Description form - PROVIDER

Language Resources: Broadcast News

B.1. GENERAL INFORMATION

Recording source: *(radio station)*

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

Type of recording:

☐ News   ☐ Phone call   ☐ Interviews

☐ Conversation   ☐ 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 (up to 12)   ☐ Teenagers (12-15)   ☐ Teenagers (16-19)

☐ Adults young (20-30)   ☐ Adults (31-45)   ☐ Adults (46-60)

☐ Elderly (over 60)   ☐ Age unknown   ☐ Other distribution:

Comments:

Origin:

☐ Native   ☐ Non native   ☐ Unknown

Comments:

Geographic distribution:

Total number of regions:

Percentage per region:

Regions included:

Information included about:

☐ Place of living   ☐ Place of birth   ☐ Place of (secondary) education

☐ Dialect/accent

Comments:

Additional speaker information included:

☐ Speaking/hearing impairments   ☐ Height   ☐ Weight

☐ Smoking habits   ☐ Trained speakers   ☐ Education level

☐ Profession

Comments:

B.3. RECORDING INFORMATION

Mode:

☐ Spontaneous   ☐ Planned

Episodes:

Total number of episodes:

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:

Computer type:

Strategy: *(if any)*

B.4. LEXICON

Lexicon included: ☐ Yes   ☐ No

Size (number of lexicon entries) :

Format:

☐ ASCII   ☐ XML   ☐ Other:

Pronunciation lexicon:

☐ Available   ☐ Not available

Transcriptions:

☐ Canonical only   ☐ Canonical + alt. pronunciation   ☐ Automatically generated

☐ Checked manually   ☐ Generated fully manually   ☐ Other:

Phoneme set:

☐ SAMPA   ☐ DARPA   ☐ Other:

B.5. LINGUISTIC INFORMATION AND SEGMENTATION

Linguistic annotation:

☐ Ortographic   ☐ Morphological   ☐ Phonetic

☐ Syntactic   ☐ Semantic   ☐ Prosodic

☐ Other:

Level of segmentation:

Level of annotation:

B.6. TECHNICAL INFORMATION

Signal encoding:

☐ PCM   ☐ Other:

File format:

☐ AIFF   ☐ Wav   ☐ Without header

☐ SAM   ☐ NIST/Sphere   ☐ Au

☐ Other:

Amount of data:

Size (Mb, Gb, etc):

Duration (minutes, hours, etc):

Sampling rate:

☐ 16 kHz   ☐ Other:

Quantisation:

☐ 16 bit   ☐ Other:

Compression:

☐ None   ☐ Zip   ☐ Shorten

☐ Other:

Number of recording channels:

☐ 1 (mono)   ☐ Other:

Annotation standard:

☐ SAM   ☐ SGML   ☐ XML

☐ NIST/LDC   ☐ Other:

Sound quality measures included:

☐ SNR   ☐ Cross talk   ☐ Clipping rate

☐ Background noise   ☐ Other:

Tools used for measuring sound quality:

B.7. FURTHER COMMENTS