

DELIVERABLE IDENTIFICATION

Identification number	LE4-8334 - D0.1
Type	Technical Report
Title	Baseline document for Interested Candidates for recording additional languages
Status	Final
Deliverable	D0.1
Work Package	WP 0
Task	Additional Task
Period covered	All project and after
Date	08/02/1999
Version	62
Number of pages	9
Author	Gaël RICHARD, Jérôme BOUDY (Matra Nortel Communications)
Workpackage (WP)/ Task (T) responsible	WP 0 – Jérôme BOUDY (Matra)
Project contact point:	Jerome Boudy, Matra Nortel Communications Rue J.P. Timbaud – 78392 Bois d’Arcy - France Phone: +33 1 34 60 , Fax: +33 1 34 60 88 32 E-mail: Jerome.Boudy@matranortel.com
CEC project officer	Mr. Mats Ljungqvist
Status	Public
Actual distribution	Consortium and CEC
Supplementary notes	
Key words	Database extension, external partners.
Abstract	This document briefly describes the procedures that external candidates (candidates for recording additional languages than those present in SpeechDatCar or for extending the recordings of a SpeechDat-Car language) should follow to become an external partner of the SpeechDat Car project and thus to achieve database exchanges with consortium members.
Status of the abstract	Public

Received on	
Recipient's catalogue number	

DOCUMENT EVOLUTION

Version	Date	Status	Notes
1.0	24/11/98	First draft	General concept. Presented in SPDC meeting
2.0	11/12/98	Second Draft	The document is better documented and integrates comments from all partners
3.0	12/01/99	Third Draft	The document is pre-final, and integrates the proposal agreed by the consortium
4.0	27/01/99	Final Draft	minor changes ; ready for vote
5.0	02/02/99	Final Draft	additional comments ;
6.0	02/02/99	Final Draft	Additional comments following negative vote from one partner.

SpeechDat-Car : *Baseline document for Interested Candidates for recording additional languages.*

6.1	02/02/99	pre-Final	Additional comments from SPEX.
6.2	08/02/99	Final	Additional comments from UPC

Contents

1. Introduction.....	4
2. General.....	4
3. Guideline for candidates.....	5
3.1 Admission procedure.....	5
3.2 Exchange procedure between the external partner and all PRL partners who signed the Commitment Letter	6
4. Additional information	7
4.1 Additional language.....	7
4.2 Future Status in SpeechDatCar	8
4.3 Requirements.....	8
4.4 Commitment Letter (CL).....	9
5. Contacts	9
APPENDIX : List of PRL partners.....	10

1. Introduction

This document briefly describes the general procedure that external candidates for recording additional languages than those present in SpeechDat-Car or for extending the database of an existing language should follow to become an external partner of the SpeechDat Car project and thus to achieve database exchanges with consortium members. This document is organised as follows. The next section provides the reader with a general summary. Section 3 gives the guidelines for candidates and details the admission and database exchange procedures. The fourth section gives additional important information.

2. General

An External organisation may propose one of the following possibilities to be eligible as an external partner :

- Record an additional language (see below for preferred languages list).
- Record an additional European language.
- Extend the database of an existing language (i.e extend the number of speakers). In this case, the candidate will need the agreement of the partner responsible of this language within SpeechDat-Car to be eligible.

The admission of the external organisation as an external partner will be effective if all requirements are fulfilled (see sections 3 and 4)

3. Guideline for candidates

3.1 Admission procedure

The general procedure for proposing the recording of an additional language or an extension of an existing language is the following :

1. The candidate mentions his interest for recording an additional language by sending a letter to the prime contractor (project leader) with copies to each PRL partner¹ of the SpeechDatCar project.
2. If the language proposed by the candidate is not in the preferred list (see section 3.1), the speechDatCar consortium (PRL partners) will vote on the acceptance of the proposed language (50 % of positive votes is required). If the language is present in the preferred language list no vote is needed since the language is de facto acceptable. If the language proposed is already present within SpeechDat-Car (i.e. the candidate proposes an extension of a database), the candidate needs the formal acceptance of the partner responsible of that language, otherwise his proposal will not be eligible.
3. If the proposed language is accepted, the SpeechDatCar project manager sends to the candidate the SPDC database specifications. These specifications might however be updated until the platform validation and pre-validation of internal SpeechDat Car databases are completed (i.e. until April 1999). If the candidate agrees on all specifications and requirements, he sends 2 (two) copies of a Commitment Letter (CL, see next section for additional explanation on the Commitment letter) to all PRL partners. Each PRL partner then returns one of the copy signed by an authorised person, if he accepts the database exchange with the candidate.
4. When the candidate has received all signed copies of his Commitment Letter, he informs the SpeechDatCar project manager who will then send to all partners an official letter confirming the admission of the candidate as an external partner. The commitment of the external and internal partners becomes effective only when the official letter confirming the admission and listing the partners who will exchange their database is sent to all partners.

It is understood here that the external candidate will exchange his database with all PRL partners who have signed the Commitment Letter but that there is no guarantee that all PRL partners will sign it.

¹ *A PRL partner is a Partner Responsible for the recording of a Language*
SpeechDat-Car : *Baseline document for Interested Candidates for recording additional languages.*

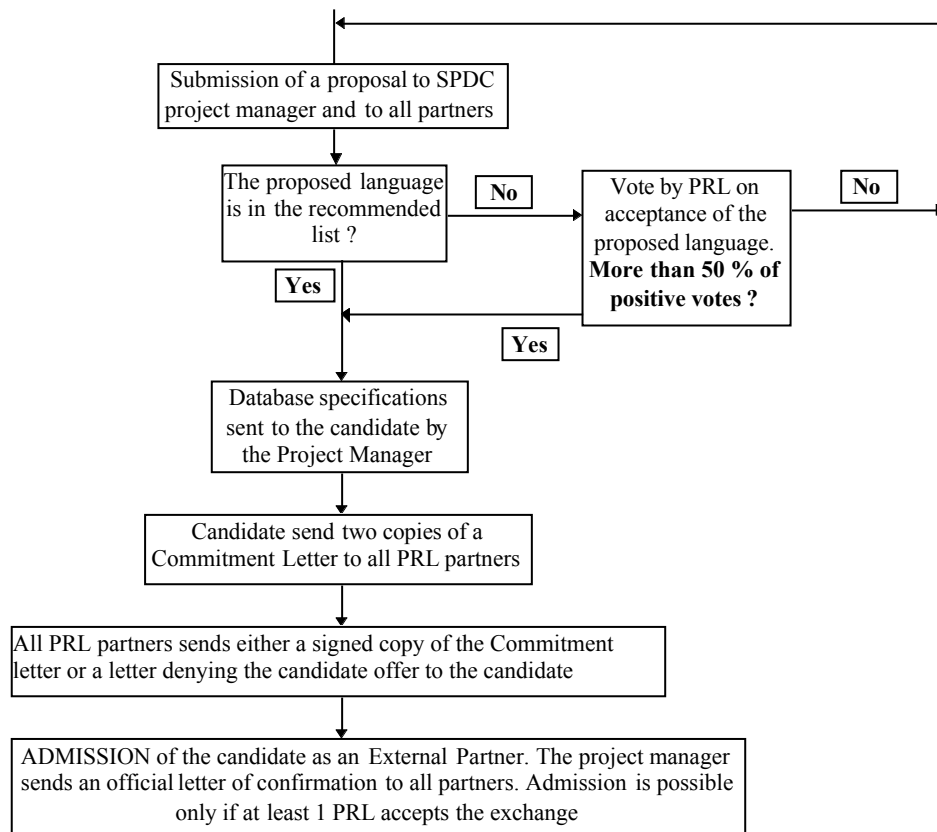


Figure 1: Admission procedure for external partners

The candidate should be aware that the recording start date will not be possible before April 1999, since platform validation and database pre-validation will not be completed before this date.

3.2 Exchange procedure between the external partner and all PRL partners who signed the Commitment Letter

The exchange procedure between accepted External partner's database and PRL partners who signed the commitment letter will follow the following steps :

1. Completion of the External Partner's database
2. Validation of his database : the database will be validated with the same procedure as for the other languages. This validation is performed by SPEX² at cost of 6 kECU (incl. prevalidation). SPEX will edit a validation report for each evaluated database.
3. If the database does not fulfil all requirements, a vote is carried out to accept the database. If rejected, the External partner is requested to correct the database problems.
4. After proper validation, the External partner produces copies of his database and sends them to all PRL partners who have completed their database (recordings + validation) on the same day and informs all partners of the SPDC consortium. The External partner sends a copy to the other PRL as soon as they have completed their database (recordings + validation).
5. All PRL partners send their database to the external partner within two weeks (after reception of the External partner database).

² SPEX : Validation Centre.

SpeechDat-Car : Baseline document for Interested Candidates for recording additional languages.

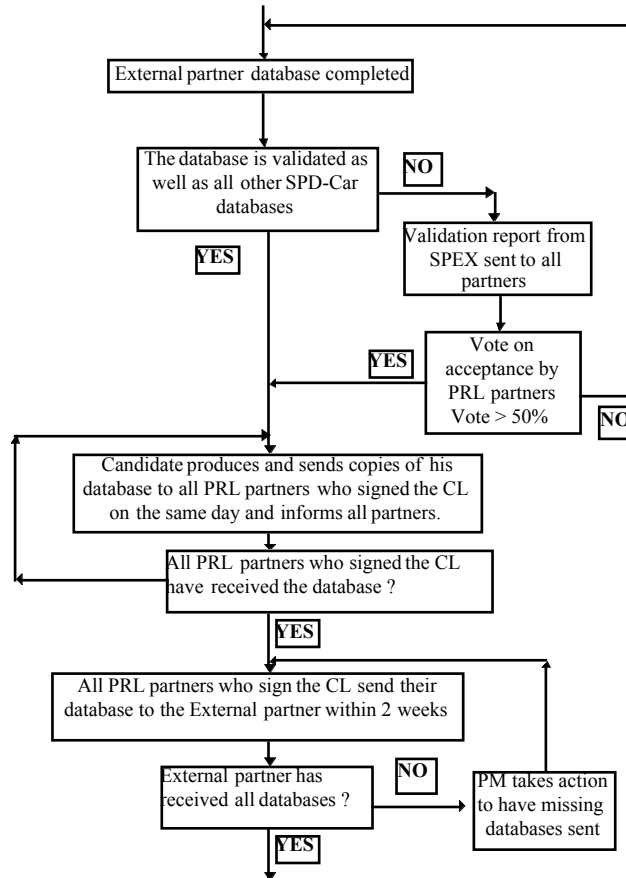


Figure 2: Exchange procedure with external partners for the case where the external partner completes his database after all the PRL who signed his Commitment Letter.

4. Additional information

This section provides additional information concerning the preferred languages, the commitment letter, the status of the external partner and requirements.

4.1 Additional language

The SpeechDatCar consortium established a list of *preferred languages* resulting from a vote of each participating partner. The order of the *preferred languages* is the result of a vote involving all partners of the consortium. The table below summarises the result of the vote (each partner could give up to three preferred languages):

Languages	Ranking
American English	1
Swedish, Japanese	2
Chinese	4
Norwegian, Portuguese	5
Brazilian Portuguese	7
Catalan, Arabian	8

In addition of this list, a candidate may propose additional European languages and/or extension of an already existing language within the SpeechDat-Car consortium. If the last choice is made, then the candidate will be asked to produce detailed specifications based on the existing specifications and the vote by all partners will be made after presentation of these specifications. A straightforward way of extending a language will be to strictly follow SPDC specifications and simply record additional speakers, but other solution are possible.

However, it should be clear that if an extension of a language is proposed, the candidate will need a formal approval of the Partner Responsible of that language prior to any vote on acceptance of the proposal.

4.2 Future Status in SpeechDatCar

If a candidate proposal is accepted by the consortium (see below for guidelines for proposals), the future status of the candidate will be “*External Partner*”. This status will confer to the organisation the right to attend the specific SpeechDatCar workshops on database status and to participate to discussions. Though, the “*External Partner*” will not be a “partner” of the project and will not be eligible for funding from the project or from the EC for their database recordings. They will also not be able to vote on decisions taken by the consortium. However, in the case where another candidate is declared, the accepted “*External Partner*” will as the other “*Internal Partner*” receive a Commitment Letter from the candidate and will have the possibility to accept or reject databases exchange with the new candidate.

4.3 Requirements

To be acceptable, the candidate will have to accept to follow the SpeechDatCar specifications (car speech database, car environments and speaker coverage) that are detailed in document D.1.12 “Specification of SPDC database items” and those related to the recording platforms reported in document D1.4 (This document gives, in particular, details about microphone position in the car, microphone type, etc...). A public version of these two documents will be made available to any potential candidate. The candidate should be aware of the fact that after acceptation, he will have to acquire a CD ROM (most probably accessible through ELRA) that will contain :

1. Baseline document for Interested Candidates for recording additional languages (this document)
2. Database specifications (document D.1.12)
3. Platform specifications (public version of Document D.1.4) with an additional user guide
4. Transcription conventions (D1.3.2a)
5. Validation criteria (D1.3.1)
6. Format specifications (D1.3.3)
7. Software of the Mobile platform
8. Software of the fixed platform
9. Platform Validation procedure

The approximate price for this CD-ROM will be around 6 kEuro.

4.4 Commitment Letter (CL)

The commitment letter will be sent by the candidate to all PRL partners. This letter will contain :

1. A commitment to deliver for exchange his speech database as specified by SpeechDat-Car and that the candidate accepts all conditions of this document and the SpeechDat-Car database specifications and validation procedures.
2. A commitment to deliver his completed (recordings + validation) database to all PRL partners who have completed their database (recordings + validation) on the same day and informs all partners of the SPDC consortium. A commitment to deliver his completed database to the other PRL as soon as they have completed their database (recordings + validation).

The commitment letter (CL) will be sent to all partners within 2 weeks after acceptance of the language proposed by the candidate. All partners will return the signed copy of the CL or the letter denying the offer within three weeks after reception of the CL.

5. Contacts

- Prime Contractor :
Matra Nortel Communications
- Project coordinator :
Jérôme Boudy
MATRA NORTEL COMMUNICATIONS
Rue Jean-Pierre Timbaud
BP 26
78392 Bois d'Arcy cedex - France
Phone : +33.1.34.60.86.69
Fax : +33.1.34.60.88.32
Email : jerome.boudy@matranortel.com

APPENDIX : List of PRL partners

Full legal name, legal status	MATRA NORTEL COMMUNICATIONS,
Registered Address Street Name and No. Town/City Post Code / Cedex Region + Region Code (a) Country Code (b)	Rue Jean-Pierre Timbaud...BP 26..... Bois d'Arcy cedex..... 78392..... FRQ - R2103..... FR.....
Contact Persons Name, First Name Job Function, Title Telephone Number Fax Number E-mail	BOUDY Jérôme +33.1.34.60.86.69 +33.1.34.60.88.32..... jerome.boudy@matranortel.com
Name, First Name Job Function, Title Telephone Number Fax Number E-mail	RICHARD Gaël +33.1.34.60.79.55 +33.1.34.60.88.32..... gael.richard@matranortel.com

Participating Organisation (legal entity)	Lernout & Hauspie Speech Products, NV..... L&H.....
Registered Address Street Name and No. Town/City Post Code / Cedex Region + Region Code (a) Country Code (b)	Sint-Krispijnstraat 7..... Ieper..... B-8900..... West-Vlaanderen (BEWVL)..... BE.....
Contact Person Name, First Name Job Function, Title Telephone Number Fax Number E-mail	Robrecht Comeyne Group Manager Databases +32 57 22 95 71..... +32 57 20 84 89..... robrecht.comeyne@lhs.be.....

Participating Organisation (legal entity)	Robert Bosch, GmbH
Full legal name, legal status Participant's short name	BOSCH
Registered Address Street Name and No. Town/City	Robert-Bosch Platz 1 Gerlingen-Schillerhöhe

Post Code / Cedex	70839
Region + Region Code (a)	DE 111
Country Code (b)	DE
Contact Person	
Name, First Name	Arévalo, Dr. Luis
Job Function, Title	Coordinator DSP-Group
Telephone Number	+49-711-811-6116
Fax Number	+49-711-811-7602
E-mail	arevalo@fli.sh.bosch.de

Participating Organisation (legal entity)	
Full legal name, legal status Participant's short name	Vocalis, Ltd Vocalis
Registered Address Street Name and No. Town/City Post Code / Cedex Region + Region Code (a) Country Code (b)	Chaston House Mill Court Great Shelford, Cambridge CB2 5LD CAMBRIDGESHIRE R7401 GB
Contact Person Name, First Name Job Function, Title Telephone Number Fax Number E-mail	SIMPSON Andrew Responsible for Language Resources and Technology +44 (0) 1223 846177 +44 (0) 1223 846178 andrew.simpson@vocalis.com

Participating Organisation (legal entity)	
Full legal name, legal status Participant's short name	KNOWLEDGE, S.A. KNOW
Registered Address Street Name and No. Town/City Post Code / Cedex Region + Region Code (a) Country Code (b)	N.E.O PATRON-ATHINON 37 PATRAS 26441 ACHAIA - RA232 GR
Contact Person Name, First Name Job Function, Title Telephone Number Fax Number E-mail	Dr. Stefanos Michos Assistant Director of R&TD Projects +30.61.452820 +30.61.453819 E-mail: smichos@knowledge.gr

Participating Organisation (legal entity)	
Full legal name, legal status Participant's short name	Universitat Politècnica de Catalunya UPC
Registered Address Street Name and No. Town/City Post Code / Cedex Region + Region Code (a) Country Code (b)	Jordi Girona 31. Edificio R Barcelona 08034 Barcelona ESB ES
Address	

Street Name and No. Town/City Post Code / Cedex	Jordi Girona 1-3 Edificio D-5 Barcelona (Spain) 08034
Contact Person Name, First Name Job Function, Title Telephone Number Fax Number E-mail	Moreno, Asuncion Profesor Titular 34 3 401 64 37 34 3 401 64 47 asuncion@gps.tsc.upc.es

Participating Organisation (legal entity)	Uni. of Aalborg – Center for Person Kommunikation
Full legal name, legal status
Participant's short name	CPK.....
Registered Address Street Name and No. Town/City Post Code / Cedex Region + Region Code (a) Country Code (b)	Fredrik Bajers Vej 7A Aalborg 9220..... DK080..... DK.....
Contact Person Name, First Name Job Function, Title Telephone Number Fax Number E-mail	Dalsgaard, Paul Professor..... +45 96358642..... +45 98151583..... pd@cpk.auc.dk.....

Participating Organisation (legal entity)	NOKIA
Full legal name, legal status Participant's short name	NOKIA Corporation, Research Center; Oyj (Ltd) NOKIA
Registered Address Street Name and No. Town/City Post Code / Cedex Region + Region Code (a) Country Code (b)	Nokia Research Center P.O.Box 407 / Heikkilantie 7 HELSINKI..... FIN-00045 NOKIA GROUP..... Uusimaa FI11..... FI.....
Contact Person Name, First Name Job Function, Title Telephone Number Fax Number E-mail	Laurila, Kari..... . Group Manager..... +358 3 272 5876..... +358 3 272 5897..... kari.laurila@research.nokia.fi.....

Participating Organisation (legal entity)	
Full legal name, legal status Participant's short name	ALCATEL BUSINESS SYSTEMS, SA ALCATEL
Registered Address Street Name and No. Town/City Post Code / Cedex Region + Region Code (a) Country Code (b)	32 avenue Kléber COLOMBES cedex 92707 ILE DE FRANCE FR
Contact Person Name, First Name Job Function, Title Telephone Number Fax Number E-mail	BOURMEYSTER Ivan Project Manager +33.1.46.52.17.06 +33 1.46.52.80.25 ivan.bourmeyster@art.alcatel.fr