



PLS, a New Standard for Pronunciation Lexicons

Paolo Baggia

Director of International Standards

March 2nd, 2009

A105 Advanced applications and technology

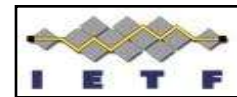
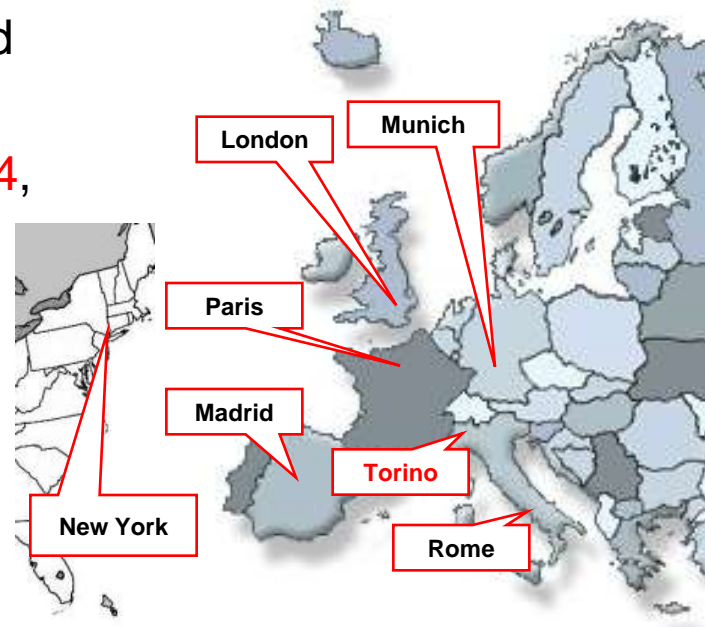
Voice Search Conference 2009



Loquendo
VOCAL TECHNOLOGY AND SERVICES

- Now is the Time for PLS!
- Advantages of PLS
- Major Use Cases of PLS
- Open Issues

- Privately held company (fully owned by Telecom Italia), founded in 2001 as spin-off from Telecom Italia Labs, capitalizing on **30yrs experience and expertise in voice processing**.
- **Global Company**, leader in Europe and South America for award-winning, **high quality voice technologies** (synthesis, recognition, authentication and identification) available in **26 languages** and **62 voices**.
- **Multilingual, proprietary technologies** protected over 100 patents worldwide
- **Financially robust, break-even reached in 2004**, revenues and earnings growing year on year
- **Growth-plan investment** approved for the evolution of products and services.
- **Offices in New York**. Headquarters in Torino, local representative sales offices in Rome, Madrid, Paris, London, Munich
- **Flexible**: About 100 employees, plus a vibrant ecosystem of local freelancers.





“2008 Frost & Sullivan European Telematics and Infotainment Emerging Company of the Year” Award



Winner of “Market leader-Best Speech Engine” Speech Industry Award 2007 e 2008



Loquendo MRCP Server: Winner of 2008 IP Contact Center Technology Pioneer Award



“Best Innovation in Automotive Speech Synthesis” Prize AVIOS-SpeechTEK West 2007



“Best Innovation in Expressive Speech Synthesis” Prize AVIOS-SpeechTEK West 2006



“Best Innovation in Multi-Lingual Speech Synthesis” Prize AVIOS-SpeechTEK West 2005



Now is the Time for PLS!

W3C Recommendation



Pronunciation Lexicon Specification (PLS) Version 1.0

W3C Recommendation 14 October 2008

This version:

<http://www.w3.org/TR/2008/REC-pronunciation-lexicon-20081014/>

Latest version:

<http://www.w3.org/TR/pronunciation-lexicon/>

Previous version:

<http://www.w3.org/TR/2008/PR-pronunciation-lexicon-20080818/>

Editor:

Paolo Baggia, Loquendo

Authors:

Paul Bagshaw, France Telecom
Daniel C. Burnett, Voxeo
Jerry Carter, Nuance
Frank Scahill, BT (until 10 October 2001)

Please refer to the [errata](#) for this document, which may include some normative corrections.

See also [translations](#)

Copyright © 2008 W3C® (MIT, ERCIM, Keio). All Rights Reserved. W3C [liability](#), [trademark](#), and [document use](#) rules apply.

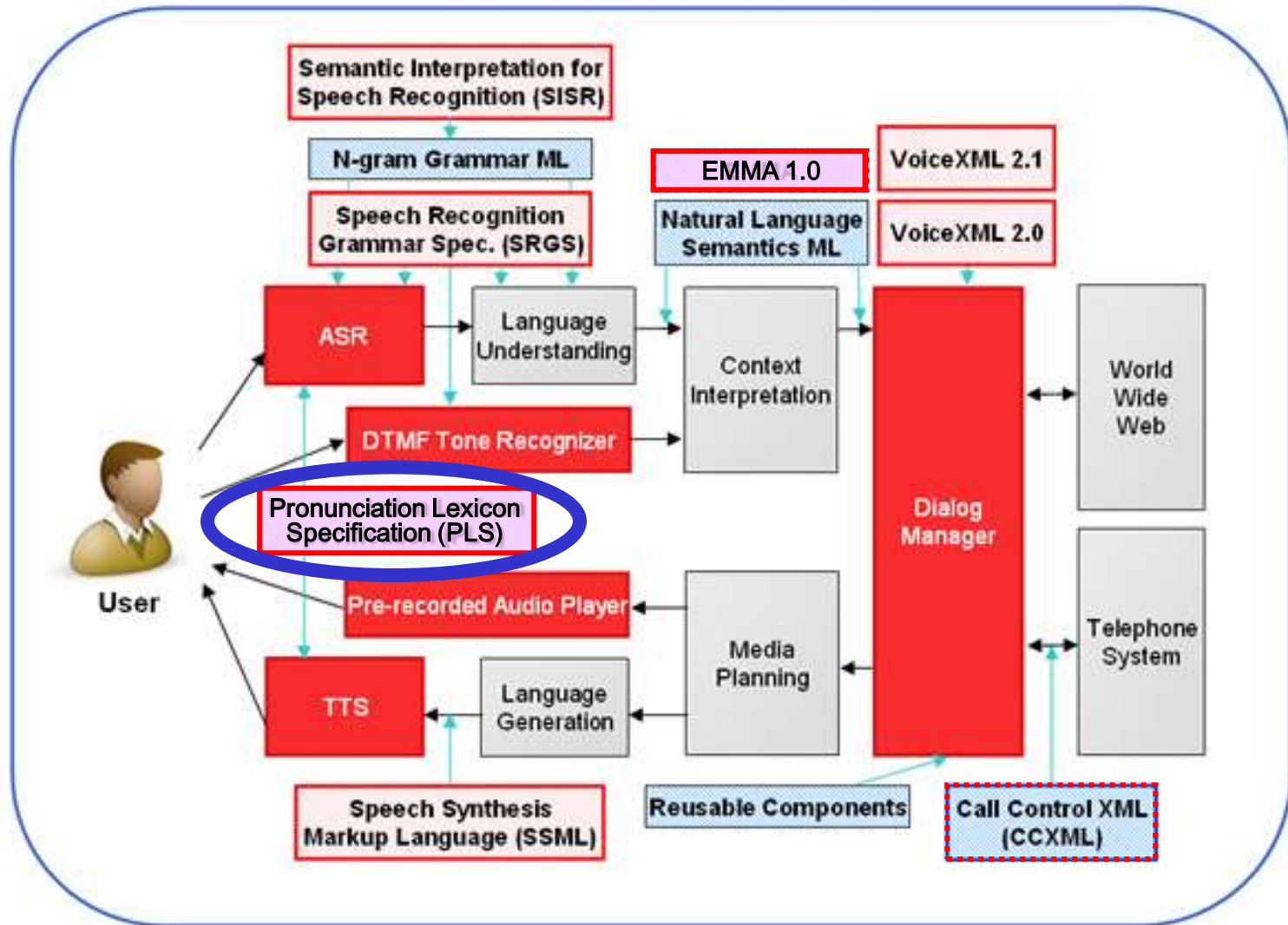
- W3C Voice Browser WG standards are:
 - VoiceXML 2.0 (16 March 2004)
 - SRGS 1.0 (16 March 2004)
 - SSML 1.0 (7 September 2004)

 - SISR 1.0 (5 April 2007)
 - VoiceXML 2.1 (19 June 2007)

 - **PLS 1.0 (14 October 2008)**

- PLS 1.0 – Pronunciation Lexicon Specification, version 1.0, is the most recent W3C Recommendation from the W3C VBWG
<http://www.w3.org/TR/pronunciation-lexicon/> or
<http://www.w3.org/TR/2008/REC-pronunciation-lexicon-20081014/>

- 2001: VBWG early works on Pronunciation Lexicons
 - Activity stopped for a period
- 2004: VBWG re-launches the PLS subgroup
 - Publication of PLS 1.0 Requirements (29 October 2004)
- 2005:
 - First Working Draft (14 January 2005)
- 2006:
 - First Last Call Working Draft (31 January 2006)
 - Second Last Call Working Draft (26 October 2006)
- 2007:
 - Candidate Recommendation (12 December 2007)
- 2008:
 - Implementation Report period
 - **Proposed Recommendation** (18 August 2008)
 - **W3C Recommendation** (14 October 2008)



Advantages of PLS

- Many speech applications need to specify pronunciation for words and phrases
 - Surnames, locations, company names
 - Acronyms
 - Names in specific contexts (restaurants, sports, movie titles, etc.)
 - Foreign words, mixed languages

- Pronunciation is critical both for TTS and ASR
 - Ensures correct reading of prompts by TTS
 - Improves ASR performance

- VoiceXML 2.0/2.1 applications are the reference scenario
 - Prompts are based on SSML 1.0 (or in future SSML 1.1)
 - Recognition grammars are based on SRGS 1.0

■ Pronunciation Lexicon

- A mapping between words (or short phrases), their written representations, and their pronunciations suitable for use by an ASR engine or a TTS engine

■ Pronunciation lexicons are not only useful for voice browsers

- They have also proven effective mechanisms for supporting accessibility for the differently able as well as greater usability for all
- They are used to good effect in screen readers and user agents supporting multimodal interfaces

■ The W3C Pronunciation Lexicon Specification (PLS) Version 1.0 is designed to enable interoperable specification of pronunciation lexicons

- Technology vendors:
 - More powerful and adaptable speech technologies

- Developers:
 - More powerful and portable speech applications

- Customers:
 - More accurate speech application

- Open Sources and Forums:
 - Open Source implementations of PLS 1.0:
 - ➔ http://www.orange.com/en_EN/innovation/patents_licensing/Software/PLS.html
 - Creation of tools to author PLS documents
 - Availability of PLS resources to speed-up application development

- A PLS document is a container (`<lexicon>`) of several lexical entries (`<lexeme>`)
- Each lexical entry contains
 - One or more spellings (`<grapheme>`)
 - One or more pronunciations (`<phoneme>`) or substitutions (`<alias>`)
- Each PLS document is related to a single unique language (`xml:lang`)
- SSML 1.0 and SRGS 1.0 documents can reference one or more PLS documents
- Current version doesn't include morphological, syntactic or semantic information associated with pronunciations

```
<?xml version="1.0" encoding="UTF-8"?>
<lexicon version="1.0"
  xmlns="http://www.w3.org/2005/01/pronunciation-lexicon"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.w3.org/2005/01/pronunciation-lexicon
  http://www.w3.org/TR/pronunciation-lexicon/pls.xsd"
  alphabet="ipa" xml:lang="en-US">

  <lexeme>
    <grapheme>Sepulveda</grapheme>
    <phoneme>səpʌlvɪdə</phoneme>
  </lexeme>

  <lexeme>
    <grapheme>W3C</grapheme>
    <alias>World Wide Web Consortium</alias>
  </lexeme>

</lexicon>
```

Major Use Cases of PLS

- **Multiple pronunciations for the same orthography**
- **Multiple orthographies**
- **Homophones**
- **Homographs**
- **Acronyms, Abbreviations, etc.**

Multiple pronunciations for the same orthography

Multiple pronunciations are represented by more than one `<phoneme>` or `<alias>` element

```
<?xml version="1.0" encoding="UTF-8"?>
<lexicon version="1.0" ... alphabet="ipa" xml:lang="en-GB">
  <lexeme>
    <grapheme>Newton</grapheme>
    <phoneme>ˈnjuːtən</phoneme>
    <phoneme>ˈnuːtən</phoneme>
  </lexeme>
</lexicon>
```

Alternative textual representations for the same word or phrase are represented by more than one `<grapheme>` inside the same `<lexeme>`

All the pronunciations given within the `<lexeme>` apply to each and every `<grapheme>` within the `<lexeme>`

```
<?xml version="1.0" encoding="UTF-8"?>
<lexicon version="1.0" ... alphabet="ipa" xml:lang="jp">
  <lexeme>
    <grapheme>nihongo</grapheme>
    <grapheme>日本語</grapheme>
    <grapheme>にほんご</grapheme>
    <phoneme>nihonjo</phoneme>
  </lexeme>
</lexicon>
```

Words with the same pronunciation but different meanings are represented as different lexemes:

```
<?xml version="1.0" encoding="UTF-8"?>
<lexicon version="1.0" ... alphabet="ipa" xml:lang="en-US">
  <lexeme>
    <grapheme>cede</grapheme>
    <phoneme>si□d</phoneme>
  </lexeme>
  <lexeme>
    <grapheme>seed</grapheme>
    <phoneme>si□d</phoneme>
  </lexeme>
</lexicon>
```

Words with the same spelling but pronounced in different ways are represented using the `role` attribute of the `<lexeme>` element

This mechanism allows for the referencing of defined taxonomies of word classes (part of speech, meaning, etc.)

```
<?xml version="1.0" encoding="UTF-8"?>
<lexicon version="1.0"
  xmlns:claws="http://www.example.com/claws7tags"
  alphabet="x-myorganization-pinyin" xml:lang="zh-CN">
  <lexeme role="claws:VV0"> <!-- base form of lexical verb -->
    <grapheme>处</grapheme>
    <phoneme>chu3</phoneme> <!-- pinyin: "chǔ" in 处罚 处置 -->
  </lexeme>
  <lexeme role="claws:NN"> <!-- noun, neutral for number -->
    <grapheme>处</grapheme>
    <phoneme>chu4</phoneme> <!-- pinyin: "chù" in 处所 妙处 -->
  </lexeme>
</lexicon>
```

SSML 1.1 will support the **role** attribute

```
<speak version="1.1"
  xmlns:claws="http://www.example.com/claws7tags"
  xml:lang="zh-CN">
  <lexicon uri="http://www.example.com/lexicon.pls"
    type="application/pls+xml" xml:id="mylex"/>
  <lookup ref="mylex">
    他这个人很不好相<w role="claws:VV0">处</w>。
    此<w role="claws:NN">处</w>不准照相。
  </lookup>
</speak>
```

PLS 1.0 doesn't define/mandate any values for **role** attribute

PLS 1.0 defines **role** values as qualified names (QNames)

Pronunciations expressed as a sequence of other orthographies (acronyms, abbreviations, etc.) are represented by the `<alias>` element

```
<?xml version="1.0" encoding="UTF-8"?>
<lexicon version="1.0" ... alphabet="ipa" xml:lang="en-US">
  <lexeme>
    <grapheme>W3C</grapheme>
    <alias>World Wide Web Consortium</alias>
  </lexeme>
  <lexeme>
    <grapheme>101</grapheme>
    <alias>one hundred and one</alias>
  </lexeme>
</lexicon>
```

Pronunciations of the <alias> element contents MUST be generated by the processor, using pronunciations described by the <phoneme> element of any constituent graphemes in the PLS document, and without invoking recursive access to the PLS document on the <alias> elements of any constituent graphemes

```
<?xml version="1.0" encoding="UTF-8"?>
<lexicon version="1.0" ... alphabet="ipa" xml:lang="en-US">
  <lexeme>
    <grapheme>GNU</grapheme>
    <alias>GNU is Not Unix</alias>
    <phoneme>gə nu</phoneme>
  </lexeme>
  <lexeme>
    <grapheme>Unix</grapheme>
    <grapheme>UNIX</grapheme>
    <alias>a multiplexed information and computing service</alias>
    <phoneme>ju niks</phoneme>
  </lexeme>
</lexicon>
```

GNU is pronounced:
gə'nu: is Not 'ju:nɪks

ASR

- If more than one pronunciation for a given `<lexeme>` is specified, an ASR processor **MUST** consider each of them as valid pronunciations for the `<grapheme>`

TTS

- If more than one pronunciation for a given `<lexeme>` is specified, a TTS processor **MUST** use the first one in document order that has the **prefer** attribute set to `"true"`
- If none of the pronunciations has **prefer** set to `"true"`, the TTS processor **MUST** use the first one in document order **unless the TTS processor is documented as having a method of selecting pronunciations**, in which case the processor **MUST** use any one of the pronunciations

An ASR processor will recognize both pronunciations, whereas a TTS processor will only use the first one (because it is the first in document order that has **prefer** set to "true").

```
<?xml version="1.0" encoding="UTF-8"?>
<lexicon version="1.0" ... alphabet="ipa" xml:lang="en-US">
  <lexeme>
    <grapheme>lead</grapheme>
    <alias prefer="true">led</alias>
    <phoneme prefer="true">li□d</phoneme>
  </lexeme>
  <lexeme>
    <grapheme>led</grapheme>
    <phoneme>led</phoneme>
  </lexeme>
</lexicon>
```

Open Issues

- Registration registry for alternative standard phonetic alphabets:
 - IPA covers all languages, but ...
 - Mandarin Chinese commonly uses PinYin,
 - Japan standardized JEITA phonetic alphabet,
 - etc.
- ➔ This is addressed by SSML 1.1

- Multilingual PLS lexicons

- Standardization of a common set of `role` attribute values, e.g. common syntactic types, common semantic classes

- Further integration with other W3C standards to promote a broader usage of PLS, e.g. accessibility, multimedia, multimodal, internationalization

- PLS 1.0 W3C Recommendation (14 October 2008)
<http://www.w3.org/TR/pronunciation-lexicon/> or
<http://www.w3.org/TR/2008/REC-pronunciation-lexicon-20081014/>
- Voice Browser Activity Page (VoiceXML, SSML, SRGS, ...)
<http://www.w3.org/Voice/>
- VoiceXML Forum
<http://www.voicexml.org/>
- International Phonetic Association
<http://www.arts.gla.ac.uk/IPA/>
- IPA online keyboard and pickers:
<http://weston.ruter.net/projects/ipa-chart/view/keyboard/>
<http://people.w3.org/rishida/scripts/pickers/ipa>

For more information please

- Keep an eye on: www.loquendo.com
- Subscribe Loquendo Newsletter
- Contact me: paolo.baggia@loquendo.com

THANK YOU!