

Multilingual mark-up of
text-audio synchronization
at a word-by-word level, how
HTML5
may assist *in-browser* solutions

Gavin Brelstaff (gjb@ crs4.it)
CRS4, Sardinia, Italy

Francesca Chessa
University of Sassari, Italy

Multilingual Web Workshop Rome
March 2013



A.D. MDLXII



First, to the movies



Movies subtitles





```
<video src="Cyrano.ogv">  
  <track  
    kind="subtitles" label="English"  
    src="Cyrano_en.vtt" srclang="en"  
    default/>  
</video>
```



Timed Text Markup Language (TTML)
1.0 (Second Edition)

W3C Working Draft 31 January 2013

HTML5 audio



```
<audio>
  <track
    kind="subtitles" label="English"
    src="file_en.srt" srclang="en"
    default/>
  <source src="file-RU.ogg" type="audio/ogg"/>
  <source src="file-RU.mp3" type="audio/mpeg" />
  Your browser does not support the audio element.
</audio>
```

Simply supply the **vtt** or **srt** timed-text file and the browser does it all for you *line by line*.

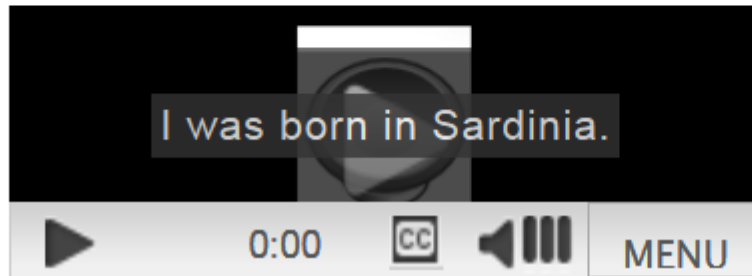
Timed-text audio on the web:



Voice Recording

[edit]

The voice of Grazia Deledda speaking (in Italian) at the Nobel Prize Ceremony in 1926.



<http://commons.wikimedia.org/wiki/TimedText:GraziaDeledda.ogg.en.srt>

Timed-text audio *srt*:



[Main Page](#)
[Welcome](#)
[Community portal](#)
[Village pump](#)

▼ [Participate](#)
[Upload file](#)
[Recent changes](#)
[Latest files](#)
[Random file](#)
[Help](#)
[Contact us](#)
[Donate](#)

TimedText

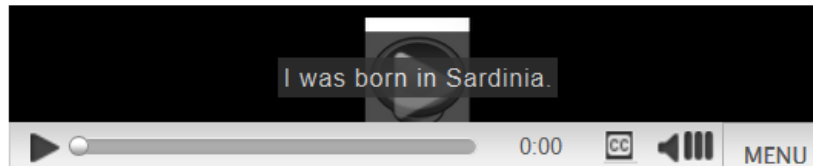
[Discussion](#)

[View](#)

[Edit](#)

[View history](#)

English subtitles for clip: File:GraziaDeledda.ogg



```
1
00:00:00,100 --> 00:00:02,000
I was born in Sardinia.


2
00:00:02,100 --> 00:00:05,100
My family contained wise people,

3
00:00:05,200 --> 00:00:07,000
but also violent ones

4
00:00:07,100 --> 00:00:09,000
and productive artists;
```

Speech to Text

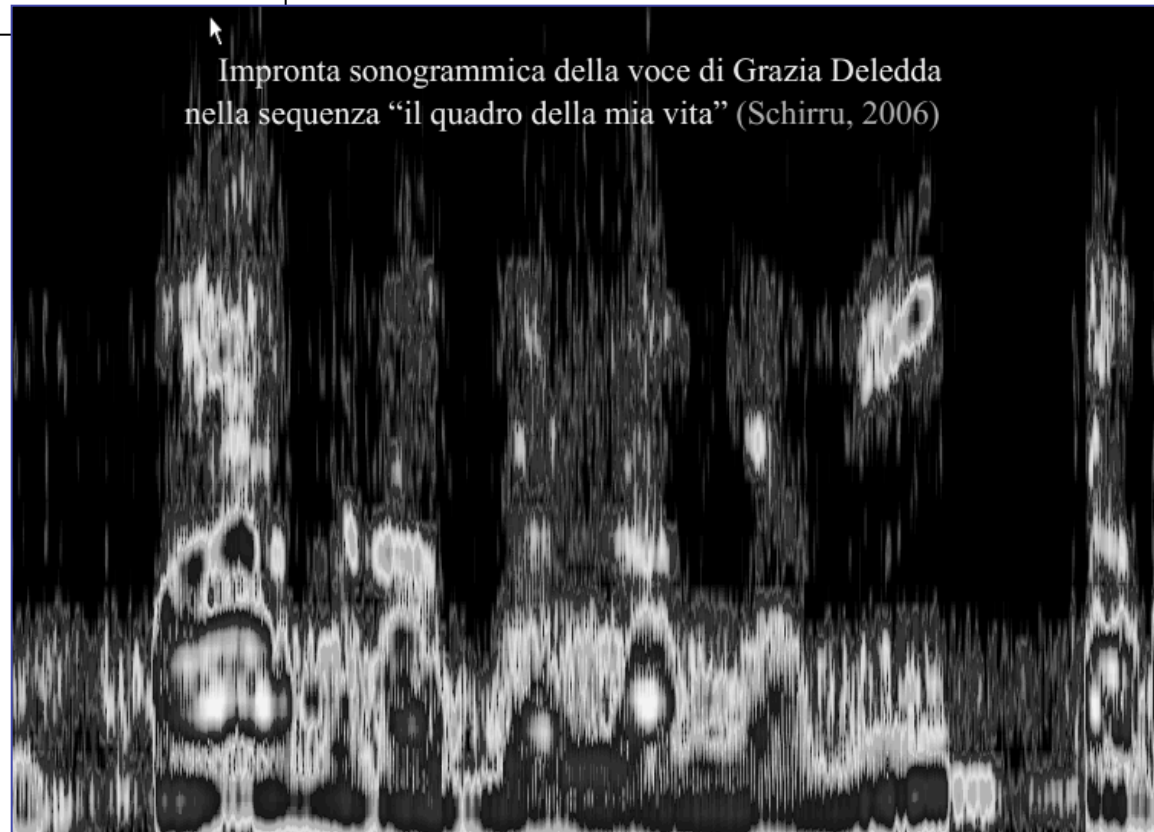
digital spectrogram

 **Voice Recording** [\[edit\]](#)

The voice of Grazia Deledda speaking (in Italian) at the Nobel Prize Ceremony in 1926.

I was born in Sardinia.

0:00 CC MENU



Speech analysis credit: **Carlo Schirru, Univ. Sassari**

Aspetti fonetico-fonologici introduttivi all'analisi strumentale sull'intonazione del sardo (2006)

W3C Workshop Report: Content on the Multilingual Web

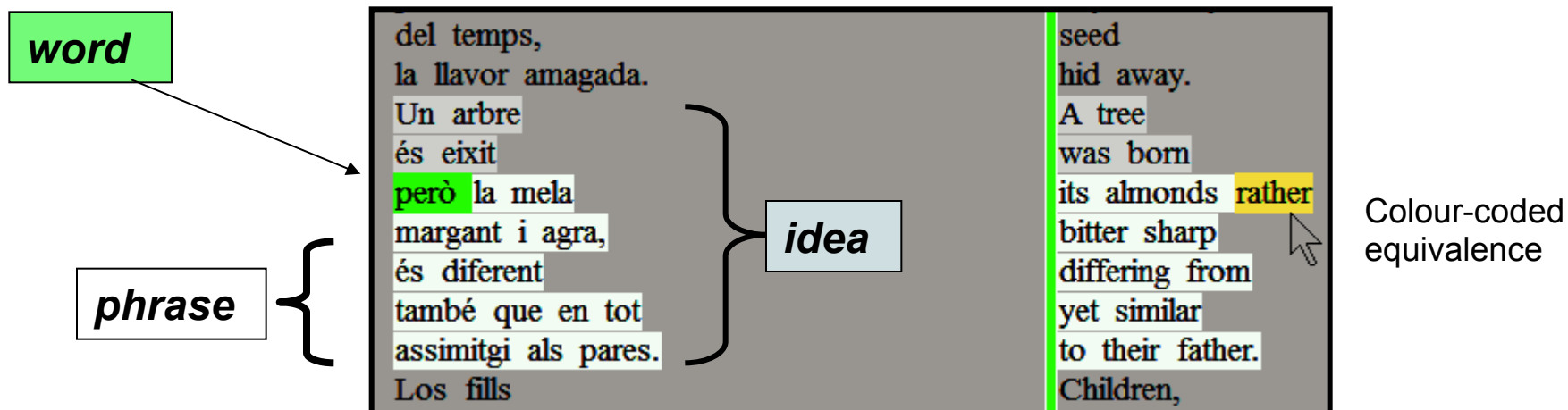
4-5 April 2011, Pisa, Italy



Demo + timed-text

Multilingual markup - recap

A **human** marks up the equivalences between bilingual texts at three different levels: **word**, **phrase**, **idea**.



Web-based alignment and presentation of semantic equivalence [XHTML + CSS + jQuery]

HTML under the hood

```
<audio> ... </audio>
<span class="poem" lang="en">
...
<br unit="stanza" class="milestone"/>
<span class="s">
  <span class="phr">
    <span class="w" n="ru:И_удивило" type="parap"
      start="22.9s" end="24.8s">Astonished was</span>
    <span class="w" n="ru:меня"
      start="24.8s" end="25.5s">!</span>
  </span>
  <span class="phr">
    <br class="lb"/>
    <span class="w" n="ru:как" type="parap"
      start="25.5s" end="25.9s">by </span>
    <span class="w" n="ru:спокойны"
      start="25.9s" end="26.7s">the hush over </span>
    <span class="w" n="ru:воды"
      start="26.7s" end="27.9s">water </span>
  </span>
</span>
...
</span>
```

Timed Text Markup

(TTML)

W3C WD 1.0

31 Jan 2013

Archive format: XML TEI

```
<text><body>
<div type="poem" xml:lang="en"><p>...
<s><phr>
<milestone unit="stanza"/>
<lb/>
<milestone unit="cue" n="22.9s"/>
<w n="ru:И_удивило" type="parap">Astonished was</w>
<milestone unit="cue" n="24.8s"/>
<w n="ru:меня">I:</w>
</phr><phr>
<lb/>
<milestone unit="cue" n="25.5s"/>
<w n="ru:как" type="parap">by</w>
<milestone unit="cue" n="25.9s"/>
<w n="ru:спокойны">the hush over</w>
<milestone unit="cue" n="26.7s"/>
<w n="ru:воды">water</w>
</phr>
<lb/>
...
```

Add one
TEI
milestone
“anchor”
per audio
cue-point

Text Encoding Initiative P5, 2012

HTML5 audio tag

HTML code

```
<audio id="audio" nocontrols>  
  <source src="01-RU.ogg" type="audio/ogg">  
  <source src="01-RU.mp3" type="audio/mpeg">  
  Your browser does not support the audio element.  
</audio>
```

No `<track>` subtitles here

Javascript audio play

```
var myAudio=$('#audio'); // jQuery selector  
myAudio.get(0).currentTime = 15.5 //secs  
myAudio.get(0).play(); // start HTML5 audio
```

Javascript text sync

(scarry stuff instead)

```
setTimeout('switch_on (... )', start_ms ); // times in  
setTimeout('switch_off(... )', end_ms ); // milliseconds
```

See also: [westonruter-html5-audio-read-along on github](#)

Cue-point mark-up tools?

www.nikse.dk/subtitleedit



The screenshot displays the Subtitle Edit application interface. At the top, there is a menu bar (File, Edit, Tools, Video, Spell check, Synchronization, Auto-translate, Options, Networking, Help) and a toolbar with various icons. Below the toolbar, the 'Format' is set to 'SubRip (.srt)' and 'Encoding' is 'ANSI - 1252'. The main window is divided into several sections:

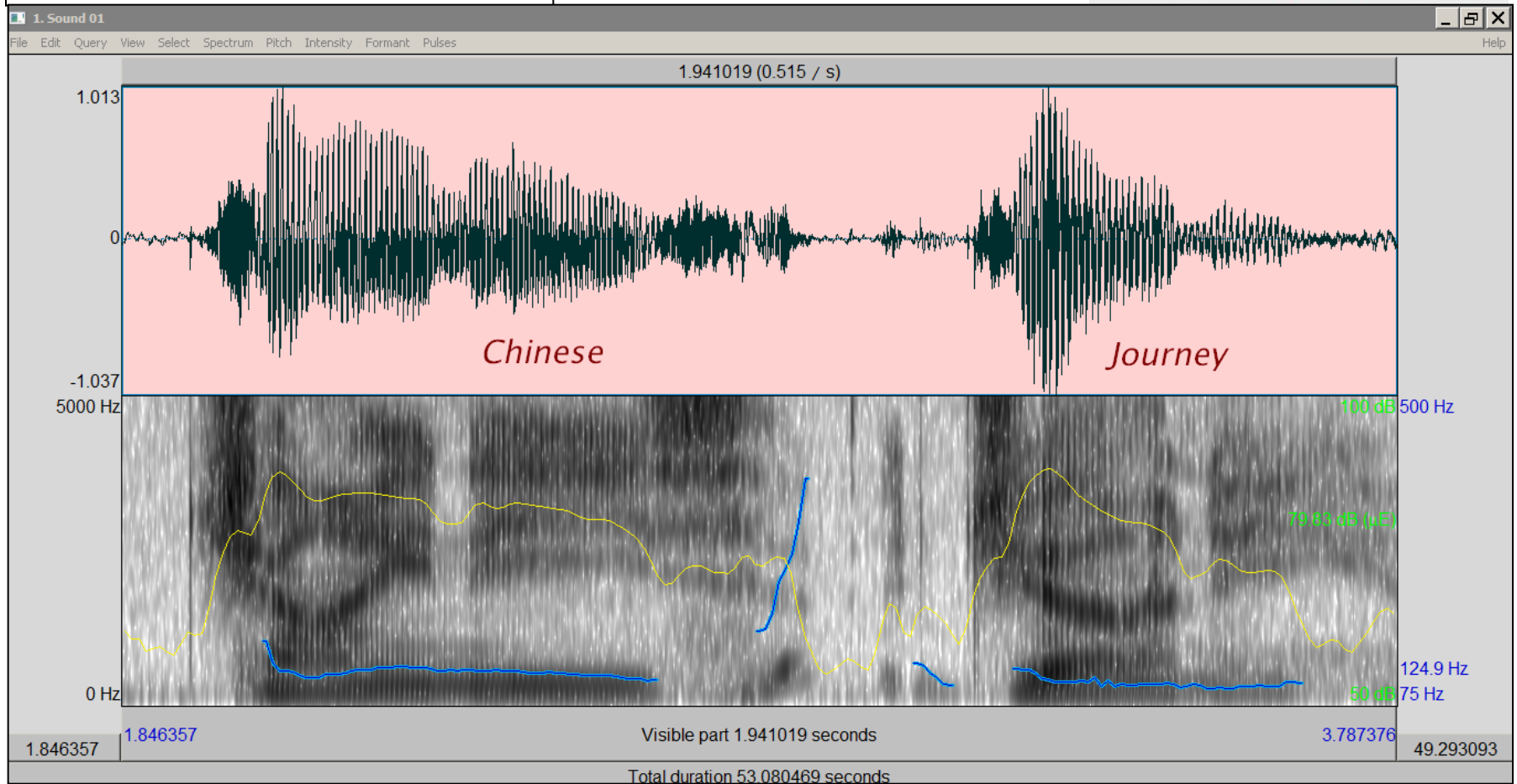
- List view / Source view:** A table listing subtitle entries with columns for #, Start time, End time, Duration, Text, and Original text. Entry #18 is highlighted in blue.
- Preview:** A section showing the selected subtitle's start time (00:01:16.900), duration (2.659), and text in both the original and translated versions.
- Translate:** A panel with options for 'Auto repeat' (checked, 3 times), 'Auto continue' (unchecked), and 'Delay' (3 seconds). It includes buttons for '< Previous', 'Play current', 'Next >', 'Pause', 'Google it', 'Google translate', and 'MS Encarta Thesaurus'.
- Waveform:** A visualization of the audio waveform for the selected subtitle, showing peaks and valleys in red and green.

On the right side of the interface, a video player shows a man speaking, with a subtitle overlay that reads: "Dr. Zachary Walzer, Committee on the Future of Manned Spaceflight, Washington, DC. Vi taler om ansvar for fremtiden." The video player controls are visible at the bottom of the video frame.

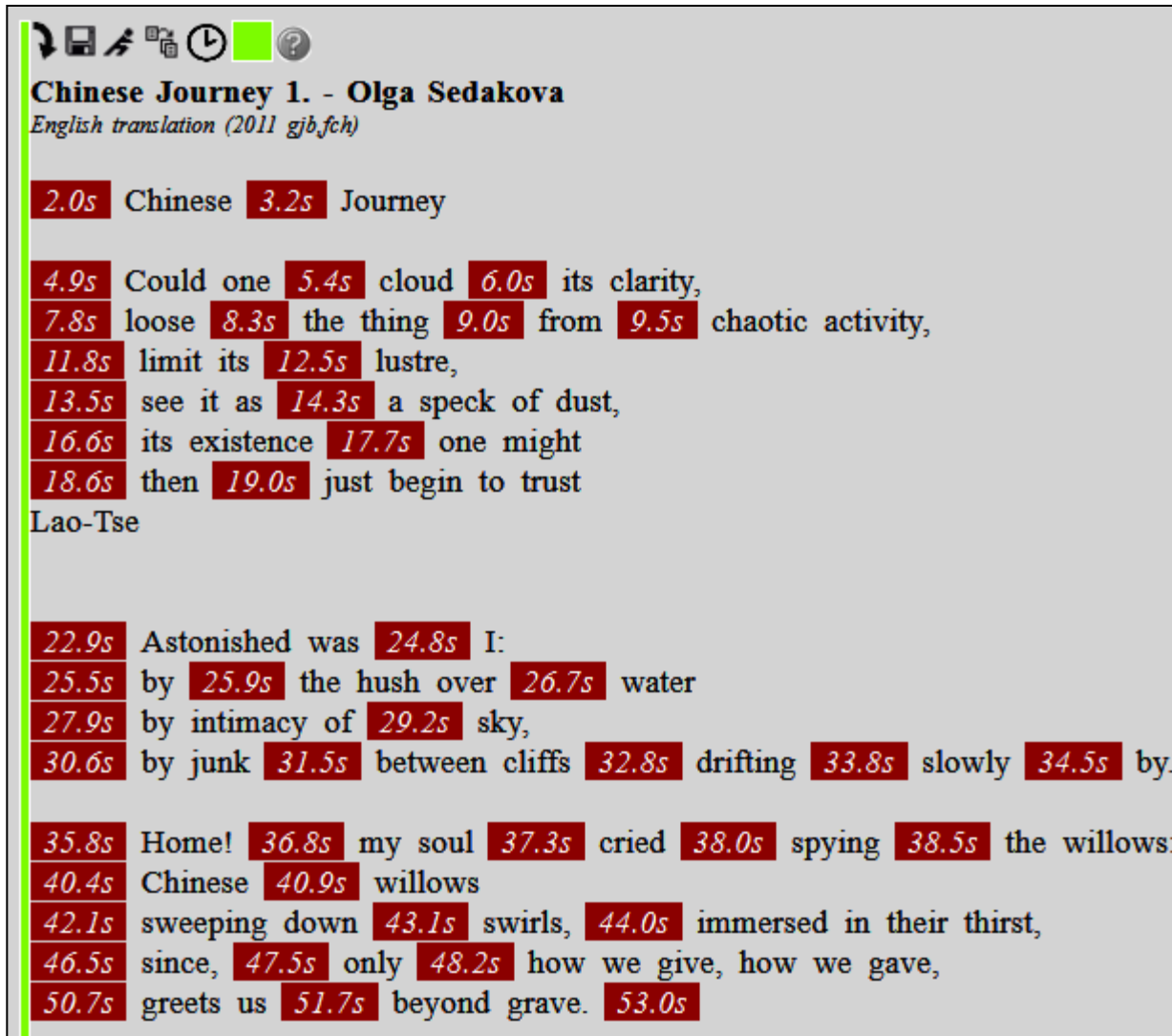
Cue-point mark-up tools?

www.fon.hum.uva.nl/praat

PRAAT
doing phonetics by computer



Cue-point mark-up (visual interface)



The screenshot shows a web page with a toolbar at the top containing icons for back, forward, search, and other navigation functions. Below the toolbar, the title "Chinese Journey 1. - Olga Sedakova" is displayed, followed by the subtitle "English translation (2011 gjb,fch)". The main content is a poem with red rectangular markers placed over specific words and phrases, each containing a time value in seconds (e.g., "2.0s", "3.2s"). The poem is divided into three paragraphs. The first paragraph starts with "2.0s Chinese 3.2s Journey" and ends with "18.6s then 19.0s just begin to trust Lao-Tse". The second paragraph starts with "22.9s Astonished was 24.8s I:" and ends with "30.6s by junk 31.5s between cliffs 32.8s drifting 33.8s slowly 34.5s by.". The third paragraph starts with "35.8s Home! 36.8s my soul 37.3s cried 38.0s spying 38.5s the willows:" and ends with "50.7s greets us 51.7s beyond grave. 53.0s".

2.0s Chinese **3.2s** Journey

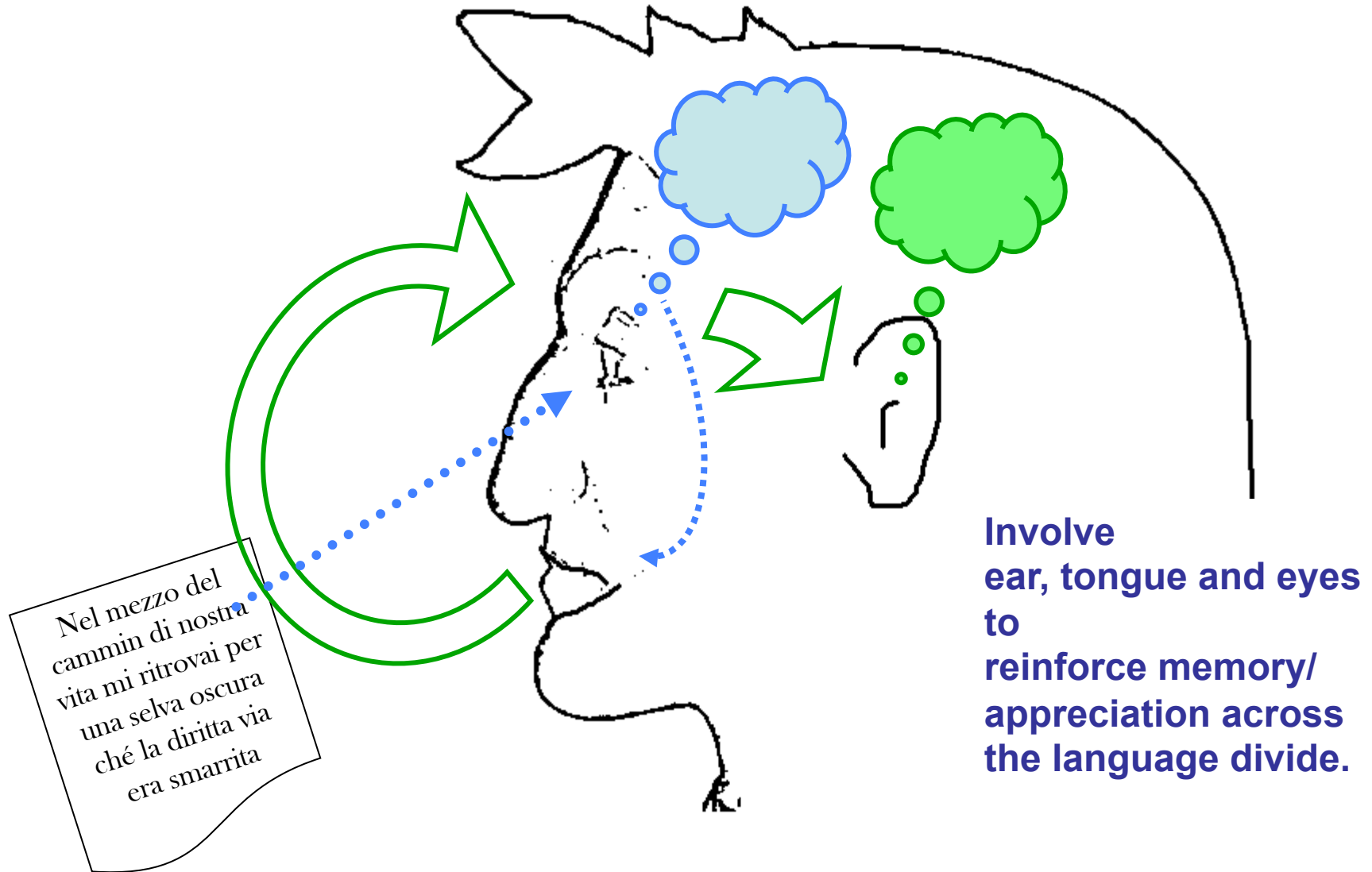
4.9s Could one **5.4s** cloud **6.0s** its clarity,
7.8s loose **8.3s** the thing **9.0s** from **9.5s** chaotic activity,
11.8s limit its **12.5s** lustre,
13.5s see it as **14.3s** a speck of dust,
16.6s its existence **17.7s** one might
18.6s then **19.0s** just begin to trust
Lao-Tse

22.9s Astonished was **24.8s** I:
25.5s by **25.9s** the hush over **26.7s** water
27.9s by intimacy of **29.2s** sky,
30.6s by junk **31.5s** between cliffs **32.8s** drifting **33.8s** slowly **34.5s** by.

35.8s Home! **36.8s** my soul **37.3s** cried **38.0s** spying **38.5s** the willows:
40.4s Chinese **40.9s** willows
42.1s sweeping down **43.1s** swirls, **44.0s** immersed in their thirst,
46.5s since, **47.5s** only **48.2s** how we give, how we gave,
50.7s greets us **51.7s** beyond grave. **53.0s**

Insert & nudge
cue-points
directly on the
web-page
while
listening

Our aim: to activate poetic memory



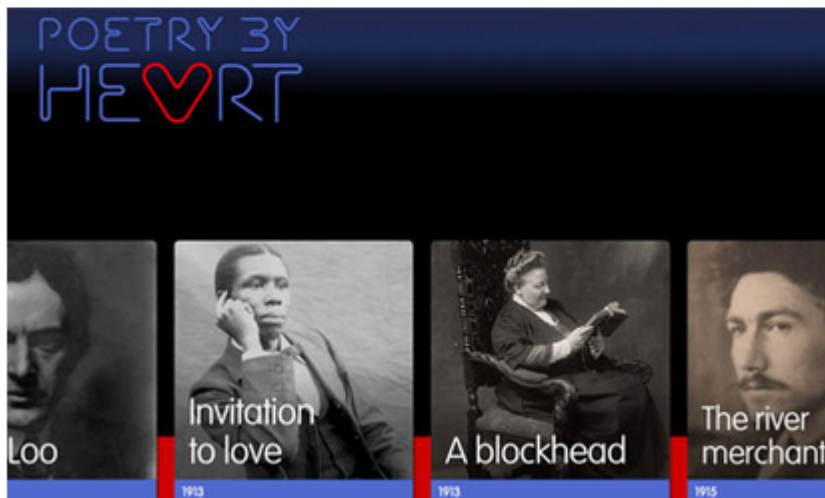
Government backs drive for young to learn poetry by heart

Department for Education funds contest for schoolchildren to learn and recite verse

Peter Walker

guardian.co.uk, Tuesday 8 January 2013 13.59 GMT

[Jump to comments \(45\)](#)



Learning lines ... the Poetry By Heart website

"We preferred poems that make a powerful impact when they are heard aloud - not because they are theatrical, but because they dramatise experiences that surprise us into a new apprehension of ourselves and our capacity for imagining, thinking and marvelling."

Mr Gove said the project would ensure that more children would be captivated by great poetry and it would help "pass our cultural legacy on to the next generation".

Caesar's Europe: poetic memory

<http://www.perseus.tufts.edu/hopper/text?doc=Perseus%3Atext%3A1999.02.0001%3Abook%3D6%3Achapter%3D14>



C. Julius Caesar, *Gallia War*

book:
chapter:

This text is part of:

[Greek and Roman Materials](#)

Table of Contents:

- ▶ [book 1](#)
- ▶ [book 2](#)
- ▶ [book 3](#)
- ▶ [book 4](#)
- ▶ [book 5](#)
- ▼ [book 6](#)
 - [chapter 1](#)
 - [chapter 2](#)
 - [chapter 3](#)
 - [chapter 4](#)
 - [chapter 5](#)
 - [chapter 6](#)
 - [chapter 7](#)
 - [chapter 8](#)
 - [chapter 9](#)



Caes. Gal. 6.14

14.

The Druids do not go to war, nor pay tribute together with the rest; they have an exemption from military service and a dispensation in all matters. Induced by such great advantages, many embrace this profession of their own accord, and [many] are sent to it by their parents and relations. They are said there to learn by heart a great number of verses; accordingly some remain in the course of training twenty years. Nor do they regard it lawful to commit these to writing, though in almost all other matters, in their public and private transactions, they use Greek characters.

to me to have adopted
doctrines to be divulged
to devote themselves
since it generally occurs
they relax their diligence
the memory. They wish to inculcate this as one of their leading tenets, that souls do not become extinct, but pass after death from one body to another, and they think that men by this tenet are in a great degree excited to valor, the fear of death being disregarded. They likewise discuss and impart to the youth many things respecting the stars and their motion, respecting the

Latin (T. Rice Holmes, 1914)

[focus](#) [hide](#)

14. Druides a bello abesse consuerunt neque tributa una cum reliquis pendunt; militiae vacationem omniumque rerum habent immunitatem. [2] Tantis excitati praemiis et sua sponte multi in disciplinam conveniunt et a parentibus propinquisque mittuntur. [3] Magnum ibi numerum versuum ediscere dicuntur. Itaque annos nonnulli vicanos in disciplina permanent. [4] Neque fas esse existimant ea litteris mandare, cum in reliquis fere rebus, publicis privatisque rationibus Graecis litteris utantur. Id mihi duabus de causis instituisse videntur, quod neque in vulgum disciplinam efferri velint neque eos, qui discunt, litteris confisos minus memoriae studere: quod fere plerisque accidit, ut praesidio litterarum diligentiam in perdiscendo ac memoriam remittant. [5] In primis hoc volunt persuadere, non interire animas, sed

The Druids ... learn by heart a great number of verses;
.... Nor do they regard it lawful to commit these to writing ...

[hide](#)

Citation URI: <http://data.perseus.org/citations/urn:cts:latinLit:phi0448.phi001.perseus-eng1:6.14>
Citation URN: urn:cts:latinLit:phi0448.phi001.perseus-eng1:6.14
Document URN: urn:cts:latinLit:phi0448.phi001.perseus-eng1

Poetic memory

internal



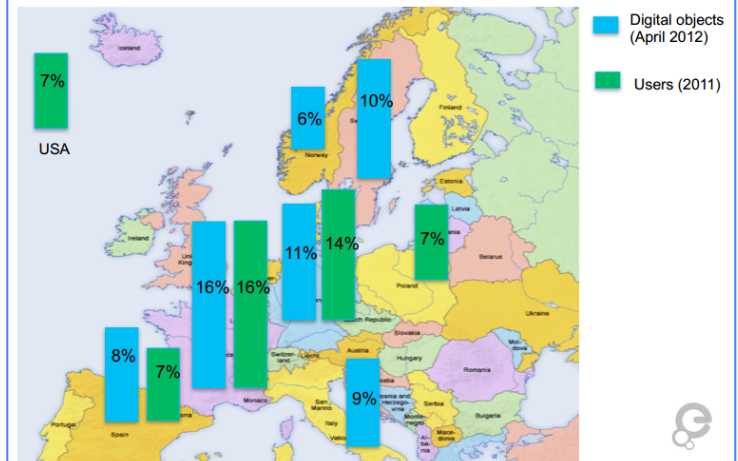
Internal:
Immediately available
to society (in cache)

*Appreciation,
comprehension across
the language divide?*

external



Multilingual digital objects & users



External:
Available
on demand
(in digital archive)

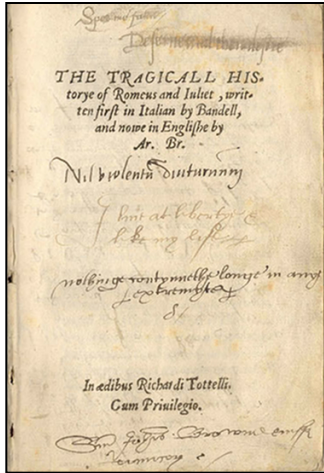
Juliane Stiller, Marlies Olenksy MLW Dublin 2012



- Information is not knowledge
 - knowledge is not wisdom
 - wisdom is not truth ...

F.Zappa 1979

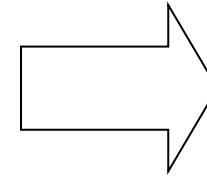
Poetic memory informs society



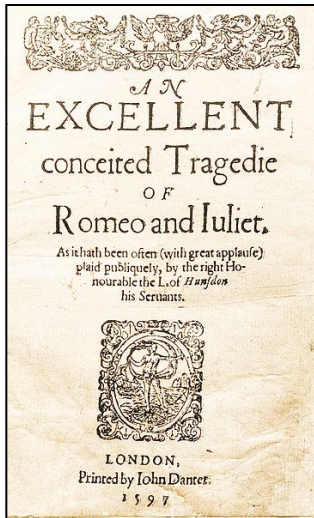
1562

Arthur
Brooke

Prose plot
(information)



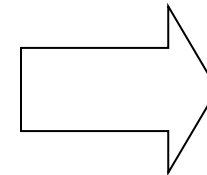
Lost on us



1597

William
Shakespeare

Poetic language
(information *plus*)



**Able to
inform
society**

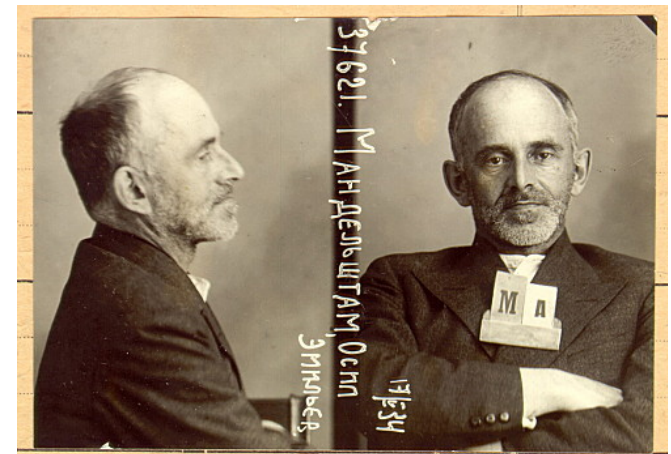
Back to the movies – an extreme social network



Learning by rote

or

Learning by heart?





That's all folks:

Gavin Brelstaff (gjb@crs4.it)
CRS4 09010 Pula (CA) – Sardinia, Italy

Francesca Chessa
University of Sassari, Italy



A.D. MDLXII

