε την «μνήμη» του ακροατή κατά την ακρόαση ενός εγγράφου (στην οπτική μορφή ο αναγνώστης μπορεί γρήγορα να μεταφερθεί μπρος-πίσω. Στην ακουστική...;).

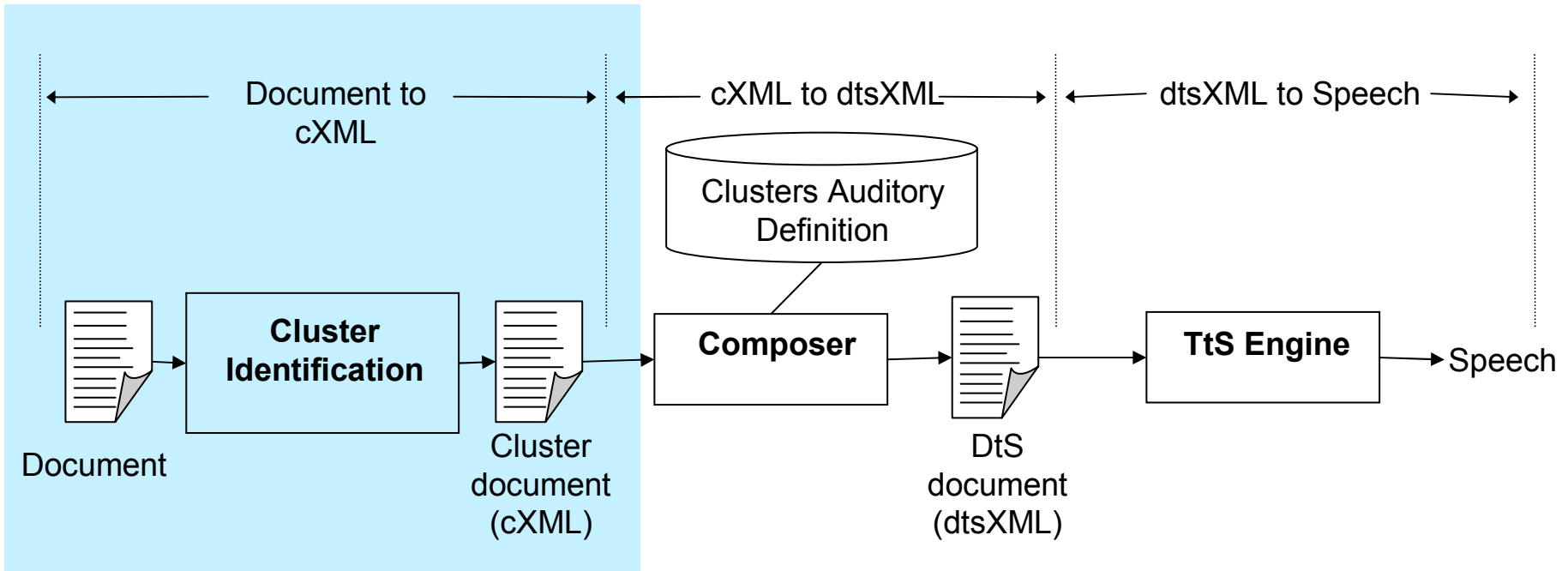
Document-to-Speech (D-t-S)

- Μία προσέγγιση για την ακουστική αξιοποίηση της μετα-πληροφορίας στα κείμενα.
- Χαρακτηριστικά
 - Ανοικτό και επεκτάσιμο
 - Δύναται να προσαρμοστεί «σφικτά» σε συγκεκριμένες υλοποιήσεις TtS (για καλύτερα αποτελέσματα) ή «χαλαρά» σε οποιαδήποτε υλοποίηση (*).
 - Παρέχει μία script γλώσσα για τον έλεγχο της προσωδίας.
 - Βασισμένο σε XML (θετικά/αρνητικά).
 - Ενεργοποιεί ένα φωνητικό/ακουστικό κανάλι επικοινωνίας σε έγγραφα χωρίς φωνητική αντίληψη.

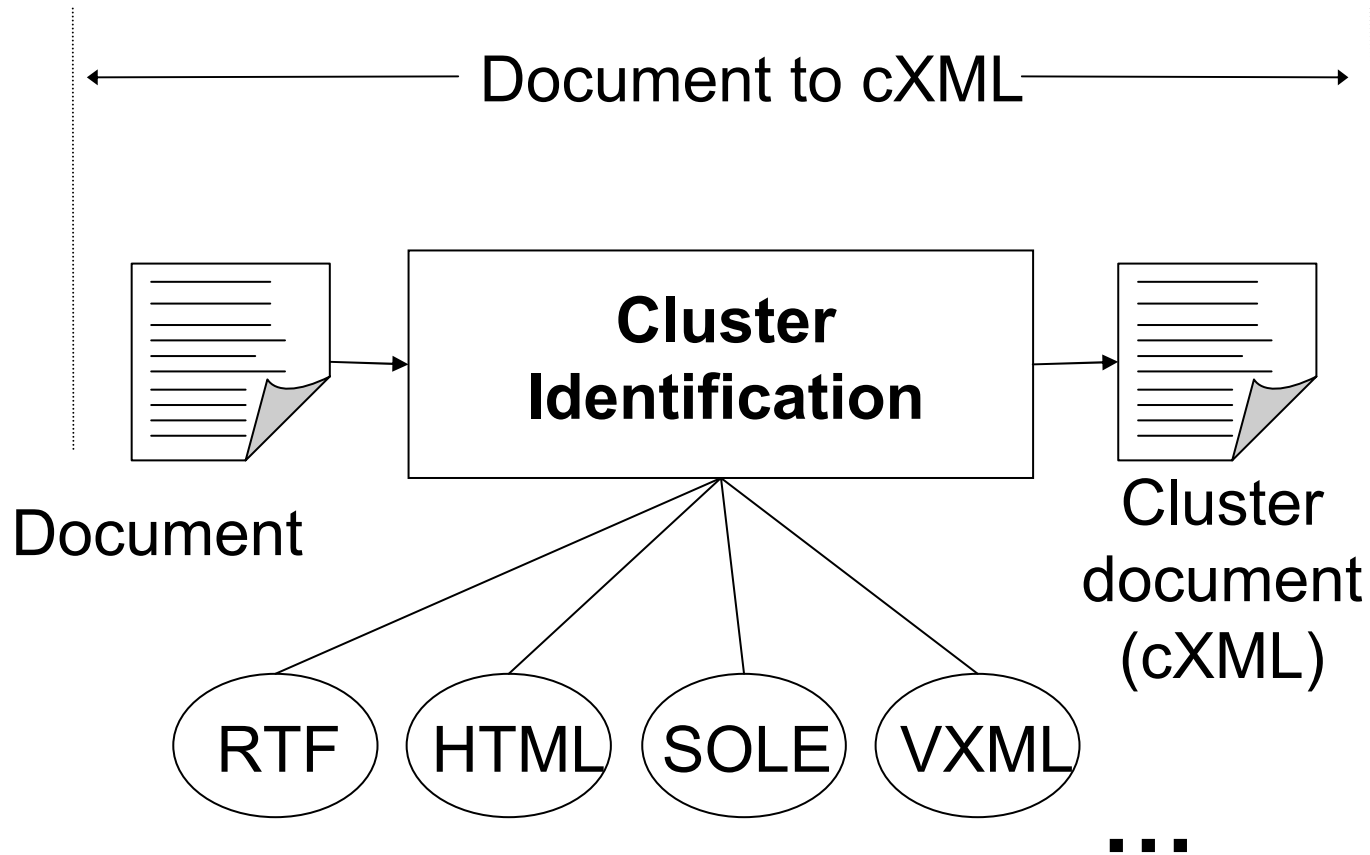
Αρχιτεκτονική



Εντοπισμός και οργάνωση συστάδων



Εντοπισμός συστάδων

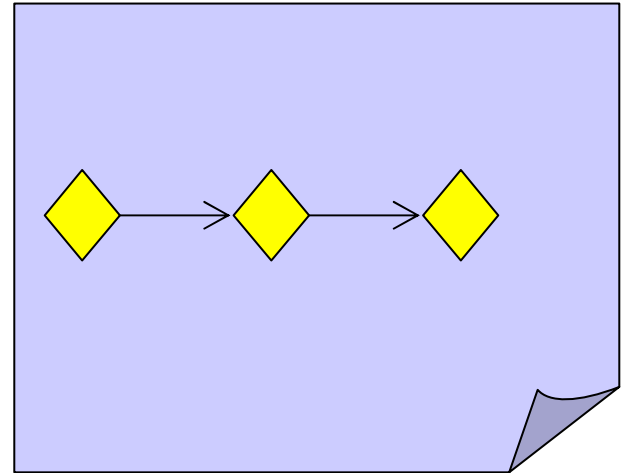
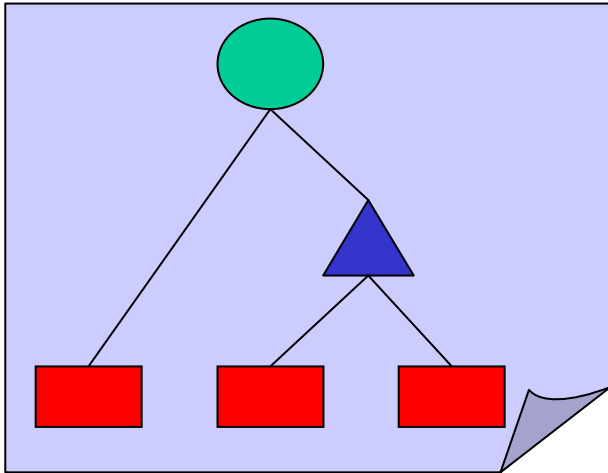


Εντοπισμός συστάδων

- Κάθε τύπος εγγράφου έχει ένα προσαρμοστή που το διαχειρίζεται προκειμένου να εντοπίσει και να οργανώσει τις συστάδες που το αποτελούν.
 - Δημιουργία δέντρου φράσεων και χαρακτηρισμός κάθε κόμβου
- Τεχνικές προσαρμοστών: Finite-state transducers ή γλώσσες μετασχηματισμού.
- Σε κάθε τύπο συστάδας παραχωρείται μία μοναδική διαδρομή (**XPath**).

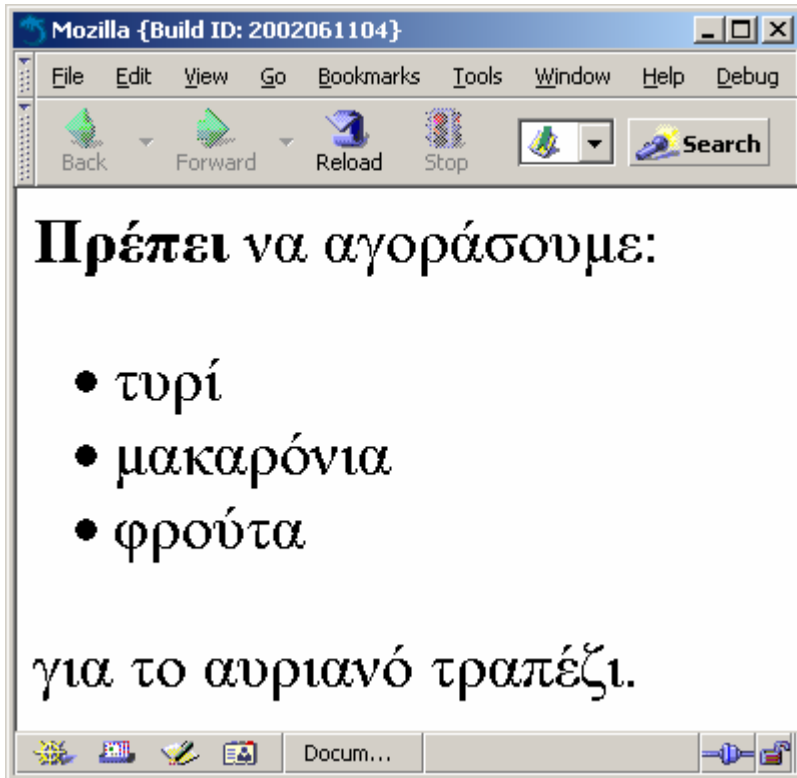
Οργάνωση συστάδων

- Οι συστάδες στα αρχεία που μας ενδιαφέρουν φέρουν οργάνωση δέντρου ή λίστας.



- Στην περίπτωση του δέντρου πρέπει η φωνητική αναπαράσταση να υπενθυμίζει/τονίζει την ιεραρχία, για λόγους νοήματος.

Συστάδα λίστας



```
<html><title></title><body>
```

```
<b>Πρέπει</b> να αγοράσουμε:
```

```
<ul>
```

```
<li>τυρί</li>
```

```
<li>μακαρόνια</li>
```

```
<li>φρούτα</li>
```

```
</ul>
```

```
για το αυριανό τραπέζι.
```

```
</body>
```

```
</html>
```

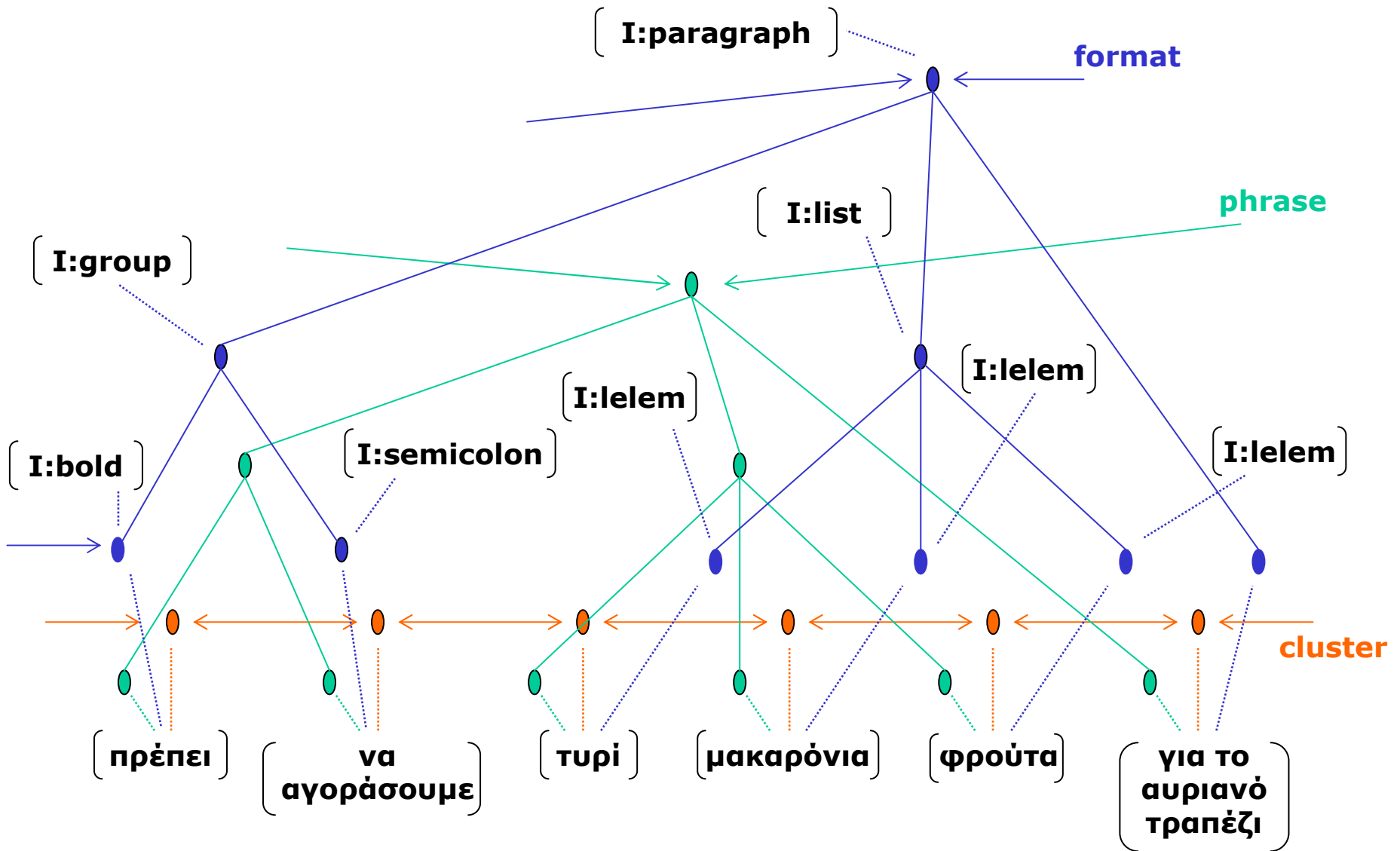
5 καταφατικές προτάσεις-φράσεις. ΛΑΘΟΣ

1 πρόταση με 5 δευτερεύουσες φράσεις. ΣΩΣΤΟ

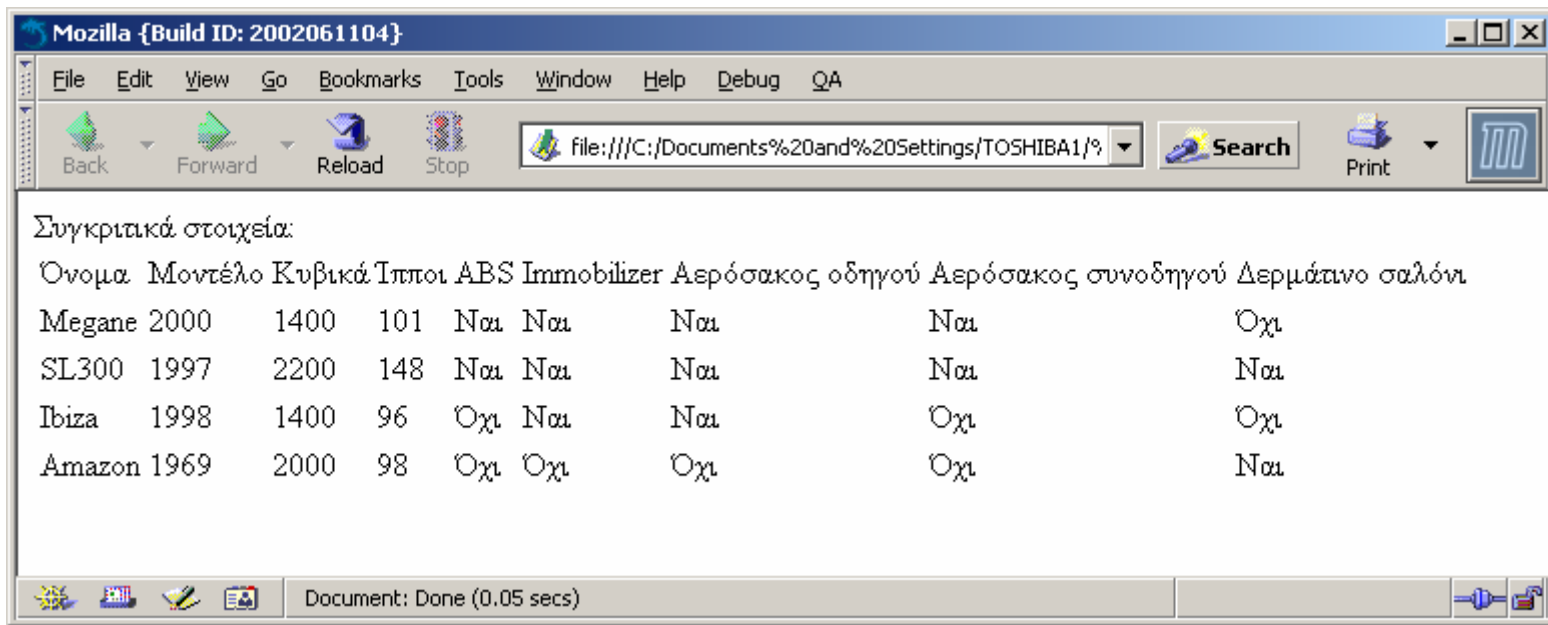
Οργάνωση συστάδων

```
<paragraph>
  <group>
    <emph href="cluster(1)">Πρέπει</emph>
    <semicolon href="cluster(2)">να
    αγοράσουμε</semicolon>
    <list href="cluster(3) .. cluster(5)">
      <lelem href="cluster(3)">τυρί</lelem>
      <lelem href="cluster(4)">μακαρόνια</lelem>
      <lelem href="cluster(5)">φρούτα</lelem>
    </list>
    <plain href="cluster(6)">για το αυριανό
    τραπέζι</plain>
  </group>
</paragraph>
```

HRG



Συστάδα δέντρου



The screenshot shows a Mozilla browser window with the following content:

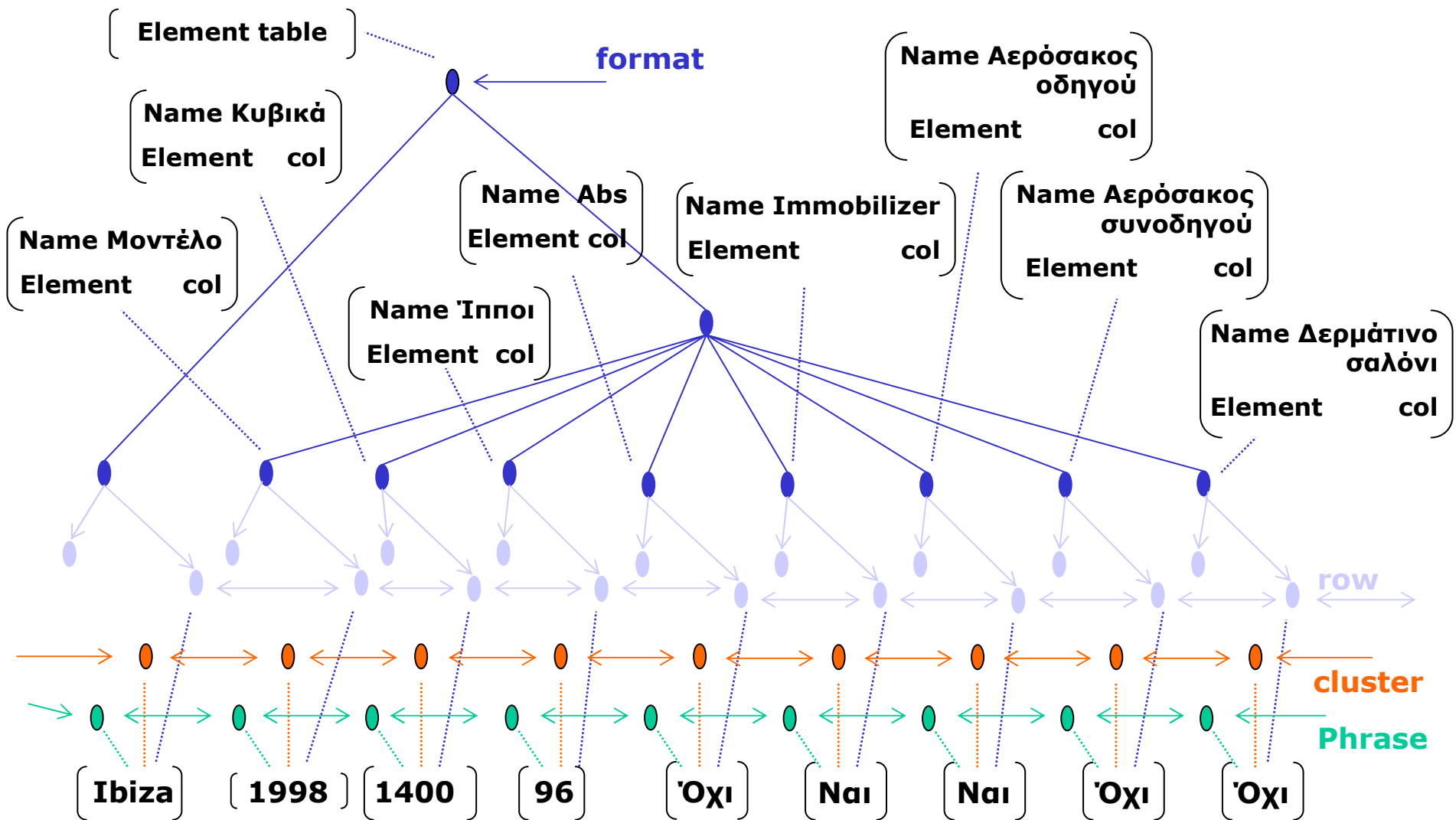
Συγκριτικά στοιχεία:

Όνομα	Μοντέλο	Κυβικά	Ίπποι	ABS	Immobilizer	Αερόσακος οδηγού	Αερόσακος συνοδηγού	Δερμάτινο σαλόνι
Megane	2000	1400	101	Ναι	Ναι	Ναι	Ναι	Όχι
SL300	1997	2200	148	Ναι	Ναι	Ναι	Ναι	Ναι
Ibiza	1998	1400	96	Όχι	Ναι	Ναι	Όχι	Όχι
Amazon	1969	2000	98	Όχι	Όχι	Όχι	Όχι	Ναι

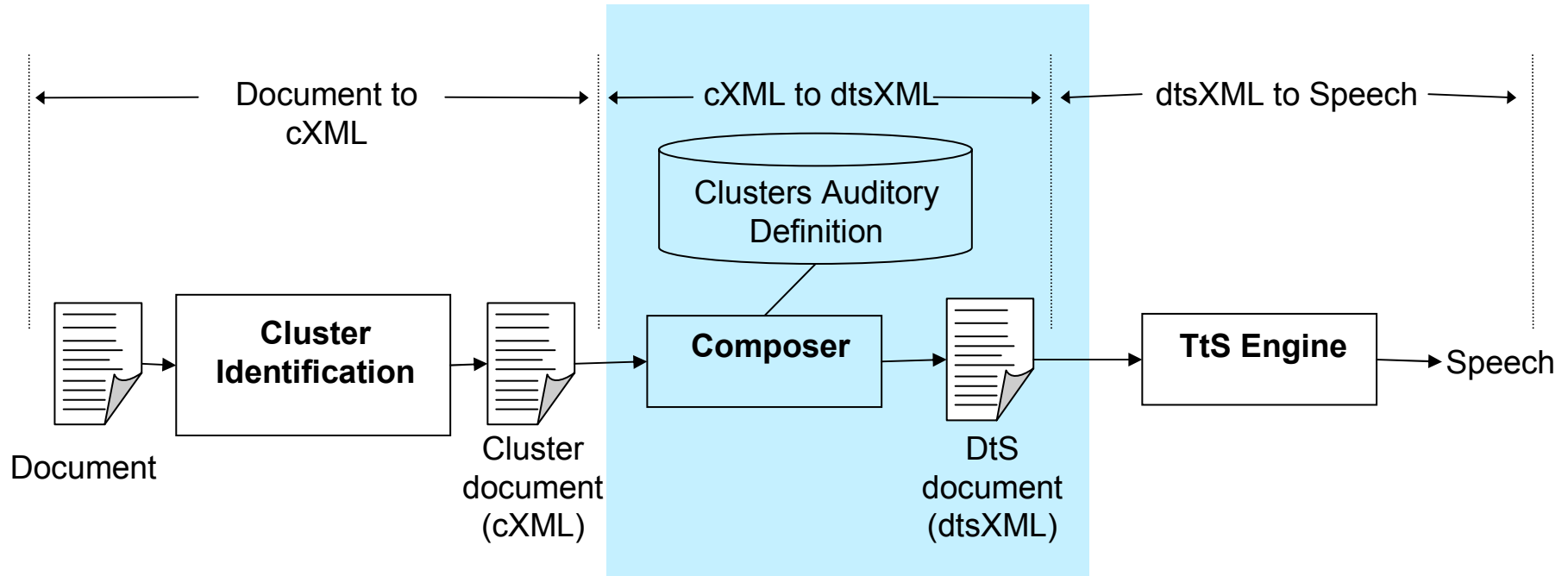
Document: Done (0.05 secs)

```
<table>
<tr><td>Όνομα</td><td>Μοντέλο</td><td>Κυβικά</td><td>Ίπποι</td><td>ABS</td><td>Immobilizer</td><td>Αερόσακος οδηγού</td><td>Αερόσακος συνοδηγού</td><td>Δερμάτινο σαλόνι</td></tr>
<tr><td>Megane</td><td>2000</td><td>1400</td><td>101</td><td>Ναι</td><td>Ναι</td><td>Ναι</td><td>Ναι</td><td>Όχι</td></tr>
<tr><td>SL300</td><td>1997</td><td>2200</td><td>148</td><td>Ναι</td><td>Ναι</td><td>Ναι</td><td>Ναι</td><td>Ναι</td></tr>
<tr><td>Ibiza</td><td>1998</td><td>1400</td><td>96</td><td>Όχι</td><td>Ναι</td><td>Ναι</td><td>Όχι</td><td>Όχι</td></tr>
<tr><td>Amazon</td><td>1969</td><td>2000</td><td>98</td><td>Όχι</td><td>Όχι</td><td>Όχι</td><td>Όχι</td><td>Ναι</td></tr>
</table>
```

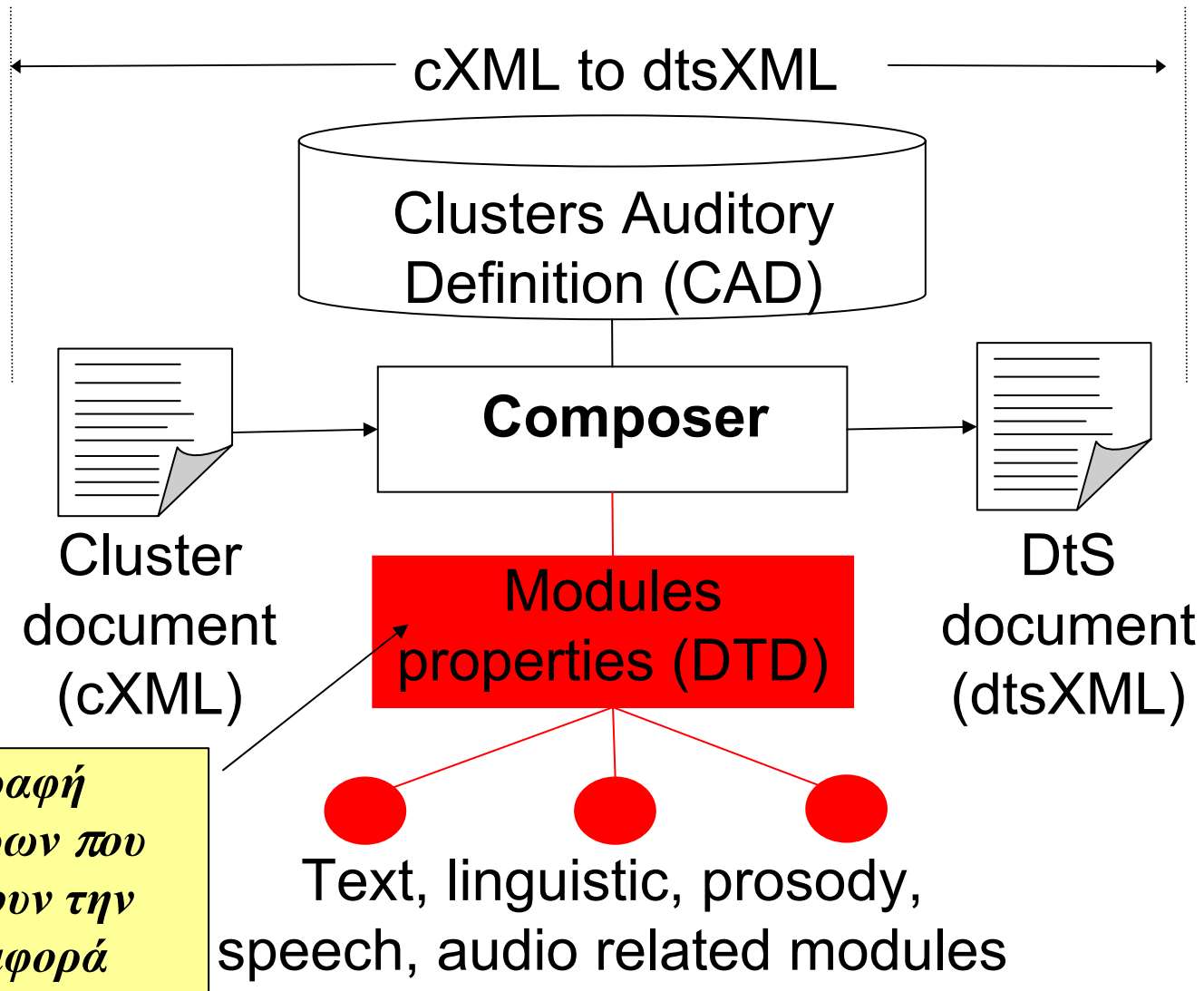
HRG+format



Στοιχειοθεσία



Στοιχειοθεσία



DTD αρθρωμάτων

Για κάθε παράμετρο P_{xi} τύπου T_x με πεδίο τιμών $V=\{V_{x1}, V_{x2}, \dots, V_{xN}\}$ ορίζουμε:

<!ELEMENT P_{xi} (...)>

<!ATTLIST P_{xi} ($V_{x1} | V_{x2} | \dots | V_{xN}$)>

TtS Module

Parameters:

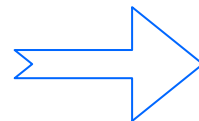
$T_i P_{ik} = V_{im}$

$T_q P_{qj} = V_{qn}$

function1()

function2()

...



TtS Module DTD

<!ELEMENT P_{ik} >

<!ELEMENT P_{qj} >

**<!ATTLIST P_{ik}
($V_{i1} | V_{i2} | \dots | V_{im}$)>**

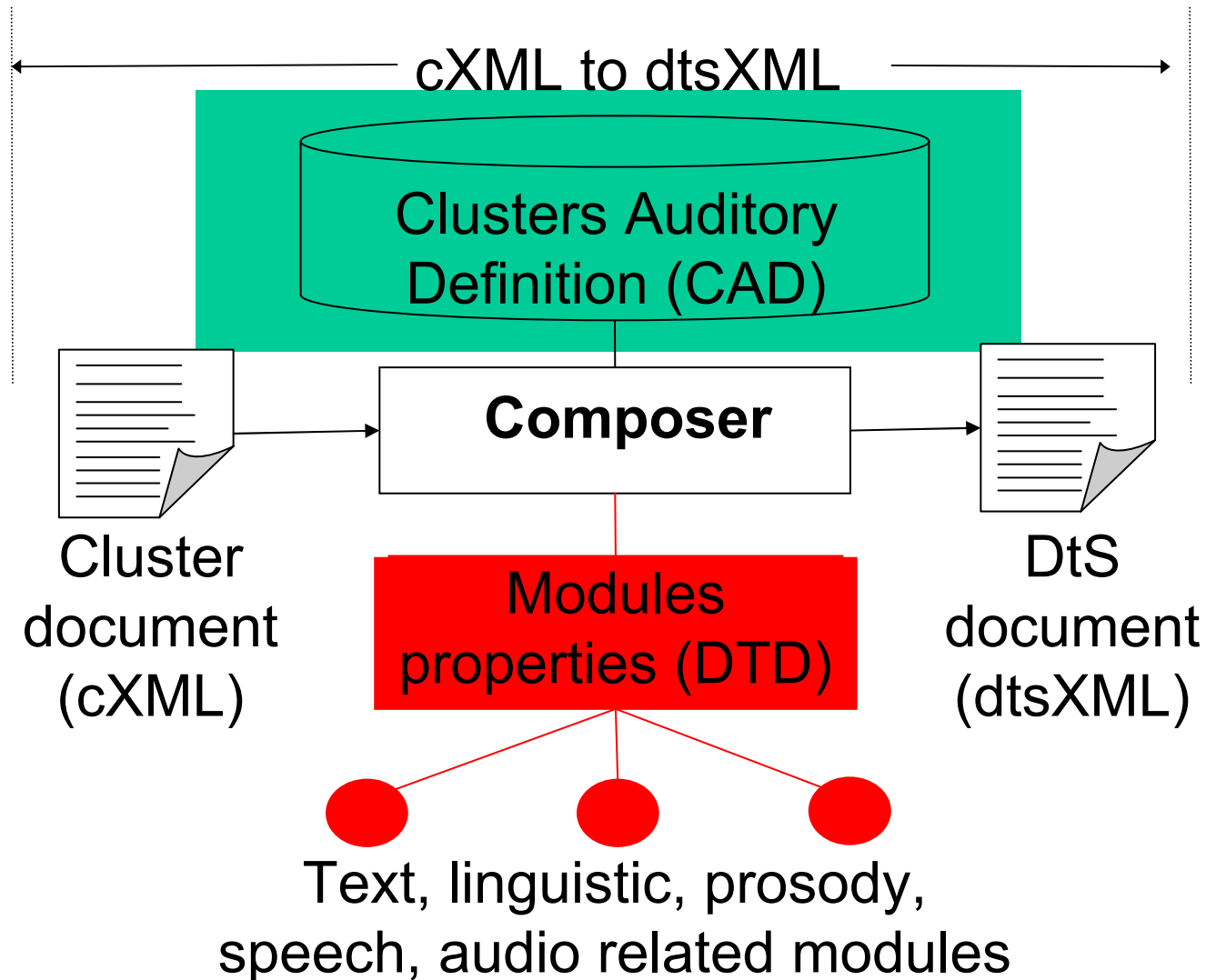
**<!ATTLIST P_{qj}
($V_{q1} | V_{q2} | \dots | V_{qn}$)>**

Παράδειγμα

Το άρθρωμα RHYTHMIC στον ΔΗΜΟΣΘΕΝΗ μπορεί (α) να εφαρμόζει διάφορα tempo εκφοράς και (β) να τροποποιεί την διάρκεια των συλλαβών σε συνάρτηση με την θέση τους μέσα σε μία φράση.

```
<!ELEMENT RHYTHMIC (Tempo*, Syllable*)>
<!ELEMENT Tempo (#PCDATA?)>
<!ATTLIST Tempo speed (lento|andante|allegro)>
<!ATTLIST Tempo acc_var (legato|staccato)>
<!ELEMENT Syllable (#PCDATA?)>
<!ATTLIST Syllable place (begin|end)>
<!ATTLIST Syllable ratio (#PCDATA)>
```

Εφαρμογή των προσωδιακών σεναρίων



Cluster Auditory Definition (CAD)

- Βασισμένα στη **XSL** (eXtensible Stylesheet Language)
- Αναθέτουν σε μία πηγαία συστάδα...

```
<xsl:template match=CLUSTER_XPATH>
```

- ...μία επιθυμητή προσωδιακή συμπεριφορά

```
<MODULEa:FUNCTIONk ATTRIBUTE=VALUEu ...>
```

```
<MODULEb:FUNCTIONm ATTRIBUTE=VALUEw ...>
```

```
<MODULEc:FUNCTIONn ATTRIBUTE=VALUEv ...>
```

- Παράδειγματα:
 - /paragraph/table/tr: *«Εισήγαγε ένα βόμβο 220HZ»*
 - /paragraph/table/tr/td: *«Διάβασε τον τίτλο της κολώνας με γυναικεία φωνή και διάβασε το κείμενο με ανοδικό τονισμό»*

Παράδειγμα

Mozilla {Build ID: 2002061104}

File Edit View Go Bookmarks Tools Window Help Debug QA

Back Forward Reload Stop

file:///C:/Documents%20and%20Settings/TOSHIBA1/%

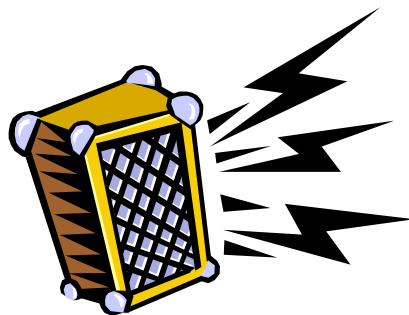
Search Print

Συγκριτικά στοιχεία:

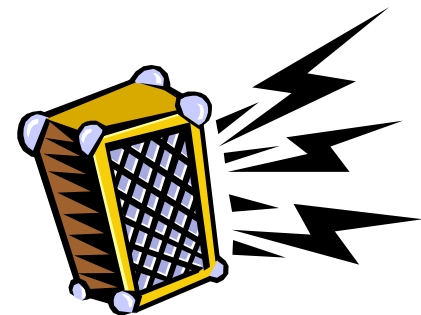
Όνομα	Μοντέλο	Κυβικά	Τπποι	ABS	Immobilizer	Αερόσακος οδηγού	Αερόσακος συνοδηγού	Δερμάτινο σαλόνι
Megane	2000	1400	101	Ναι	Ναι	Ναι	Ναι	Όχι
SL300	1997	2200	148	Ναι	Ναι	Ναι	Ναι	Ναι
Ibiza	1998	1400	96	Όχι	Ναι	Ναι	Όχι	Όχι
Amazon	1969	2000	98	Όχι	Όχι	Όχι	Όχι	Ναι

Document: Done (0.05 secs)

plain...



με D-t-S...



paragraph.table.CAD

```
<xsl:template match="*/table/tr">  
  <Wave:Insert file="220hz-1sec.wav"/>  
  <xsl:apply-template/>  
</xsl:template/>
```

```
<xsl:template match="*/table/tr/td">  
  <VoiceFont:Select name="female.vof">  
    <HRG:item_prepend name="R:Format.parent.name"/>  
    <HRG:item_set feat path to item="previous"  
name="endtone" value="H̄"/>  
  </VoiceFont:Select>  
  <Intonation:Assign endtone="L-L%">  
    <xsl:apply-template/>  
  </Intonation:Assign>  
</xsl:template/>
```

D-t-S

- Το υπο-σύστημα D-t-S του ΔΗΜΟΣΘΕΝΗ επιτρέπει την επαυξημένη σε ποιότητα και πληροφoρία ακουστική αναπαράσταση εγγράφων.
- Πεδία εφαρμογής: τηλεφωνική πρόσβαση σε ηλεκτρονικές βάσεις ή στο Web, πιστότερη απεικόνιση εγγράφων σε άτομα με προβλήματα όρασης, συστήματα IVR, υπηρεσίες καταλόγου, e-mail μέσω τηλεφώνου, ...

Συμπεράσματα

- ΔΗΜΟΣΘΕΝΗΣ: Μία νέα αρχιτεκτονική συστήματος μετατροπής κειμένου σε ομιλία ου μπορεί να φιλοξενήσει τις παραδοσιακές τεχνικές σύνθεσης καθώς και καινοτόμες.
- ΔΗΜΟΣΘΕΝΗΣ: Ανοιχτό εργαλείο για έρευνα στη σύνθεση ομιλίας.
- Μία προσέγγιση για D-t-S: Δομημένα Έγγραφα ή Έγγραφα με οπτική πληροφορία μπορούν να αναπαρασταθούν ακουστικά με μεγαλύτερη ακρίβεια.