ETRW on Speech Technology
in Language Learning (STiLL)
Marholmen, Sweden
May 24-27, 1998

Research in Speech Processing for Breton Language Training

Jean-Pierre Messager, Hervé Gourmelon, Guy Mercier, Jacques Siroux
LLI/IRISA, CORDIAL team – KGB project*, ENSSAT, Lannuon, France
{Jean-Pierre.Messager, Guy.Mercier, Jacques.Siroux}@enssat.fr

Abstract

The aim of the research reported on here is to develop educational software tools for Breton language[1], based on speech technology. These programs are intended to be used by children in classrooms in order to improve their language knowledge (lexical, phonological, morpho-syntactic and prosodic). As they are tools involved in a pedagogical relationship with teachers, individual learners of the Breton language may have an interest in using them too.

Three main applications are planned : a bilingual vocal dictionary based on speech synthesis, a prosodic trainer and a spell trainer (dictation program) using text to speech synthesis.

1. Introduction

Breton is a celtic language, traditionally spoken in the west half of Brittany. At the beginning of twentieth century about one million persons were speaking Breton, most of them as a native and unique language. During the two next generations, french linguistic policy as well as schools development (where French were the only admitted language) make French the dominant language on the whole country. Now, about three hundred thousand persons speak Breton, most of them being more than sixty years old.

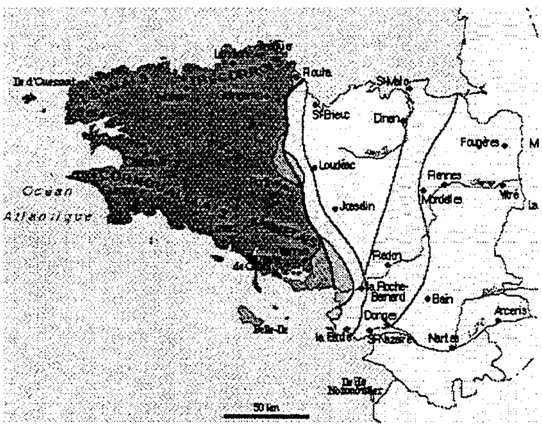

Figure 1. Evolution of the Breton speaking area during past centuries

During the past twenty years, Breton usage in schools increases quickly, going from few to 4000 pupils either in bilingual schools or in breton schools. Considering pupils receiving breton courses, the total amount of children concerned, more or less, by Breton teaching is about 23000. It is now taught in the whole Brittany, as well as (for adults) in some cities in France and overseas.

The main problem of Breton teaching, is the influence of the preeminent french language in linguistic environment (TV, newspapers, administrations, public life, etc.) The French influence on the *new* Breton spoken now by young generation is perceptible on syntactic, lexical, as well as phonological points of view. Even teachers are rarely native speakers, and can have problems in such competences.

The KGB project (*Kenaoz ar Gomz e Brezhoneg* : Speech Synthesis in Breton Language) is a joint initiative of TES (*Ti-Embann ar Skoliou Brezhoneg* : an institutional producer of Breton teaching materials) and the research team CORDIAL (part of IRISA Labs, working on human/computer interaction). The emergence of high quality, robust speech technologies, makes them suitable to take place in classrooms for language learning.

Several applications have been scheduled as steps for the research project, all of them involving speech synthesis, speech analysis and/or natural language processing.

2. A Bilingual Vocal Dictionary

The first application, now on an evaluation stage, is a vocal Breton/French dictionary containing about 40000 definitions on both sections. Giving ability to navigate on both breton and french lexical entries, it allows to hear diverse breton pronunciation variants of breton words (figure 2).

2.1 Data

The original data is taken from a paper book work [1]. It was available on electronic form, without any document structuring but typographic ones. The first point was to parse the original rich text (translated from *Word* proprietary format to ANSI *Rich Text Format*, and then to HTML) to obtain a lexical database with logical distinction of both words definitions and elements within definitions (as well as variants forms for each entry).

The typographic conventions commonly used in dictionary printing give us the ability to split definition between several fields :

*This project was funded by Breton Regional Council and European Community.

[1]Breton is a celtic language, belonging to brittonic group, like welsh or cornish.

- Key field : the word himself (and often variants) ;
- Grammatical category ;
- Phonetic of common pronunciations ;
- Mostly unstructured text of definition (including sometimes some subdefinitions).

The style coding of text (mostly of the unstructured part) is a subset of HTML and the phonetic coding, remains the one of the *Windows True Type* SIL-IPA font, as used in the original document.

2.2 Lookup System

The search of a word in the dictionary can be performed by different ways, if one wants something better than exact matching (because of orthographic mistakes or variants, derived forms for nouns, verbs, etc). Between dynamic programming (to compute distances between words) and search based on regular expressions, we opted for the last one, because of regularity of variants in Breton (three different orthographic systems are used, french influence make some kind of mistakes quite usual, such as consonants doubling).

Another point, specific to celtic language, which can be easily handled by regular expression is the grammatical mechanism called *mutation*. This variation, very usual in Breton, affect the initial consonant of a word given its left context. (*"bugale"* the word for *children*, can be written *"vugale"* when preceded by the definite article).

For example, when searching for a word such as *"vreiz"*, that one can hear or read, the system will try to match the regular expression `(B|V)rei(Z|ZH)` (possible mutational origin of initial V, and Z/ZH orthographic classical variant) against the words list, so the word *"BREIZH"* (Brittany) will be found.

This provides, moreover, the ability for the user to use wildcards in order to explore the dictionary (one may look for all words beginning by any sequence, or finishing by another – for poetry – or containing any radical... Even conditions about words not matching some expression can be specified).

2.3 Speech Synthesis

The phonetic information for each definition on the Breton/French part of the dictionary, gives the usual pronunciations for main dialects of Breton, (*Trégor/Goëlo, Léon, Cornouaille, Vannetais*). A synthesizer for breton language have been set up for some months ; we are using the synthesizer MBROLA [2]. As it is based on diphone concatenation technology, we had to record each combination of two phonemes in Breton, then segmenting the signal files.

No text to phonemes translation is needed here, since data include, for each word, pronunciations in IPA style including stress marks. We built a small rule-based prosodic model (at word level) to generate appropriate prosodic parameters (mainly duration and piecewise linear frequency curve), in order to control the acoustic synthesizer which generates then a natural signal (moreover the user can change the global speech rate, the usual conversation speed being the fastest).

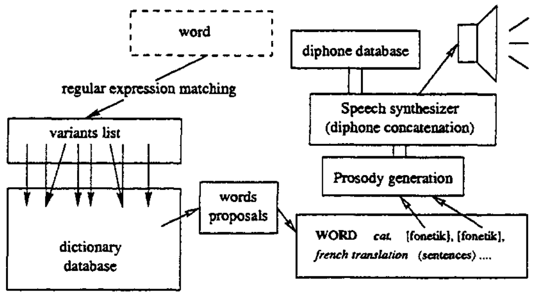

Figure 2. Diagram of the dictionary databases and processing

2.4 Interfaces

Two front-ends to lexical and acoustic database have been developed : one for *Windows*, with graphics, animations and an interface as friendly as possible since aimed users are mostly children (figure 3); the other is a WEB version, giving us faster development and evaluation.

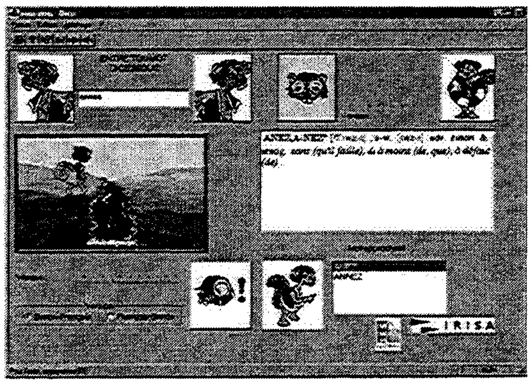

Figure 3. The Windows interface

3. Further Projects

3.1 Prosodic Trainer and Evaluator

Breton teachers focus now on acquisition of prosodic competences in Breton by learners, given the fact that french scheme of intonation is influencing new speakers pronunciation. A prosodic trainer software is planned for next months.

It will allow the learner to hear a sentence recorded by an experimented or native speaker, then, after repeating it through a microphone, receive evaluation of quality of his own prosody as well as graphic feedback on difference between his prosodic scheme (lengths and frequency variations) and the reference signal.

3.2 Automatic Spell Trainer

This project, which has ever been developed by CORDIAL team for French [3] and is on a rewriting stage, is an orthographic dictation program. Through a multimodal interface, the user is invited to write a text the software pronounces. This one scans the keyboard entry to ensure synchronization between typing and dictation, and therefore can repeat current sentence or decide to pause for a moment.

Central elements for such a program are, a fine multimodal interface, high quality speech synthesis on a sentence basis and liable detection of spelling and grammatical mistakes.

4. Speech Synthesis for Language Teaching

Most text-to-speech synthesizers have now a quite good acoustic rendering, but fail in prosodic naturalness. It can be good enough for some kind of applications such as meteorology prediction broadcast, telephone information services, etc. where a quite "robotic" voice may even be desirable (for pragmatic reasons).

Beyond such applications, where only intelligibility is needed, the aim is to bring speech synthesis to the classrooms for language learning, a place where phonology and prosody can't be underestimated, as they are an important part of the competences to be taught.

NLP technologies and linguistic foreground, especially for generation of prosody by lexical, syntactic and semantic analysis, for the moment can't provide such level of quality for arbitrary text, without any additional information. In collaboration with CSTR (Edinburgh University Speech Technology Laboratory), where a formalism, called STML (Spoken Text Markup Language) [4], to enrich text for speech applications has been designed ; we are investigating the ways a text can be tagged for an educational purpose in the field of language learning.

Such tagging has to provide information to help the linguistic processing part of speech synthesis to generate adequate prosody and rhythm. It does not have of course to supersede normal natural text analysis (otherwise the user would have to specify an excessive amount of information such as evident syntactic boundaries, pauses, etc). The aim is to provide a way to help the automatic processing when information is impossible or difficult to obtain given the syntactic context; furthermore it makes possible to express some kind of variability which cause can be given by anything else than naturalness.

Here is a quite enriched text, with tags expressing enumeration, voice variation and language switching :

```
<STML>
<LANGUAGE ID=français> <SPEAKER ID=man1>
<GENRE TYPE=plain>
Here is a text in STML format, including
some useful <EMPH> tags </EMPH>,
such as :
 <GENRE TYPE=list>
 language specification <DIV TYPE=item>
   <SPEAKER ID=woman2>
   change of voice <DIV TYPE=item>
   </SPEAKER>
 text category <DIV TYPE=item>
 <DEFINE WORD="etc" PRO="et'setr@"
 SCHEME=sampa> etc.
 </GENRE>
<LANGUAGE ID=brezhoneg>
<SPEAKER ID="Annie Ebrel">
 Yec'hed mad !
<LANGUAGE ID=cymraeg>
 Yeched da !
</STML>
```

Such enrichment is to speech synthesis what typography is to printing. So a typographically enriched view of this text is quite straightforward :

> " Here is a text in STML format, including some useful *tags*, such as :
>
> - language specification ;
> - change of voice ; *[feminine voice]*
> - text category ;
> - etc.
>
> Yec'hed mad ! *[Breton sentence]*
> Yeched da ! *[Welsh one]* "

STML gives ability to express a large set of linguistic information, which can be useful for speech synthesis of texts still enriched on others bases (typography, computer generated texts, databases) [5], as well as for texts where user exigencies are high enough, so he would want to prepare text data. A well-designed markup language with an adequate graphical front-end could help teachers to obtain a good enough speech synthesis to fit their pedagogical needs.

References

[1] F. Favereau (1992). Geriadur ar Brezhoneg A-Vremañ, *Skol Vreizh*, Montroulez.

[2] T. Dutoit, V. Pagel, N. Pierret, F. Bataille, O. van der Vrecken (1996). The MBROLA project : Towards a Set of High-Quality Speech Synthesizers Free of Use for Non-Commercial Purposes, *proceedings of ICSLP'96*, Philadelphia, 1996.

[3] M. Guyomard, J. Siroux, D. Pernici, C. Royer (1997), A General Public Application of Pedagogic and Linguistic Vocations of Speech Synthesis : Ordictée, *Proceedings of ROCLING X*, Taipei, Taiwan, 1997.

[4] R. Sproat, P. Taylor, M. Tanenblatt, A. Isard, (1997). A Markup Language for Text-To-Speech Synthesis, *CSTR Edinghburg*, 1997.

[5] P. A. Taylor, A. W. Black (1994). Synthesizing Conversational Intonation from a Linguistically Rich Input, *Second ESCA/IEEE Workshop on Speech Synthesis*, New York, 1994.