

LE DÉBIT DE PAROLE : UN FILTRE UTILISÉ POUR LA
GÉNÉRATION DES VARIANTES DE PRONONCIATION EN
FRANÇAIS PARISIEN

Anne Lacheret-Dujour

LIMSI-CNRS BP 133 91 403 Orsay-Cedex France

ABSTRACT

This paper describes the development of a grapheme-to-several phoneme strings module according to speech rate in French. Some examples of the phonological variations linked to the speech rate and the basic principles of the system are presented.

1. INTRODUCTION

Puisqu'il n'existe pas de prononciation standard en reconnaissance de la parole multilocuteur, l'intégration de modules de génération automatique des variantes de prononciation dans des systèmes de reconnaissance phonétique est nécessaire pour l'accès au lexique. De tels systèmes ont été réalisés au LIMSI: GRAPHER [5] et VARION.0 [1]. Les tests de VARION.0 ont révélé la complexité du problème lié à la génération automatique des variantes de prononciation: du fait de la production équivalente et maximaliste des variantes, l'explosion combinatoire des chemins allophoniques produits est inévitable. Il est donc irréaliste de penser pouvoir utiliser de façon optimale ces systèmes dans des complexes de reconnaissance; l'utilisation d'heuristiques, qui permettent le cas échéant de bloquer la génération de certaines variantes, est indispensable.

Nous avons développé pour le français parisien un module de transformation graphème-phonème avec variantes,

VARION.1 [3], dans lequel la génération des allophones est conditionnée par le débit de parole (lent ou rapide). Les règles du module ont été développées à partir de l'observation de corpus de parole continue en situation de lecture, prononcés par quatre locuteurs à différents débits (lent, normal et rapide) [2]. L'objet de cet article est de présenter deux exemples de variantes liées au débit prises en compte par le système: la prononciation variable du schwa et la fusion vocalique. L'architecture générale du module de règles (logiciel utilisé et formalisme adopté) est également décrite.

2. LE SYSTÈME DE RÈGLES

2.1 Génération des variantes en fonction du débit de parole: présentation des règles.

* Elision facultative du schwa

Indépendamment du débit, l'élosion en début de mot n'est envisageable que si le schwa est précédé d'une seule consonne dont le contexte gauche est autre que le graphème 'e' en finale de lexème (*la dame demande* → /ladamdəmād/ et non /ladamdād/).

Quand plusieurs 'e' se suivent les séquences graphémiques suivantes sont à distinguer:

- 'monosyllabe#ne' (*mais je ne sais pas*) ou 'ne#monosyllabe' (*ce pantalon ne te va pas*). Dans ces contextes, seul le 'e' de la négation peut être éliidé en débit lent. En débit rapide, la négation

peut être omise totalement. Le schwa restant peut ne pas être prononcé dans la mesure où le contexte gauche le permet (*je ne sais pas* → /Ssɛpa/).

- Dans la séquence 'ce#que', seul le premier 'e' peut être éliidé quel que soit le débit (*je sais ce que tu penses*).

- Dans les autres séquences, un 'e' sur deux peut tomber à partir du premier si le contexte gauche le permet, à partir du second dans les autres cas (*mais redemande-le*).

Les règles sur la chute du schwa isolé en fonction du débit de parole varient relativement à sa position dans le mot:

- En début de polysyllabe ou dans un monosyllabe, le 'e' est toujours maintenu en débit lent, il peut être éliidé en débit rapide (*demander* → /dəmād/ en débit lent, /dəmād/ ou /dmād/ en débit rapide).

- En milieu de polysyllabe, le schwa est éliidé obligatoirement en débit rapide, facultativement en débit lent (*seulement* → /soelmā/ en débit rapide, /soelamā/ est également prévu en débit lent).

- En fin de polysyllabe, l'élosion facultative, obligatoire ou interdite des finales '-e', '-es', '-ent' dont le contexte droit est consonantique dépend du nombre de consonnes à gauche de la finale. Les règles sont les suivantes pour les contextes consonantiques gauches ci-dessous:

R1 [+1cons]: la finale est éliidée quel que soit le débit (*une robe verte* → /ynrɔbvɛrt/).

R2 [+cons,-liq][+liq]: la finale est toujours maintenue en débit lent, facultativement en débit rapide, entraînant avec elle la chute de la liquide qui la précède (*ils peuplent Paris* → /ilpœplɔpari/ en débit lent, /ilpœppari/ est une variante possible en débit rapide).

R3 [+liq][+cons,-liq]: la chute du 'e' est facultative quel que soit le débit (*une valse de Vienne*).

R4 [+2cons]: la chute de la finale est tolérée en débit rapide uniquement (*un texte de base*).

Quand la finale '-ent' est suivie d'un mot à initiale vocalique, le 'e' est tou-

jours éliidé en débit rapide, en débit lent il peut être entendu si la liaison est effectuée (*ils aiment y aller* → /ilzɛmɑ̃tjalɛ/).

Un certain nombre d'exceptions sont à noter à ces règles générales, pour lesquelles, quel que soit le contexte et le débit, le 'e' est toujours maintenu (*femelle, relier*). Il en va de même dans le déterminant 'le' accentué, dans le démonstratif 'ce' suivi d'une voyelle (*ce en quoi*). En revanche, dans les formes du futur, le graphème 'e' correspond à un phonème 0 si le contexte gauche est autre que /+obstr/[+liq] (*il aidera* → /ilɛdra/). Il en va de même pour le pronom 'je' postposé au verbe (*qui suis-je?*) ainsi que dans la tournure interrogative 'est-ce' (*qui est-ce qui vient?*).

* Fusion vocalique

Si 2 voyelles identiques sont séparées par une frontière de mots et éventuellement un 'h' aspiré, elles peuvent être réduites en un seul et même long segment. Il s'agit de *fusion vocalique*. Dans le module, la fusion est produite pour toutes les voyelles en débit rapide uniquement puisqu'aucune considération de type syntactico-sémantique n'est prise en compte pour affiner les règles. Si les graphies correspondant aux phonèmes 'o', 'e', 'ø', qui en syllabe inaccentuée non fermée par la liquide 'r' peuvent être facultativement réalisées ouvertes ou fermées, sont à l'initiale de polysyllabes éventuellement précédées d'un 'h' muet (*offert, aimable, heureux*), si elles ont comme contexte gauche respectivement les phonèmes 'o', 'e' ou 'ø', leur degré d'aperture doit être identique à celui du contexte gauche (*le corbeau officie le mardi* → * /lɔkɔrboʔfisi lɔmardi/. La fusion vocalique est alors une variante libre produite en débit rapide: /l(ɔ)kɔrbo:fisil(ø)mardi/.

2.2 Le système de règles VARION.1.

Les principaux objectifs lors du développement de VARION.1 étaient les suivants :

(1) Formaliser les variantes par le biais de règles et non en faisant usage d'un lexique. Car, la structure de la langue française ne justifie pas l'emploi d'un dictionnaire couteux en espace mémoire; les irrégularités rencontrées sont dues pour une faible part à la structure morphologique de la langue; l'accent lexical et le ton, qui jouent un rôle distinctif dans certaines langues, sont inexistantes en français. Enfin, l'utilisation d'un lexique nécessite une maintenance rigoureuse afin de traiter correctement les néologismes.

(2) Effectuer une transcription graphème-phonème sans passer par l'intermédiaire de formes de base au sens chomskyen du terme; concept qui sous-entend une notion d'écart par rapport à une norme abstraite dont la définition est loin d'être claire.

(3) Adopter un formalisme permettant une écriture simple, économique et compacte des règles, qui doivent être facilement testables et le cas échéant modifiables. Pour ce faire, nous avons utilisé le compilateur de règles LEX [4] qui permet d'effectuer n'importe quel traitement linguistique sur une chaîne de caractères donnée en entrée (Fig.1).

Les règles, dans lesquelles on peut inclure des traitements procéduraux, sont écrites sous forme déclarative et compilées en langage C. Le programme effectue ainsi un certain nombre d'actions spécifiées par l'utilisateur. Il génère ensuite un automate déterministe d'états finis. Le temps requis pour l'exécution des règles dépend de la taille du texte à phonémiser. Dans un fichier, les règles sont ordonnées, non cycliques, elles sont déclenchées de gauche à droite de la forme à transcrire, elles sont du type :

A o("B");

trée où l'on trouve le caractère 'A', appliquer la fonction o(s) pour le remplacer par le caractère 'B'. On substitue ainsi un buffer phonémique à un buffer graphémique au fur et à mesure de la transcription. 'A' peut être un mot, un graphème ou même une séquence de mots. 'B' peut correspondre à 0, 1 ou plusieurs phonèmes.

La figure 2 représente la phonémisation de la séquence graphémique "Le corbeau officie seulement le mardi" en débit lent et en débit rapide.

3. CONCLUSION

Les tests du système sur de la parole lue mettent en lumière l'amélioration des résultats (98.8% de variantes prononcées et prévues par le système en débit lent, 98.4% en débit rapide contre 95.5% pour VARION.0 tout débit confondu). Néanmoins, l'analyse de ces résultats amène les conclusions suivantes : des connaissances prosodiques supplémentaires (la distribution des pauses par exemple) sont nécessaires pour améliorer les performances d'un tel système. Il en va de même des connaissances syntaxiques. Enfin, le débit de parole est une variable relative. Les variations de débit ne sont pas toujours exécutées de façon identique d'un sujet à un autre. Chez un locuteur donné les variations sont également possibles. De ce fait, à un débit particulier, les choix allophoniques peuvent varier d'un groupe de locuteur à un autre et chez un même sujet lorsqu'il répète la même séquence de parole. Il est donc nécessaire d'étudier, outre les facteurs linguistiques, les mécanismes extralinguistiques (sociolectes, idiolectes) et para-linguistiques (situation de discours, émotivité du locuteur, etc) qui sous-tendent les stratégies allophoniques pour des classes de locuteurs données. Une telle étude permettrait de déclencher les ensembles de règles appropriés à un groupe de locuteurs spécifique.

En tout point du texte donné en en-

Légende des symboles utilisés

R: règle.

[+ 1cons]: une et une seule consonne.

[+ 2cons]: deux consonnes ou plus.

(e/es/ent/s): finales facultatives.

[-son]: consonne moins sonante.

[+ocl]: consonne occlusive.

REFERENCES

- [1] A. Dujour, Octobre 1987, "Conversion graphèmes-phonèmes avec variantes du français par règles". DEA interuniversitaire de phonétique, Université de Paris 7 (Jussieu).
 [2] A. Lacheret-Dujour, Octobre 1989, "Automatic Generation of Phonological Variations", EUROSPEECH, vol 2, pp 376-379, Paris.
 [3] A. Lacheret-Dujour, Juin 1990, "Contribution à l'analyse de la variabilité phonologique pour le traitement automatique de la parole continue multilocuteur". Thèse de Doctorat de l'Université de Paris 7 (Jussieu).

FIGURES

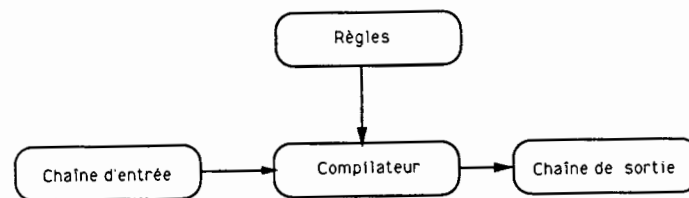


Fig.1: Le compilateur de règles, LEX.

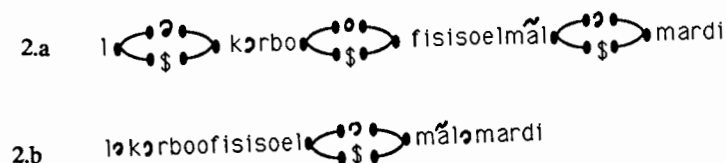


Fig.2: Phonémisation de la séquence "Le corbeau officie seulement le mardi", en débit rapide (Fig.2.a) et en débit lent (Fig.2.b)