

English Project

Contents

Introduction

2. Many-Talker Prompt-File Distribution

3. Few-Talker Prompt-File Distribution

4. Very-Few-Talker Prompt Files

Introduction

This report documents the subjects, equipment, recording environments, materials and the file structure details involved in making the speech and laryngograph recordings for the English section of the SAM EUROM1 speechdatabase.

The corpus consists of 4 components:

- a) systematically structured C(C)VC monosyllables to be produced in isolation and in a number of controlled precursive and following contexts (see section 5.1).
- b) selected numbers from 0-9999, such that all the phonotactic possibilities of the English number system were covered (see section 5.2).
- c) short passages containing 5 thematically connected sentences (see section 5.3).
- d) sentences composed to compensate the phoneme-frequency imbalance resulting from the thematic (i.e. not structurally orientated) composition of the passages (see section 5.4).

Different parts and increasing amounts of this material were recorded by three sets of native speakers, recruited from Southeast England:

- i) Many Talker Set (MT) - (30 women, 30 men):
 - 100 numbers
 - 3 passages
 - 5 sentences

ii) Few Talker Set (FT) - (5 women and 5 men selected from MT):

Isolated C(C)VCs
5 x 100 numbers
15 passages
25 sentences

For the few talker data set laryngograph signals were recorded together with the acoustic speech signal. However, for the C(C)VC material only the first of the five repetitions is available on CD5. See CD5_E.TXT for more information.

iii) Very Few Talker Set (VFT) - (1 woman and 1 man selected from FT):

Contextualised C(C)VCs
5 x Context words

For the very few talker data set laryngograph data is available for all the recordings.

The recordings were carried out in anechoic rooms at University College London and at the National Physical Laboratories (NPL), Teddington, and SAM colleagues from UCL and NPL provided technical support in calibration tests of the recording room to comply with the conditions stipulated in the Recording Protocol document (SAM-RSRE-15, Dec. 1990). (Files recorded before 30/5/91 were made in the Anechoic Chamber at UCL. Files recorded on or after 30/5/91 were made in the Anechoic Chamber at NPL).

Two operators shared the task of recording. Calibration was carried out for each subject prior to recording, and continual monitoring of each speaker's performance ensured that a minimum of deviations from the prompt text, and a minimum of articulatory lapses are contained in the recordings. Any error noted by the operator led to a repeat recording of the prompt item (i.e. a block of CVCs, a block of 20 numbers, a five-sentence passage, or a block of five sentences). Inspection and back-up (on Exabyte) of the recorded material followed immediately after each session.

The subjects were selected so that there was an equal number of women and men, as good a coverage of age groups, and as wide a range of voice types as possible (cf. SAM-UCL-030, May 1991). There was also a considerable variation in body size and no direct means of calculating vocal-tract dimensions was available for the subjects who were not recorded using the laryngograph. For the FT and VFT subjects, who all gave simultaneous speech and laryngograph recordings, the precise positioning of the microphone relative to the subjects lips gives a basis for vocal tract length estimation (see Appendix A).

Age groupings of subjects are given in the following table:

Age Group	Subject Code (male)	Subject Code (female)
-----------	---------------------	-----------------------

20-29	15 *MP, MR, MU, MV, MX, MY, MZ, NA, ND, NI, NJ, NK, NL, OG, OL	13 *MC, *MD, MN, MS, MT, MW, NB, NM, NN, NS, NW, NY, OA
30-39	6 *MB, *ME, NE, NF, NT, NV	5 *MK, ML, NH, OE, OJ
40-49	4 !MA, MM, MO, NO	7 MQ, OB, OD, OF, OH, OI, OK
50 +	5 NC, NG, NQ, *NX, NZ	5 !MJ, NP, NR, NU, *OC

Table 1: The subject codes given are those used in the speech-signal filenames. The codes preceded by * refer to the subjects belonging to the FT set, and those preceded by ! are those who were also in the Very-Few-Talker set. All subject details are given in section 7.

The distribution of texts by subject is given in the following table. All MT-subjects produced one repetition of the numbers (prompt files N1-N5).

2. Many-Talker Prompt-File Distribution

Passage	Speaker (male)	Passage	Speaker (female)
O1	MA MT NG NU OG	Q1	MD NA NN, OA
O2	MA MT NH NU OG	Q2	MN NA NN OB
O3	MA MU NH NU OI	Q3	MN NA NO OB
O4	MK MU NH NV OI	Q4	MN NB NO OB
O5	MK MU NI NV OI	Q5	MO NB NO OC
O6	MK MV NI NV OH	Q6	MO NB NP OC
O7	MB MV NI NW OH	Q7	MO NC NP OC
O8	MB MV NJ NW OH	Q8	MP NC NP ME
O9	MB MW NJ NW OJ	Q9	MP NC NQ ME

O0	ML MW NJ NX OJ	Q0	MP ND NQ ME
P1	ML MW NK NX OJ	R1	MQ ND NQ OD
P2	ML MX NK NX OK	R2	MQ ND NR OD
P3	MC MX NK NY OK	R3	MQ NE NR OD
P4	MC MX NL NY OK	R4	MR NE NR OE
P5	MC MY NL NY OL	R5	MR NE NS OE
P6	MM MY NL NZ OL	R6	MR NF NS OE
P7	MM MY NM NZ OL	R7	MS NF NS OF
P8	MM MZ NM NZ MJ	R8	MS NF NT OF
P9	MD MZ NM OA MJ	R9	MS NG NT OF
P0	MD MZ NN OA MJ	R0	MT NG NT OG

Sentence

Speaker

F1 MA MQ NA NK NU OD
F2 MK MR NB NL NV OE
F3 MB MS NC NM NW OF
F4 ML MT ND NN NX OG
F5 MC MU NE NO NY OI
F6 MM MV NF NP NZ OH
F7 MD MW NG NQ OA OJ
F8 MN MX NH NR OB OK
F9 MO MY NI NS OC OL
F0 MP MZ NJ NT ME MJ

3. Few-Talker Prompt-File Distribution

The Few-Talker set subjects were drawn from the Many-Talker Group. The relations between the Many-Talker and Few-Talker codes are as follows:

MA = FA MB = FB MC = FC MD = FG ME = FE
MJ = FJ MK = FF MP = FD NX = FI OC = FH

All FT-subjects recorded 5 repetitions of the numbers (N1 - N5) and 5 repetitions of the isolated C(C)VC material (S1 - S5).

Passages	Speaker	Sentences	Speaker
O1 - O5	FA FE FD FI	F1 - F5	FA FB FG FE FI
O6 - O0	FA FF FD FI	F6 - F0	FC FJ FF FD FH
P1 - P5	FA FG FF FI	P6 - P0	FC FG FF FJ
Q1 - Q5	FB FC FG FJ	Q6 - Q0	FB FC FJ FH
R1 - R5	FB FE FH	R6 - R0	FE FD FH

4. Very-Few-Talker Prompt Files

The two VFT subjects were members both of the Many-Talker set and the Few-Talker set. The same subject codes were used in the Very-Few-Talker set as in the Few Talker set. As immediately above, FA corresponds to MA and FJ to MJ. These two subjects recorded ALL the contextualised C(C)VC stimuli (files T1 - T5, U1 - U5, V1 - V5, W1 - W5, X1 - X5) and 5 repetitions of the context words (Z1) in isolation. All their recordings were made using both condenser microphone and laryngograph signals.