

French Project

Contents

General information

Definitions

Subjects

Experimental recording conditions

Recording session

Recording procedure

Recording sequence

Statistics

References

General information

The multilingual speech database EUROM_1 has been recorded in the framework of the ESPRIT n°2589 - SAM (Multilingual Speech Input/Output Assessment Methodology and Standardisation) [1][7]. The recordings that form the EUROM1_F (French part of EUROM1 [2]) were carried out at ICP, Grenoble, during the year 1991-92 using the EUROPEC (V 4.11) software developed in the framework of the ESPRIT n°2589 SAM Project. The total amount of data is approximately 3.2 Gigabytes, that are stored on five CDROM disks. The fifth disk contains compressed waveform files (laryngographic signal, see disk 5), so that the full database could be fitted on 5 disks.

The volume ID names of the 5 discs of EUROM1_F are:

CD1_EURO1_F to CD5_EURO1_F

Sampling rate: 20000 Hz
coding: 16 bits.

The acquisition equipment used was also in accordance with SAM recommendations:

- B&K 4165 condenser measurement microphone
- B&K 2603 amplifier
- OROS AU22 I/O board
- Sampling frequency: 20000 Hz.
- samples coded on 16 bits.

The realisation of this database was made possible in the frame of the SAM project and with the support of French CNRS and MESR (D.I.S.T. and D.M.T.I). The "Institut de la Communication Parlée" was in charge of the project, in collaboration with the "Institut de Phonétique" in Aix and "IRIT" in Toulouse.

Corpus adaptation:	D. Autesserre (IPA), J-F. Malet (IRIT-CERFIA)
Phonetic transcription:	D. Autesserre, C. Meunier (IPA)
Recording campaign:	J. Zeiliger (ICP)
Speaker selection:	D. Autesserre
Organisation, structuration, premastering:	J. Zeiliger
Supervision:	J-F. Serignat (ICP)

Definitions

CORPORA:

Corpora consist in a French adaptation (and not a strict translation) of the typical corpus commonly agreed [3].

- Continuous speech :

- 40 passages made of five thematically linked sentence, showing a coherent semantic structure so that to induce a correct prosodic structure at each sentence level.
- 50 patching sentences, to let appear the distinctive features of the language that are not present in the previous texts.

- Logatomes :

- 82 CVC (Consonant + Vowel + Consonant) standing for the possible variations in the French language of the initial consonant and of the final consonant with the extremes vowels [a,i,u]. These

"logatomes" are uttered in isolation or in context, using five different contexts. The words used for the contexts have also been pronounced individually.

initial consonant variation	form: C + /il/	(20)
initial consonant variation	form: C + /al/ or /oul/	(23)
two initial consonants	form: CC + /il/	(5)
final consonant variation	form: /li/ + C	(19)
vowel variation in context	form: /t/ + V + /t/	(15)
	total	82

- Numbers :

5 groups of 20 numbers (from 0 to 9999) selected as to cover all the phonotactic possibilities in the set of numbers under consideration.

Data

Corpus code	Corpus name	No. of items	Item type
F0	Sentences	5	S
F1	Sentences	5	S
F2	Sentences	5	S
F3	Sentences	5	S
F4	Sentences	5	S
F5	Sentences	5	S
F6	Sentences	5	S
F7	Sentences	5	S
F8	Sentences	5	S
F9	Sentences	5	S
N1	Numbers	20	N

N2	Numbers	20	N
N3	Numbers	20	N
N4	Numbers	20	N
N5	Numbers	20	N
O0	Passage	1	P
...			
O9	Passage	1	P
P1	Passage	1	P
...			
P9	Passage	1	P
Q0	Passage	1	P
...			
Q9	Passage	1	P
R0	Passage	1	P
...			
R9	Passage	1	P
S1	CVC_1 with initial consonant variation	20	W
S2	CVC_2 with initial consonant variation	23	W
S3	CVC_3 with initial consonant variation	5	W
S4	CVC_4 with initial consonant variation	19	W
S5	CVC_5 with vowels variation in a context	15	W
T1	CVC_1 in context 1	20	S

T2	CVC_2 in context 1	23	S
T3	CVC_3 in context 1	5	S
T4	CVC_4 in context 1	19	S
T5	CVC_5 in context 1	15	S
U1	CVC_1 in context 2	20	S
U2	CVC_2 in context 2	23	S
U3	CVC_3 in context 2	5	S
U4	CVC_4 in context 2	19	S
U5	CVC_5 in context 2	15	S
V1	CVC_1 in context 3	20	S
V2	CVC_2 in context 3	23	S
V3	CVC_3 in context 3	5	S
V4	CVC_4 in context 3	19	S
V5	CVC_5 in context 3	15	S
W1	CVC_1 in context 4	20	S

Table 1: Table of corpus

Legend: D Digits
L Letters
N Numbers
S Sentences
P Passage
W Isolated words

The EUROM Speakers

All the speakers are French native-speakers. They belong to three sets. A first group of 60 speakers, the Many Talker Set, of whom only twelve have a strong regional accent. A second group of 10 speakers, the Few Talker Set, was selected from the

Many Talker Set, none of whom had a regional accent. Lastly, a third group, the Very Few Talker Set comprised of 4 speakers selected from the Few Talker Set.

In the second group, the glotto-pharyngeal signal has been recorded together with the speech signal, using a laryngograph.

1st Group (Many Talker Set)

SP. CODE	SEX	AGE	HEIGHT	WEIGHT	EDUC. LEVEL	SMOKING	LINGUISTIC CHARACTERIS.
AC	M	46	170	68	High	Stop 10	Std French
AD	M	50	180	84	High		South. French
AY	M	39	182	70	High	Little	Std French
BA	F	33	160	51	High		Std French
BD	M	26	173	65	High		Std French
BF	M	27	173	60	High		Std French
BG	M	44	172	76	High		Std French
BL	M	52	186	85	High	Little	South. French
BO	M	33	178	75	High		Parisian
BS	F	37	160	55	Sec		Std French
BT	M	50	180	82	High	Stop 13	South F strong
BU	F	26	167	52	High		Parisian light
BO	M	33	178	75	High		Parisian
BS	F	37	160	55	Sec		Std French
BT	M	50	180	82	High	Stop 13	South. F strong
BU	F	26	167	52	High		Parisian light
CB	M	23	172	65	High	Much	Std French
CC	F	27	159	42	High		Std French
CM	F	42	160	55	High		Std French
DA	F	29			High	Much	Std French
DI	F	23	162	56	High		Std French light lisp
DJ	M	45	171	80	High	Much	Std French
DN	F	45	158	55	Sec		Std French
DO	M	24	183	77	High	Yes	Breton
DV	F	20	160	47	Sec	Yes	of the Dauphiné
EF	F	45	172	64	High	Stop 2	Std French
EH	M	25	178	70	High		Std French
EP	M	54	175	72	High	Stop 13	Std French
FA	F	28	162	52	High		Std French
GC	F	32	159	53	Sec	Little	Std French
GG	M	42	172	70	High	Little	Std French
GJ	M	60	170	62	Sec	Little	Std French
HJ	M	48	177	68	High		Std French

HM	F	45	160	50	High		Std French
JA	F	44	168	60	High		Std French
JB	M	41	178	68	High		Std French
JF	F	52	164	56	High		South. French
JN	F	24	167	62	High	Much	Std. French
KM	F	38	165	59	High		Std French
MC	M	27	175	75	High		Std French
MD	F	44	167	60	Sec	Little	Std French Larynx?
MH	F	29	163	56	High	Little	Parisian light
MJ	M	50	182	74	Sec		of the Dauphiné
PA	F	51	161	59	High		Std French
PG	M	52	168	83	High	Stop 20	Std French
PJ	M	28	177	70	High		South. French strong
PL	M	24	171	60	High		Std French
PN	F	62	163	57	High		Std French
PV	F	23	163	57	High		Dauphiné ? light lisp
RG	M	25	175	65	High	Little	Parisian
RJ	F	42				Yes	Std French Larynx
RO	F	26	163	65	High	Much	Std French
RS	F	44	165	58	Sec		Std French
SC	M	23	182	75	High		Parisian
SH	M	34	180	65	High	Very little	Std French
SJ	M	44	170	72	High	Little	Std French
SL	M	50	174	67	High		Std French
TD	F	36	160		High		Std French
VC	F	39	165	62	Sec	Little	of the Dauphiné
VF	F	36	160	55	Sec	Little	Std French
VI	F	26	172	78	High		Std French
VJ	M	45	174	75	High	Stop 15	Std French Larynx ?
VS	M	28	175	69	High		Std French ?
VT	F	27	180	62	High		Std French

Number of speakers: 60

2nd Group (Few Talker Set)

SP. CODE	SEX	AGE	HEIGHT	WEIGHT	EDUC. LEVEL	SMOKING	LINGUISTIC CHARACTERIS.
----------	-----	-----	--------	--------	-------------	---------	-------------------------

BF	M	27	173	60	High		Std French
BO	M	33	178	75	High		Parisian
FA	F	28	162	52	High		Std French
JA	F	44	168	60	High		Std French
MH	F	29	163	56	High	Little	Parisian light
RO	F	26	163	65	High	Much	Std French
SC	M	23	182	75	High		Parisian
SH	M	34	180	65	High	Very little	Std French
SL	M	50	174	67	High		Std French
VI	F	26	172	78	High		Std French

Number of speakers: 10

3rd Group (Very Few Talker Set)

SP. CODE	SEX	AGE	HEIGHT	WEIGHT	EDUC. LEVEL	SMOKING	LINGUISTIC CHARACTERIS.
BF	M	27	173	60	High		Std French
BO	M	33	178	75	High		Parisian
JA	F	44	168	60	High		Std French
RO	F	26	163	65	High	Much	Std French

Number of speakers: 4

Legend:

SP. CODE	Speaker code
EDUC LEVEL	Education level
High	Higher education
Sec	Secondary school
SMOKING	Smoking habits
Stop 10	Stopped 10 years ago
LINGUISTIC CHARACTERIS.	linguistic characteristic (accent if any)
Std French	standard French speaking
South. F	Southern French speaking

Experimental recording conditions

A common methodology has been defined for all the recording sites in the Project [5][6]. French recordings were conducted in a soundproof room at the "Institut de la Communication Parlée" (Grenoble), by one operator.

Pilot recordings have been carried out at the beginning of the recording campaign, and analysed by the NPL laboratory (Teddington, UK), in order to assess the quality of the soundproof room and of the acquisition equipment. Regarding the whole acoustic chain, the background noise level (inter-words silence) is -70 dB below the maximum recording level for a signal peak level about -10 dB.

Test and calibration procedures have been conducted to ensure the constant quality of the recordings. Recordings have been performed using the "Continuous" mode, in order to collect utterances as well as extra-linguistic events (lip smacks, coughs, breath noise ...), and thereby satisfy the needs in "natural" speech.

Note: The amplification material used was true to the ESPRIT SAM n°2589 recommendation; ; the B&K 2603 amplifier being set on the 3V scale for all the sessions (except for calibration: 1V scale). The Input/Output board was a OROS AU22 board (OROS company, 38240 Meylan, France). The value of the analogical input gain of the board, is given for each item in the associated description file (usually 12 dB).

Recording session

A session starts when the speaker enters the soundproof room and ends when he leaves. A background audio recording (DAT) is running for the entire session. During the session the calibration signal is first recorded, then adjustment tests are performed with the speaker, and then the relevant corpus are recorded. In order to avoid speaker's fatigue and stress, a session is generally no more than a hour in length. Given this condition and the corpus to be recorded, speakers in each group have completed the following respectively:

- 1st group: 1 session
- 2nd group: 3 to 4 sessions (calibration being conducted again only if the sessions extend over 2 days)
- 3rd group: 1 session

Recording procedure

According to the group he belonged to, each speaker pronounced:

1st Group:

3 Passages	Pi, Pj, Pk
5 Sentences	(1 group of 5) Fi
100 Numbers	(5 groups of 20) N1 to N5

2nd Group:

10 Passages
25 Sentences (5 groups of 5)

5 times the 100 numbers
5 times the CVC S1 to S5

3rd Group:

All the CVC in context, (in the 5 contexts)
the context words in isolation (10 times)

Recording sequences

1st Group

CO calibration
N1 -> N2 -> N3 -> Fi -> N4 -> N5 -> Pi -> Pj -> Pk
(where P stands for Passage, F for a set of 5 sentences)

2nd Group

CO calibration
N1 -> N2 -> N3 -> S1 -> S2 -> N4 -> N5 -> S3 -> S4 -> S5 -> Pi ->
Pj -> Fi
(where P stands for Passage, F for a set of 5 sentences) and all together 5 times
this sequence (in several sessions)

3rd Group

CO calibration
Y1 10 times
T1 -> U1 -> T2 -> T3 -> U2 -> V1 -> W1 -> V2 -> U3 -> T4 -> T5 -> U4 -> V3 ->
W2 -> X1 -> X2 -> W3 -> V4 -> U5 -> V5 -> W4 -> X3 -> X4 -> W5 -> X5

Statistics

The corpus design was intended to cover the linguistic characteristics, and it has been decided not to have the whole corpus uttered by all the speakers.

1st Group: 60 speakers

In the 1st group, each speaker uttered 3 passages among the 40, 1 set of 5 sentences among the 50, and the 5 sets of 20 numbers. The average number of utterances is, therefore, for each passage 4 to 5, for each sentence 6, and for each number, 60.

1st Group

PASSAGES

SENTENCES

NUMBERS

Each speaker

pronounced: 3 5 100

Each item° has been

pronounced 4 to 5 times 6 times 60 times

item°: an item is the basic component of a corpus, it is so according to the corpus type: 1 passage or 1 sentence or 1 number...

2nd Group: 10 speakers

In the 2nd group, each speaker uttered 10 passages, 5 sets of 5 sentences, 5 times the 100 numbers and 5 times the all "logatomes". The average number of utterances is therefore for each passage 2 to 3, for each sentence 5, for each number 50 and for each "logatome" 50 again.

2nd Group	PASSAGES	SENTENCES	NUMBERS	CVC
-----------	----------	-----------	---------	-----

Each speaker

pronounced	10	25	100 * 5	82 * 5
------------	----	----	---------	--------

Each item has been

pronounced	2 to 3 times	5 times	50 times	50 times
------------	--------------	---------	----------	----------

3rd Group: 4 speakers

In the 3rd group, each speaker uttered all the "logatomes" in the 5 different contexts, and 10 times the context words in isolation. The average number of utterances is therefore 4 for each "logatome" in each context, and 40 for each context word.

3rd Group context	CVCCTX1	CVCCTX2	CVCCTX3	CVCCTX4	CVCCTX5
each speaker 10*10	82	82	82	82	82
each item 40 times	4 times				

All speakers together, the database contains 280 passages, 550 sentences, 11000 numbers, 4100 "logatomes", 1640 "logatomes" in context, and 400 utterances of the contexts words.

Volume of data:

1st Group:	15 Mb per speaker	total: $15 * 60 = 900$ Mb
2nd Group:	200 Mb per speaker (half speech signal and half laryngographic signal)	total: $200 * 10 = 2000$ Mb
3rd Group:	60 Mb per speaker	total: $60 * 4 = 240$ Mb

OVERALL TOTAL: about 3140 Mb

Speaker population:

60 speakers were selected among 70 people recorded. Various speaker information has been collected: sex, age, geographical origin, height, weight, education level, smoking habits... The coverage, regarding relevant factors, is balanced: 30 females and 30 males. All the French regions are more or less represented, with two poles at Rhone-Alpine and Parisian regions. Age distribution is good, between 20 and 62

years old. Three quarters among the speakers have an high education level and one quarter a secondary school one. None has a primary school level. This is balanced for women as well as men, so as the smoking habits: both one third of men and of women are smokers.

References

[1] ESPRIT n°2589 - SAM Final Report, Year three. SAM-UCL-G004, June 92.

[2] Zeiliger J., Serignat J-F., Autesserre D., Dolmazon J-M (1992). "EUROM1:une base de données "Parole" multilingue". Partie française. 19th J.E.P.,SFA, pp 303-306, 19-22 may, Bruxelles (Belgique).

[3] D. Autesserre, C. Meunier et J-F. Malet. "French corpus adapted from the SAM English version", 20 April 1990. SAM internal document, source: Institut de Phonétique, Aix-en-Provence.

[4] Barry W.J., Fourcin A.J., "Selection of speakers", SAM report: SAM-UCL-030, 20 May 1991.

[5] Tomlinson M.J.: "Guide to Database Generation - Recording Protocol", Final version - SAM-RSRE-015 - January 91. SAM internal document, source: RSRE, Malvern, UK.

[6] Zeiliger J. - Serignat J-F., "EUROPEC software (v4.1), User's Guide Release 4.1" - SAM-ICP-045 - March 91.

[7] Sherwood T., Fuller H. "Guide to EUROM_1 speech database", SAM report SAM-NPL-102, 20 April 1992.

[8] Zeiliger J., "Publishing CD-ROMs from EUROM_1", SAM-A report: SAM A/ICP/004/V1, 17 March 1993.

