

Description form - PROVIDER

Language Resources: Broadcast News

B.1. GENERAL INFORMATION

Recording source:

- Radio Station: Band:
 TV Station:
 Internet URL:

Recording period: (*from date-to date*)

Type of recording:

- News broadcast Phone calls Interviews
 Conversations Other:

Comments:

B.2. SPEAKER SPECIFIC INFORMATION

Sex and number of speakers:

- Male Number:
 Female Number:
 Impostors Number:

Total number:

Age class: (indicate number)

- Children I (__ to __) Adults I (__ to __) Elderly (over 60)
 Children II (__ to __) Adults II (__ to __) Age unknown
 Teenagers I (__ to __) Adults III (__ to __) Other distribution:
 Teenagers II (__ to __) Adults IV (__ to __)

Origin:

- Native Non native Unknown

Comments:

Information included about:

- Place of living Place of birth Education level
 Dialect/accent Trained speakers Speaking/hearing impairments
 Other

B.3. RECORDING INFORMATION**Mode:**

Read Spontaneous Both

Topics:

List of topics:

Total number of topics:

Channel:

Telephone Studio

Fidelity:

Low Medium High

Speech percentage: (range: 70-95%)**Background noise:**

Music Speech Commercials
 Sound effects Other:

Equipment:

Radio type:

Sound card type:

Platform type:

Strategy: (if any)**B.4. LEXICON****Lexicon included:** Yes No

Size (number of lexicon entries) :

Format:

ASCII SGML TEI
 Other:

Pronunciation lexicon:

Available Not available

Transcriptions:

Canonical only Canonical + alt. pronunciation Automatically generated
 Checked manually Generated fully manually Other:

Phoneme set:

SAMPA DARPA IPA
 CPA Other:

B.5. TRANSCRIPTIONS

Linguistic information:

<input type="checkbox"/> Orthographic	<input type="checkbox"/> Morphological	<input type="checkbox"/> Phonetic
<input type="checkbox"/> Syntactic	<input type="checkbox"/> Semantic	<input type="checkbox"/> Prosodic
<input type="checkbox"/> Other:		

Level of segmentation:
Level of transcription:

Transcription standard:

<input type="checkbox"/> SAM	<input type="checkbox"/> SGML	<input type="checkbox"/> XML
<input type="checkbox"/> NIST/LDC		
<input type="checkbox"/> Other:		

Annotations:

<input type="checkbox"/> None	<input type="checkbox"/> Speaker noise	<input type="checkbox"/> Background noise
<input type="checkbox"/> Mispronunciation	<input type="checkbox"/> Truncation	<input type="checkbox"/> Other:

B.6. TECHNICAL INFORMATION

Signal encoding:

<input type="checkbox"/> PCM	<input type="checkbox"/> Other:
------------------------------	---------------------------------

File format:

<input type="checkbox"/> AIFF	<input type="checkbox"/> Wav	<input type="checkbox"/> Without header
<input type="checkbox"/> Real audio	<input type="checkbox"/> NIST/Sphere	<input type="checkbox"/> Au
<input type="checkbox"/> Other:		

Amount of data:

Size (Mb, Gb, etc):
Duration (minutes, hours, etc):

Sampling rate:

<input type="checkbox"/> 8 kHz	<input type="checkbox"/> 16 kHz	<input type="checkbox"/> 32 kHz
<input type="checkbox"/> 44,1 kHz	<input type="checkbox"/> 48 kHz	<input type="checkbox"/> Other:

Quantisation:

<input type="checkbox"/> 8 bit	<input type="checkbox"/> 16 bit	<input type="checkbox"/> Other:
--------------------------------	---------------------------------	---------------------------------

Byte order:

<input type="checkbox"/> Lo-hi (Intel)	<input type="checkbox"/> Hi-lo (Motorola)
--	---

Compression:

<input type="checkbox"/> None	<input type="checkbox"/> Shorten	<input type="checkbox"/> Other:
-------------------------------	----------------------------------	---------------------------------

Number of recording channels:

<input type="checkbox"/> 1 (mono)	<input type="checkbox"/> 2 (stereo)	<input type="checkbox"/> Other:
-----------------------------------	-------------------------------------	---------------------------------

Annotation standard:

<input type="checkbox"/> SAM	<input type="checkbox"/> SGML	<input type="checkbox"/> XML
<input type="checkbox"/> NIST/LDC		
<input type="checkbox"/> Other:		

Sound quality measures included:

<input type="checkbox"/> SNR	<input type="checkbox"/> Cross talk	<input type="checkbox"/> Clipping rate
<input type="checkbox"/> Background noise	<input type="checkbox"/> Other:	

Tools used for measuring sound quality:

B.7. FURTHER COMMENTS

ELRA/ELDA
9 rue des Cordelières
F-75013 PARIS, FRANCE
Tel +33 1 43 13 33 33 - Fax +33 1 43 13 33 30 - E-mail mapelli@elda.org
<http://www.elda.org>