

# Generating proper name pronunciation variants for automatic speech recognition

**Katarina Bartkova**

France Telecom R&D

DIH/IPS, 2 avenue Pierre Marzin

22300 Lannion, France

E-mail: katarina.bartkova@rd.francetelecom.com

## ABSTRACT

Generating correct pronunciation of proper names remains one of the most difficult tasks in text-to-phoneme transcription. Although phonetic rules can be efficient in processing proper names of one language, foreign family names cannot be always correctly generated without additional pronunciation rules.

The present study addresses the problem of pronunciation variants for French and foreign family names in a French grapheme to phoneme translator. 12 language and language group filters process the names and when necessary additional, language dependent pronunciation variants are added to the basic pronunciation yielded by the French transcription rules. In speech recognition based interactive vocal services, adding pronunciation variants to the basic phonetic transcription reduces the substitution error rate on proper names and makes applications using proper names more user friendly.

## 1. INTRODUCTION

Pronunciation rules of a language cannot account for all proper name pronunciations. The reason of this is that proper name etymology can be very diverse. Even names of one specific language can hold traces of their historical evolution having thus a pronunciation not covered by modern language pronunciation rules. In languages such as French or English, where the spelling is far from phonetic, a same proper name can have several phonetic realizations. Names in a country witness emigration waves of different origins. The mixing of the population makes the processing of proper name multilingual and often only language dependent pronunciation rules can predict the correct pronunciation [1].

Proper name processing became crucial with automated services such as, reverse directory assistance, voice-mail, and directory assistance. The problem of proper name pronunciation is first addressed in speech synthesis [2,3]. Although there one correct pronunciation prediction is sufficient, it is to be decided which among all the possibilities is the right one. This task is not always easy for French proper names and is even less easy if proper names are of foreign origin.

## 2. PROPER NAMES IN SPEECH RECOGNITION

Many proper names in French can have more than one correct pronunciation form. The problem faced in speech recognition is that allowing only one pronunciation form for a proper name, even if it is a correct one, can turn out as too rigid, for not all the speakers would utter the name in the same phonetic form. Even phonological forms of native French proper names are not always correctly generated by general pronunciation rules. The problem is even more acute for proper names of foreign origin and when the origin of the name is detected it can help to determine correct phonetic forms [1]. A foreign origin proper name pronunciation can oscillate between a correct pronunciation according to its origin and a pronunciation applying French rules exclusively. French-rule generated pronunciation cannot be simply replaced by pronunciation rules from the language of origin, for even if the name is genuinely foreign, it still can be uttered by speakers not acquainted with those pronunciation rules. In fact, speakers, uttering a proper name of foreign origin can apply only French pronunciation rules but can also apply some or all the pronunciation rules of the origin language. Moreover, the spelling of a foreign name does not always testify the degree of its phonological assimilation. A person living in France for generations would utter his family name according to French pronunciation rules while somebody with the same family name but just arriving there or living there for a short period of time, would maintain its original pronunciation. Therefore the goal of a text-to-phoneme translation in speech recognition is not to yield one single correct pronunciation but to predict all the possible pronunciation variants in order to be able to recognize the proper name even if its pronunciation deviates from the phonetic form considered as the most correct.

In order to address the problem of proper name pronunciation, this study adopts the option to detect the language of origin of the name and then to add language dependent pronunciation variants to the phonological form obtained by French standard pronunciation rules.

### 3. TEXT-TO-PHONEME CONVERSION

To avoid problems encountered in proper names rule generated pronunciations, it can be supposed that the best way to deal with proper names is the use of a special lexicon. However, the proper name class is an open one, therefore it is hardly realistic to create an exhaustive lexicon. Moreover, proper name spelling can contain spelling variants (double consonants, diacritic signs...), which would augment even more the number of lexicon entries. For these reasons pronunciation rule processed phonetic forms are to be considered as an acceptable alternative to yield a rough phonetic form to be adjusted, if necessary, by complementary proper name specific rules.

The French text-to-phoneme translator used in this study works with a lexicon of about 30,000 proper names. When a name is not found in the lexicon, it is processed with rules. The rules together with the lexicon are very robust in phonetic form generations of French origin names and they cover the most widely spread foreign origin names as well.

### 4. DATA BASE

The French directory database used in this study contains 1.3 million different proper names covering 25 million telephone subscribers. As it was found for American English [4], in French the proper name distribution follows roughly a Zipf's law: that is a few names account for a large number of subscribers whereas a large number of names remain rare. For example, the 216 most frequent first names account for 80% of the distribution of all first names (out of a total of 204226 first names). The same tendency is followed by surnames. As it is shown on Figure 1, from 1,102,089 different surnames, 581 surnames account for 20% of the distribution, 8,255 for 50% and 66,562 for 80%. The most common surname name in France is "Martin" and the most frequent first name "Jean".

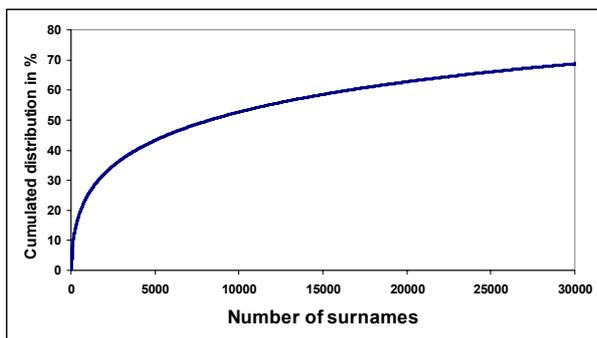


Figure 1: Surname cumulated distribution according to their frequency.

### 5. PROPER NAMES PROCESSING

When a proper name is processed by pronunciation rules (i.e. is not found in the lexicon), language filters are activated to detect its origin. According to the detected language, supplementary pronunciation variants, if

necessary, are added to the initial rule supplied phonetic form. The language filters use morphemes or grapheme sequences to detect 12 languages or language groups. In many cases no pronunciation variant is added to the basic phonetic form. For example, though "Bartkova" is detected as of Slavic origin and "Akimoto" of Japanese origin, their spelling is phonetic and therefore their initial rule supplied phonetic form is correct.

When processing the proper name database, 16% of the phonetic outputs is yielded by lexicon, 54% is generated by pronunciation rules with no other pronunciation variants than French mute [ə], 18% of phonetic forms contain French related pronunciation variants other than [ə] and 12% of the names contains foreign language related pronunciation variants (see Figure 2).

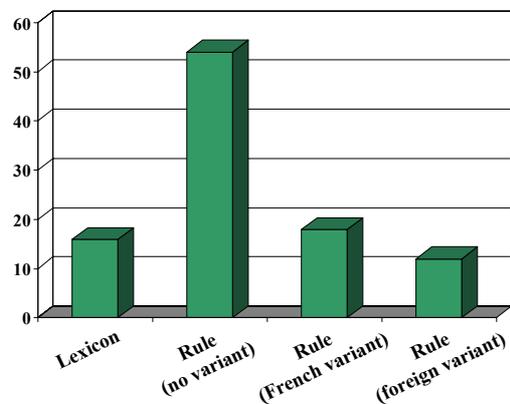


Figure 2: Proper name processing by grapheme-to-phoneme software.

### 6. LANGUAGE FILTERING

In order to generate pronunciation variants, language detection rules are introduced into the text-to-phoneme software. Distinction is made among 12 languages or language groups and phonological rules are then used to generate language dependent pronunciation variants. The languages, detected in the software are the following: French, English, Dutch, Arabic, Italian, Polish, German, Japanese and Hungarian. Three language groups are currently detected: Slavic-group containing Slovak, Czech, Serbian, Croatian Russian and Bulgarian languages; Latin-group containing Spanish and Portuguese languages; Asian group containing Chinese, Vietnamese and Korean languages. Besides these language filters, a no-language specific filter detects an unknown foreign origin of a proper name.

The problem of dealing with foreign origin proper names is highly language dependent. When a foreign name comes from a language that does not use Latin alphabet, then it is virtually phonetically transcribed in French. That is the case of most Russian origin proper names (ex: "Kougoucheff" for [kuguʃef]). Difficulties arise when the

name belongs to a language having the same alphabet system as the one providing the pronunciation rules (ex: "Aleksiejczuk" for [aleksiejtʃuk] while French rule generated transcription gives [aleksjezkyk]). In this case it is often necessary to introduce special language dependent pronunciation rules in order to cope with pronunciation variants.

In order to deal with pronunciation variants in this study, when it is necessary, the basic French rule based pronunciation is completed with variants. This way a speaker, not familiar with a name or of its origin, can utter it simply applying standard pronunciation rules [5]. However, when a speaker is familiar with a proper name or with its origin, his pronunciation can be altered from the basic form. It would be very clumsy and unacceptable to penalize such a speaker. Moreover, when a name is of a foreign origin and when a speaker is not familiar with it, he can only partly apply pronunciation rules. Local variants can account much better for these pronunciation possibilities than two complete phonological forms: one without variant and one corresponding to the language dependent variants. This way, instead of having two pronunciation variants for the proper name of Hungarian origin "Zsoldos" as (zɔldo+zɔldo) where the first one corresponds to the French rule generated pronunciation and the second one to the language dependent one, the pronunciation variants included those generated by language dependent rules are represented in the following figure. As it can be seen in Figure 3, common phoneme sequences are shared between French and native pronunciations allowing for a mix of French-based and native-based local pronunciation variants.

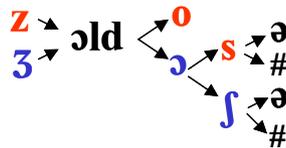


Figure 3: Combined French and language dependent variants (# = no phoneme).

Although not all pronunciation variant rules are yet introduced into the text-to-phoneme translator software (the work is still to be continued), the efficiency of the rules already active can be evaluated. Figure 4 represents the percentage of activation of foreign language filters. Italian and Polish proper names are the most frequently detected. Most of the Chinese and Vietnamese proper names are introduced into the proper name lexicon as they are mostly monosyllabic and thus not long enough to be detected by language related prefixes, suffixes or typical letter sequences. Therefore his language group automatic detection remains low. Japanese and Hungarian proper names are not very common in France compared with German or Dutch origins.

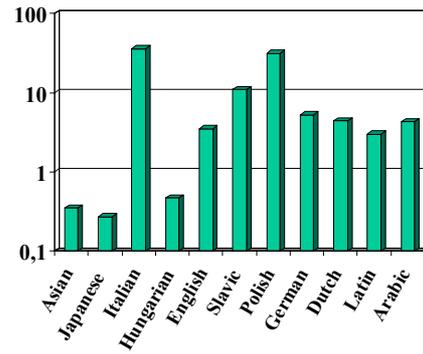


Figure 4: Percentage of activation of the different language detection filters.

## 7. FRENCH ORIGIN

A French origin proper name can also be altered from the general pronunciation rules and can have pronunciation variants. In French it is very common not to utter the last syllable of substantives, adjectives or verbs, for example "ils parlent" the suffix "-ent" of the verb is mute [il parl(ə)]. However, a proper name, though grammatically or etymologically related to common names, can have altered pronunciations. Very often in proper names the last consonant is uttered while it is mute in common names. Thus, for example, the second person of plural of the verb "bouchez" (to fill up) has its "z" mute while when the same word occurs as family name, the final "z" is to be pronounced.

In French, proper names variants are due mainly to the occurrence of mute [ə] and to the addition of the final consonant, which remains mute when general pronunciation rules are applied. This way, for example, the name "Chancenot" has its basic transcription [ʃãs(ə)no]. The special proper name processing adds to the canonical transcription the variant, which allows uttering (if needed) the last consonant (see Figure 5). Naturally not all the endings generate pronunciation variants. When the ending of a name is a very common word, such as for example, "bois" (wood) in "Dubois", no variant is created. No variant is created either when the name suffix is such that the last consonant is never uttered in French (like "x" in "-aux", "-eux").

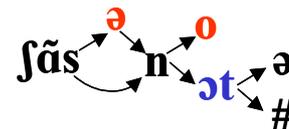


Figure 5: Pronunciation variants of the French origin proper name "Chancenot"

## 8. FOREIGN ORIGIN

In order to detect foreign origin of a proper name, prefixes, suffixes and phoneme sequences are used, when these are

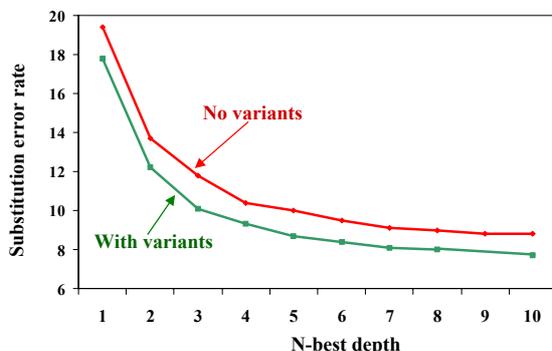
different of those found in proper names of French origin and are typical of a specific language. This way, for example, suffixes such as "-uccio", "-cchio", "-cchi", "-ggio" are considered as not-French ones and are identified as being of Italian origin. When they are detected, the correct Italian pronunciation is added to the rule generated French pronunciation, allowing combination of a French like and an Italian like pronunciation.

Some phoneme sequences are very powerful to detect a specific language. This way, for example, the sequence "-abd-" accounts for a great number of Arabic proper names. The suffix "-wski" and "-wska" are very efficient detectors of Polish proper names (of males and females). Sometimes conflict can arise between two languages as how a phoneme sequence is to be uttered. There is such a conflict between Hungarian and Polish languages in the pronunciation of the sequence "-sz-". In fact, in Hungarian graphemes "-sz-" is uttered as [s] (while grapheme "s" is uttered as [ʃ]) whereas in Polish quite the opposite is happening (the sequence "sz" is uttered as [ʃ] while the grapheme "s" is uttered as [s]). In such case, the transformation towards one or the other phoneme depends of further analysis of the grapheme sequence.

The language filter is far from being exhaustive and for example many African languages are not detected. However what is detected is the foreign origin of a name. In this case, when no specific language is identified, some general French-to-foreign pronunciation variant rules are applied. The major differences between a French pronunciation and a foreign one affect the following sounds: in French some vowels followed by nasal consonants plus an extra consonant give nasal vowels, the grapheme "u" is uttered as [+front, +high, +round], grapheme "e" is often transformed to an [ə], the grapheme "h" is always mute in French... When a word is detected as foreign, pronunciation variation creation affects these typical French grapheme-to-phoneme transformations.

## 9. PERFORMANCE EVALUATION

Pronunciation variants of proper names are currently used in an experiment of automatic directory assistance.



**Figure 6:** Evaluation test comparing proper name recognition with and without pronunciation variants

Figure 6 shows the recognition error rate figures obtained on proper names without pronunciation variants (upper line) compared with error rates obtained on proper names with pronunciation variants (lower line) when the recogniser produces from one to 10 N-best hypotheses. Using pronunciation variants improved recognition rate of the proper names as compared to using a single pronunciation basic form and thus made the service more user friendly. It appears that introducing pronunciation variants yields a reduced substitution error rate, no matter how deep the search of N-best solutions is carried out.

## 10. CONCLUSION

The present study deals with the problem of pronunciation transcription of proper names in French. It is shown here that in proper name speech recognition, where several pronunciation possibilities exist for a large number of the vocabulary words, using only a single baseform pronunciation is not satisfactory. Moreover, when a proper name is of foreign origin, detecting the language of origin is very helpful to add language dependent pronunciation variants. However, when the name is detected as of non-French origin and no specific language is identified, general French-to-foreign pronunciation variants are added. Experiments in automatic directory assistance showed that introducing pronunciation variants into the phonetic transcription of the names improves not only speech recognition performance but also the user's satisfaction.

## ACKNOWLEDGEMENTS

I would like to thank Laurence Ferrieux and Alexandre Ferrieux for providing statistics on proper name occurrence frequency and the recognition evaluations results.

## REFERENCES

- [1] A. F. Llitjos and A.W. Black, "Knowledge of language origin improves pronunciation accuracy of proper names", Proc. Eurospeech-2001, Aalborg, Denmark, 2001
- [2] P. Bagshaw, "Phonemic transcription by analogy in text-to-speech synthesis: Novel word pronunciation and lexical compression", Computer Speech and Language, 12:119-142, 1998.
- [3] F. Yvon, "Self-learning techniques for grapheme-to-phoneme conversion", Proc. 2nd Onomastica Research Colloquium, London, 1994.
- [4] A.R. Golding and P.S. Rosenblum, "A comparison of Anapron with seven other name-pronunciation systems", Journal of the American Voice Input/Output Society, 14:1-21, 1993
- [5] S. Fitt, "Spelling unfamiliar names", Proceedings: XIXth International Congress of Onomastic Sciences, Aberdeen 1996, Vol. 1, pp. 110-6