FROM PHONEME TO GRAPHEME: DIAGNOSIS IN A DICTATION

Conception Santiago-Oriola and Guy Pérennou
IRIT, Université Paul Sabatier, Toulouse (FRANCE)

ABSTRACT
The main difficulty for the elaboration of a dictation system consists in modeling the errors and the associated explanations provided to the learner. On these bases, an experimental DICTOR system is being developed as an assistant tool to learn French language spelling. DICTOR includes an automatic checking tool based on a stochastic alignment algorithm and French written linguistic knowledge. This paper focus on both aspects. We discuss the spelling learning issues due to the no one-to-one correspondence between the utterance and the written text. Then the principles of the spelling difficulty groups (sdg) and associated explanation rules are presented. Finally, we report a DICTOR observation in a school environment for the French language spelling.

1. INTRODUCTION
Research on spoken and written language engineering offers nowadays reusable tools and linguistic knowledge (such as lexicons, alignment tools, …) for the development of new interactive applications about language. At the same time, multimedia tools for the integration of vocal synthesis systems emerge. This context and the linguistic knowledge maturity have led our research team to develop an interactive vocal dictation system in the framework of language learning.

The main difficulty of the elaboration of such a system consists in modeling the errors and the associated explanations provided to the learner.

On these bases, an experimental DICTOR system is being developed at IRIT as an assistant tool to learn French language spelling. DICTOR includes an automatic checking tool based on a stochastic alignment algorithm and French written linguistic knowledge. This paper focus on both aspects.

The spelling learning problems due to the no one-to-one correspondence between the utterance and the written text are discussed. Then the principles of the spelling difficulty groups (sdg) and associated explanation rules modelization are presented. Finally, we report an observation of the system in a school environment.

2. THE DICTATION EXERCISE
In language learning, one of the basic exercise is dictation. Dictation can be considered as a transcription exercise of an utterance. This task rises several multidisciplinary problems from aural signal perception to its interpretation.

Firstly, the learner must recognize the spoken units such as words, phrases or sentences as part of the language being learned.

Then, the sounds heard must compose a series of words. This word segmentation can be a problem for learners because the spoken language has no one-to-one and clear acoustic parameter indicating the beginning or the end of words. Moreover, some liaison phenomena, which appear at the conjunction of a phoneme on a lexical frontier, or assimilation ones between phonemes due to adjacent words intervene. The oral segment can then correspond to several homophonic non homographic phrases. For example, [ləpɔliˈtɛm] can lead to le petit ami (the boyfriend) or le petit tamis (the small sieve). Another penalizing factor is the presence in the utterance of words or expressions not belonging to the learner’s lexicon. The adaptation of the dictation text to the learner’s cognitive and learning levels is essential for this phase.

Thirdly, once the series of words is determined, it must be meaningful. The terms activated into memory must be in a near relation with the theme of the dictation. In French, words such as saule (willow), sole (sole), sol (ground) can be pronounced [sɔl] depending on the speaker’s region or the language acquisition level. This phase can lead to confusions in the following terms. To distinguish between homophonic words, the context is fundamental.

The latest phase consists in the effective transcription according to the letters provided by the language being learned and to the learner’s conceptual model of the spelling. The spoken language can be very influent: the transcription can be biased by the learner’s pronunciation related to the spoken dialect.

Those various problems show the need in terms of linguistic knowledge to efficiently model the transformation of an utterance in a written phrase.

3. AN INTERACTIVE DICTATION SYSTEM
To specify an interactive dictation system using vocal restitution systems, we differentiate two principal phases in the dictation exercise. The first one consists in the learner’s typing of the spoken dictation text, through the keyboard of the computer, followed up by the diagnosis phase (Figure 1). Interactivity, explanations pertinence and diagnosis adaptation to the learner’s knowledge level are the strong criteria for the design of this interactive system.

![Figure 1. Synoptic of an interactive dictation system](image)

As shown in Figure 1, the methods, used by the teacher, to design the exercise to do for example, and the learner’s model are taken into account to permit a differentiate pedagogy. Here, the first step of our system (called typing) is a dictation from the computer to the learner. The dictation text is enunciated by a speech synthesis system and the learner transcribes it via the
After this phase, the second step consists in evaluating the learner’s input for proposing him some explanations depending on the errors committed and the learner’s spelling knowledge level. This evaluation is based on a string alignment algorithm between the correct dictation and the learner’s input and on different corrective strategies (see Figure 2).

The diagnosis module generates an annotated string corresponding to the corrected string, completed with the explanations. To implement it, a lexicon of the learner’s possible errors and a set of linguistic resources, containing among other data the explanations, are necessary.

### 3.1. String comparison algorithm

Our goal is to compare the transcript made by the learner, on an orthographic format, and the dictation correct text to induce the eventual errors with string alignment methods.

An alignment algorithm VERITEXT [2], derived from the VERIPHONE system (which realizes a phonetic alignment), based on rules has been developed. It relies on a stochastic model described in [3].

The model build on stochastic rules allows the comparison between a source string $X=x_1x_2…x_n$ with $x_i \in A$ and an observation string $Y=y_1y_2…y_m$ with $y_j \in B$, $A$ and $B$ being two distinct alphabets.

For aligning these two strings, two modules are used:

- The spelling coder which models the passage from elements of the $A$ alphabet into elements of the alphabet $B$.
- The typographic channel to transform $Z$ into $W$ belonging to $B^*$.

The alignment is obtained when $W$ and $Y$ are identical.

VERITEXT makes an alignment and points out if an erroneous rewriting has been used.

### 3.2. Correction strategies

The correction strategies allow the system to choose the kind of presentation of the alignment results and linguistic knowledge. They may fit to the learner’s knowledge level and cognitive abilities and to the teacher’s pedagogical methods. All this information cooperates with the learner’s model and the pedagogical module.

### 4. MODELIZATION OF THE ERROR DIAGNOSIS KNOWLEDGE

We focus here on the second step of this exercise, i.e. the diagnosis of the correctness of the learner’s input. Our concern is multiple: how to determine the more precisely the learner’s errors and how to present the results of our diagnosis.
To answer satisfactorily to our first concern, we must analyze the linguistic knowledge of the French language to settle down its potential difficulties. A modelization of the French language into group of letters with spelling problems (sdg) and associated rewriting rules is here presented.

The second concern relies on corrective and explanation presentation strategies which may be adapted to the learner’s capacities. Some explanations are related closely to the rewritings and therefore to the sdg.

4.1. Graphic vs. Phonetic material
As defined above, dictation is a “graphic” transcription process of utterance. In this paper we won’t consider the comprehension problems in the neurolinguistics sense.

So dictation is only considered as the establishment of a correspondence between speech and text. To do it, the orthographic alphabet must be faced to the French phonetic alphabet. If we consider those two alphabets, some facts appear:

- the orthographic one has only six vowels (a, e, i, o, u, y) while the phonetic one has generally sixteen [a], [æ], [e], [ɛ], [i], [ɪ], [o], [ɔ], [ʊ], [ʌ], [ə] and [æ].
- There are twenty consonants in the graphic one but c and q correspond to the same phoneme [k].
- The letters c and g each have two different phonic values, [k] and [z] –for example café [kaʃ] (cafe) et cerise [sɛʁiz] (cherry)– [ɡ] and [ʒ] –as in garçon [ɡaʁsõ] (boy) and girafe [ɡiraʃ] (giraf)–.
- In the orthographic alphabet there is no nasal vowel.

Catach has demonstrated that the foundations of the French written language are phonogramic, that is its basic unit is the grapheme and particularly the phonogram, i.e. the grapheme transcribing a phoneme [1]: 80 to 85% of the characters of a text correspond to sounds.

So we can see that the French spelling has many difficulties to overcome and therefore is complex to learn and master.

4.2. Identification of the sdg
During the dictation exercise, the learner must write down graphemes of the language being learned from the meaningful phonemes he can hear.

Errors can have varied origins such as for example:

- He can make bad phonogram transcriptions because either he doesn’t know them (incomplete lexicon) or his spelling knowledge is incomplete (for example, [fɔʁmæsi], *farmassie instead of pharmacie (drugstore)).
- The relations between grammatical categories aren’t respected (*ils passe instead of ils passent (they pass))
- The learner doesn’t know some words or segments the text in an unexpected way in relation with the original text [ɛnɛvʁ] can lead to *un névier instead of un évier (a sink),

So, the sdg alphabet must allow to model all the possible errors generated by the phoneme-to-grapheme shift and to model the phonetic and phonologic errors due to an incomplete knowledge of the lexicon, syntax and morphosyntax of the French language. Moreover, this sdg model must be opened to consider different pedagogues’ strategies as already mentioned.

For example, we must define correctly the sdg to render the difference between errors such as *troi, *quatre plat, *tu passe where the left final s can’t be justified in the same way:

- Trois (three); this adjectival must be always written in that way.
- quatre plats (four dishes); the ending s is the morphogram of the plural of the noun plat.
- tu passes (you pass); here, the ending s is the morphogram of the concord of the verb passer et the second singular person.

This example shows that, at least, three sdg must be defined corresponding to the unpronounced ending s.

To constitute our alphabet, we have studied each of the 172 graphemes of the French language, adding them complementary information (word position and pronunciation) to obtain a first list of sdg. Then we complete it with morphograms and logograms to allow the teachers to adapt the explanations to the learners. We have currently a set of 564 sdg for the French language learning.

4.3. The associated rewriting rules
To detect the learner’s eventual errors, the erroneous possible rewritings of each sdg must be identified.

The principal problem is to attach to each sdg the rewritings set, those corresponding to different transcription possibilities more or less evident depending on the learning level. The learner’s possible behaviors must be predicted to adapt the necessary explanations of the exercise. Reviewing each sdg category can help us to settle some of the problems.

The difficulties concerning the rewriting rules associated to phonogramic sdg are particularly linked to the vowels. The beginner learners of the French language are tempted to add the h letter at the begin of the difficult words starting with a vowel or between vowels at a hiatus position: therefore a word as aorte (aorta) can be transformed in haorte, ahorte, haborte or even aorthe. The rewriting rules take into account all those kinds of errors. We need then rewriting rules such as:

\[
\begin{array}{|c|c|c|}
\hline
\text{Sdg} & \text{Correct} & \text{Error} \\
\hline
[a] & a & ha \\
[oVx] & o & ho \\
t & t & th \\
\hline
\end{array}
\]

Table 1. Examples of rewriting rules for an insertion of an h.

The French beginner learner can have some problems with the liaison phenomenon. The liaisons can prevent him from recognizing the words starting with a vowel, he can then write, erroneously, *un navion or *des zavions instead of un avion (a plane) or des avions (planes). We have added some sdg to manage this liaison phenomenon as:

\[
\begin{array}{|c|c|c|}
\hline
\text{Sdg} & \text{Correct} & \text{Error} \\
\hline
z & A & z \\
\hline
\end{array}
\]

Table 2. A liaison rewriting rule.
Another characteristic of the French language, which may render difficult the spelling learning, is the consonant doubling which doesn’t follow any strict rule, but only a certain regularity into a word family. Those doublings can only be present at the beginning or the end of a word when the word has a foreign origin such as djinn.

So, all the difficulties met by learners must be instantiated to achieve a complete modelization of the French language. From a typology of the spelling errors covering all the French language learning we have generated the associated rewriting rules.

For a complete description of the sdg and the associated rewriting rules, please refer to [4].

5. PRESENTATION OF THE DICTATION DIAGNOSIS

After modeling the French language to diagnose exactly the learner’s input, the mediated dictation exercise has to present the results to the learner with clues allowing him to better his spelling by acquiring the lexical terms and the grammar rules, for example. Those results, which can be presented on different forms, need to fulfill the pedagogical goals. As shown in figure 3, we model this phase into two stages: the indication and the explanation of the errors. As we want the presentation to meet the pedagogical methods and to be learner’s adapted, the respective models appear in this stage.

![Figure 3. Knowledge related to the diagnosis presentation.](image)

6. DICTOR

On these bases (modelization of sdg and associated explanation rules and communication between knowledge), we’ve implemented an interactive dictation system, called DICTOR. The dictation is read out to the learner through a speech synthesis and the system corrects the learner’s input. The diagnose and the presentation of its results are implemented too and offer the learner explanations on two different levels: the first only shows the erroneous word with a short description of the kind of mistake (spelling, typographic) and the second, asked by the learner if misunderstanding remains, proposes a more complete explanation on the spelling, lexical or grammatical rules applying to the term in that context.

7. OBSERVATION IN A SCHOOL ENVIRONMENT

To evaluate the efficiency of the sdg modelization of the French language, we confronted DICTOR to the end users: the 24 pupils of a 5th level classroom. Three dictation texts were selected by the teacher to have input duration limited to twenty minutes. The texts comprised an average of 3 sentences and 23 words. There were for the three selected dictations: 91 (53), 116 (57), 113 (67) sdg (different sdg) respectively; the associated rewriting rules had each an average of 4 erroneous rewritings.

During this observation, there were a few new words for the pupils (aristoloches (plant), Nabounassar (a proper name), fourche (fork), entrelacés (interlaced)), difficult terms (s’assoupir (to doze off), lorsqu’il (when he), souverain (sovereign), majestueux (majestic), soigneusement (carefully)) and an unusual phrase (il eut tôt fait d’arriver (he soon arrived)). Those items caused many errors for the learners but there were no problems linked to the French language modelization we’ve chosen. The errors were always well detected and the explanations adapted.

We asked the children and the teachers of the school for the opinion on the system in an open questionnaire. The system had a very good acceptance by the potential users’ class that are the children. Sixty-six per cent of them want to use the system at school or at home. Even if the computer attracts them, the learners find that DICTOR is an interesting way to do an exercise that 80 % of them don’t like. The teachers could see that computers are able to dispense an individualized teaching. The system was considered an open tool, interesting and easy to use by the children, usable during the language course or in a individual training.

8. CONCLUSION

We have presented a model of the French language devoted to the dictation exercise. Its pertinence has been put into relevance by the observation made in real conditions in a school environment. However, we must validate it on a larger period, with a larger population at different language learning levels. The genericity of the model allows us to extend it to other languages and learners: French Second Language for Arabic learners in collaboration with the Institut National d’Informatique d’Algiers (current AUPELF project), English Second Language with J. Malet (Sacramento CA).

NOTES

1. “_”: word separation character; ant]: the letters ant end a word; sPL: plural morphogram s; “Ã” : the void character.

REFERENCES