

27 juillet 2023

Université de Fianarantsoa



Sous la co-tutelle de :
CNRS
ÉCOLE DES PONTS PARISTECH
UNIVERSITÉ GUSTAVE EIFFEL

La rédaction scientifique

Objectif de cette présentation

Aider les auteurs à écrire des articles scientifiques qui seront

- **acceptés et publiés**
- **lus**
- **convaincants**
- **cités**

Accessible, simple

Progressif

Clair

Précis

Fiable, vérifiable

Citations fidèles

Structuré



Accessible, simple

Écrire de façon accessible, simple

Faire des phrases courtes

Donner des exemples

Rappeler des définitions

Utiliser le découpage en paragraphes

Comment savoir si l'article est facile à lire ?

- se relire
- savoir par l'expérience du domaine ce qui est bien connu ou peu connu (lire des articles du domaine)

Renoncer aux contenus qui ne servent pas à atteindre les objectifs de l'article,
pour éviter de distraire le lecteur du contenu utile

Expliciter les abréviations et les sigles dès leur première occurrence

Les efforts de l'auteur pour écrire simplement épargnent aux lecteurs des efforts
pour comprendre

Le fait que l'auteur écrive de façon accessible ne veut pas dire qu'il est borné !

Faire des phrases courtes

Compte tenu du fait que la production d'un message oral nécessite des connaissances syntaxiques sur le texte, on a intérêt à effectuer la phonétisation du français par lexique plutôt que par règles.

Laporte, Eric, 1986. Applications de la morphophonologie à la production automatique de textes phonétiques. *Séminaire GRECO - GALF « Lexique et traitement automatique des langages »*, 16-17 janvier 1986, Toulouse.

Ici, un découpage en deux phrases aurait facilité la lecture

Une information à la fois

Chaque fin de phrase laisse au lecteur l'occasion de faire une pause pour réagir au contenu de la phrase

Lors de la phonétisation d'un texte écrit, les informations syntaxiques ne peuvent être trouvées qu'à l'aide d'une analyse syntaxique du texte à phonétiser, ce qui comporte en particulier, notons-le déjà, la consultation d'un lexique informatique donnant les catégories grammaticales des mots. Désormais, nous supposons que ces informations syntaxiques

Donner des exemples

En effet, pour rendre la parole synthétique plus naturelle, on est amené à traiter la prosodie : intonation (MARTIN, 1975), accent et rythme, phénomènes liés à la syntaxe du texte (DANLOS, EMERARD, SORIN, 1985).

Ici, un ou deux exemples de phrases auraient pu montrer concrètement que l'intonation, l'accent ou le rythme sont liés à la syntaxe de la phrase

Non, Éric et Alain sont repartis

Jean, Éric et Alain sont repartis

On s'aperçoit de ces défauts en **se relisant**

Si on n'a **pas assez de place**, mieux vaut écrire deux articles au lieu d'un

- un dictionnaire DELAPS de 50 000 formes de base (verbes à l'infinitif, adjectifs au masculin singulier, etc), extension du dictionnaire DELAS (dictionnaire électronique du LADL - mots simples) qui donne les catégories morpho-syntaxiques des mêmes mots et occupe 1 Mo sans aucun compactage (COURTOIS, 1985) ;

Rappeler des définitions

Nous allons examiner les conséquences de ces faits, dans le cas de la phonétisation de textes écrits, puis dans le cas de la génération de messages oraux à partir de représentations sémantiques. Nous ne parlerons pas de la prosodie et nous nous limiterons aux phénomènes segmentaux.

Dans un domaine pluridisciplinaire, même les termes de base d'un domaine peuvent être inconnus des lecteurs spécialistes d'un autre domaine
Une notion peut être peu connue des spécialistes du domaine eux-mêmes
Ici, cela aurait facilité la lecture que l'article définisse les phénomènes segmentaux
Le fait que l'auteur définisse un terme ne veut pas dire qu'il vient de le découvrir !

Utiliser le découpage en paragraphes

Pour qu'un ordinateur prononce un message, une étape nécessaire est la mise sous forme phonétique du texte, c'est-à-dire la transcription en une représentation directe de la prononciation. Cette opération nécessite des informations syntaxiques sur le texte.

En effet, pour rendre la parole synthétique plus naturelle, on est amené à traiter la prosodie : intonation (MARTIN, 1975), accent et rythme, phénomènes liés à la syntaxe du texte (DANLOS, EMERARD, SORIN, 1985).

Même idée ou même raisonnement, même paragraphe

Nouvelle idée ou nouveau raisonnement, nouveau paragraphe

Ici, la 2^e phrase aurait dû faire partie du 2^e paragraphe au lieu du 1^{er}



Progressif

Progressivité

Quand on peut choisir l'ordre des informations
placer un contenu plus simple avant un contenu plus complexe
placer le connu avant l'inconnu

En lisant d'abord une partie plus simple, le lecteur

- construit petit à petit sa compréhension de l'article
- rencontre les difficultés les unes après les autres et non en même temps
- peut avoir une idée de ce qui suit, le sauter s'il ne s'y intéresse pas, et reprendre sa lecture après

La motivation d'abord

Contexte : tout début du résumé

Compte tenu du fait que la production d'un message oral nécessite des connaissances syntaxiques sur le texte, on a intérêt à effectuer la phonétisation du français par lexique plutôt que par règles. Les phénomènes d'interaction entre mots (liaisons, élisions, ...) suggèrent l'utilisation d'une représentation intermédiaire morpho-phonologique telle que celle élaborée au LADL.

Certains lecteurs ne s'intéressent pas exactement au même sous-domaine que l'auteur
Pour s'intéresser à l'article, ils ont besoin de savoir

- de quel problème il s'agit
- pourquoi on l'étudie, à quoi peut servir la solution du problème

Ici, le début du résumé aurait dû dire que

- le problème est de transcrire du texte sous forme phonétique
- pour remplacer les sorties écrites des ordinateurs (à l'époque) par des sorties sonores

Le lecteur peut le deviner d'après l'article, mais cela complique son travail de lecture

Objectif de cette partie : permettre au lecteur de décider s'il va lire l'article

L'essentiel avant les détails

c) PHONETISATION PAR DICTIONNAIRE MORPHO-PHONOLOGIQUE

Cette représentation morpho-phonologique (LAPORTE, 1984) tient compte de la régularité et de la généralité des variations phonétiques en fonction du contexte : liaisons obligatoires et faits analogues, élision (SCHANE, 1968 ; DELL, 1973 ; NOSKE, 1982). Elle permet de prévoir les prononciations d'un même mot dans des contextes différents, pourvu que le conditionnement syntaxique des liaisons obligatoires soit représenté.

Le début d'une section doit introduire l'essentiel

Ici, le début de la section aurait dû dire qu'elle expose une nouvelle représentation du français, matérialisée par

- un dictionnaire qui donne la représentation des mots
- un algorithme qui les transcrit en phonétique

Objectif : permettre au lecteur de décider s'il va lire la suite

L'essentiel avant les détails

Titre
Sous-titres
Résumé
Introductions

INTRODUCTION

Pour qu'un ordinateur prononce un message, une étape nécessaire est la mise sous forme phonétique du texte, c'est-à-dire la transcription en une représentation directe de la prononciation. Cette opération nécessite des informations syntaxiques sur le texte.

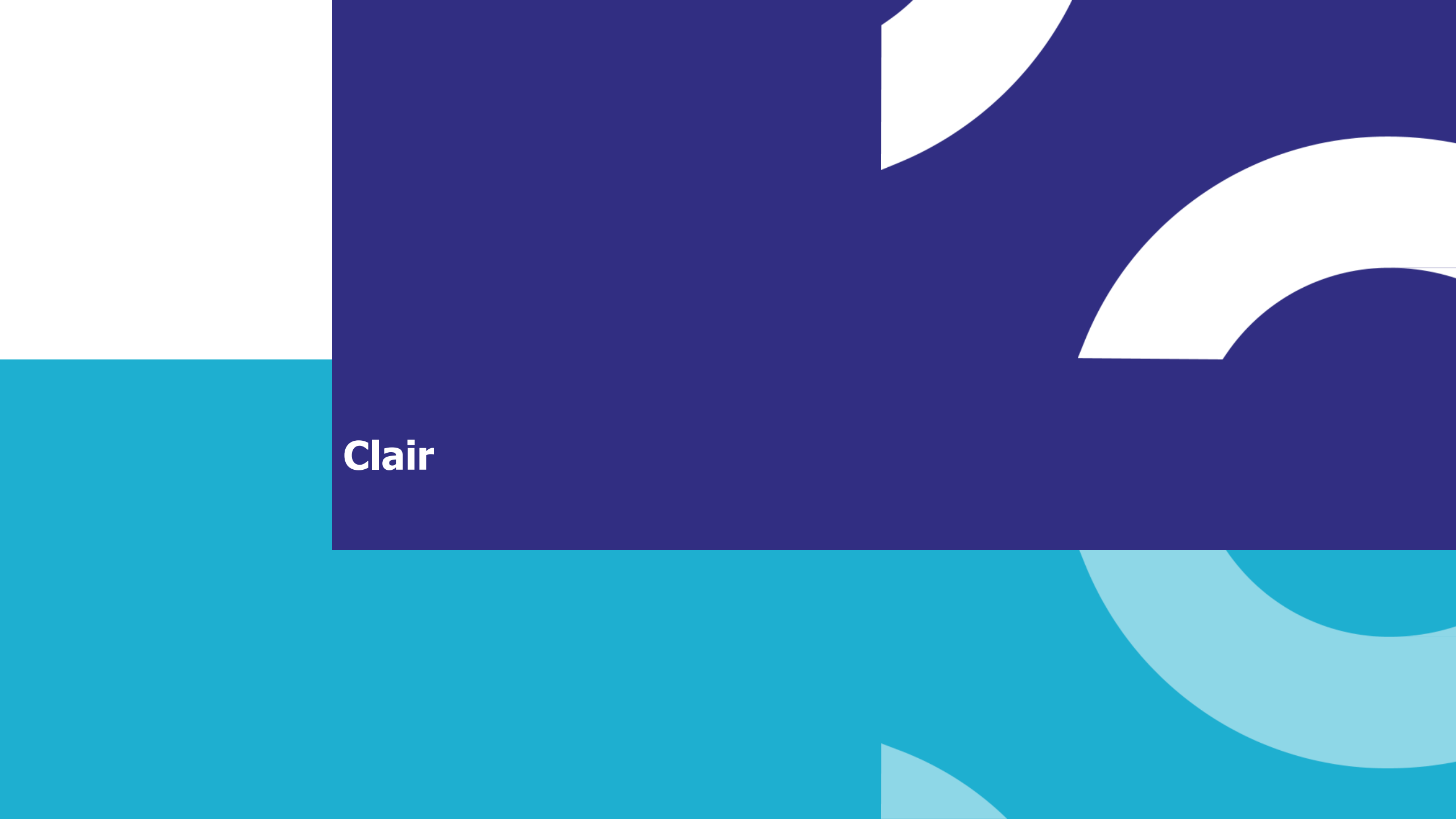
[ici trois paragraphes qui exposent une argumentation]

Nous allons examiner les conséquences de ces faits, dans le cas de la phonétisation de textes écrits, puis dans le cas de la génération de messages oraux à partir de représentations sémantiques. Nous ne parlerons pas de la prosodie et nous nous limiterons aux phénomènes segmentaux.

L'introduction et le résumé doivent donner l'essentiel sans les détails pour permettre au lecteur de décider s'il va lire la suite

Ici, il manque l'essentiel et il y a trop de détails

Le titre et les sous-titres doivent décrire le contenu correspondant



Clair

Écrire clairement

Ce qui n'est pas clair n'est pas convaincant

Éviter les ambiguïtés

a) PHONETISATION PAR REGLES

Ce procédé (DIVAY, GUYOMARD, 1977) modifie progressivement le texte. Une des transformations, par exemple, consiste à remplacer tous les *s* entre voyelles (ex. *rasoir*) par des *z*. Ces règles sont plus compliquées que, par exemple, pour l'espagnol, où 50 règles suffisent à phonétiser des textes avec un taux d'erreur virtuellement nul (SANTOS, NOMBELA, 1982). En général, en français, un son ne correspond pas à un seul graphème, et l'inverse est également vrai.

Ici :

Interprétation 1 - une séquence de plusieurs lettres : *ain* pour [ɛ]

Interprétation 2 - un choix entre plusieurs graphèmes : *o, ô, au...*

Étapes d'un raisonnement

Le passage d'une étape à la suivante doit être compréhensible

On peut avoir besoin d'une expression charnière comme *donc*, *en effet*, *de plus* ou *cependant*, mais on peut aussi ne pas en avoir besoin



Précis

Terminologie

des mots. Désormais, nous supposons que ces informations syntaxiques sont présentes, et en particulier que le conditionnement syntaxique des liaisons obligatoires est indiqué par un **marqueur** inséré entre les deux mots concernés.

Symbole indiquant une position dans laquelle la syntaxe remplit les conditions pour qu'une liaison soit obligatoire

mon [-] ami

le patron arrive

Le terme de **symbole** aurait été plus adapté, plus précis

La bonne terminologie est celle qui est

- précise, non ambiguë
- connue dans le domaine



Fiable, vérifiable

Citer les sources

tique de chaque mot est donc insuffisant. L'élision est un autre cas de variation phonétique en fonction du contexte : dans *la fenêtre*, le *e* souligné peut tomber, mais dans *une fenêtre* il se prononce obligatoirement.

Ce n'est pas un résultat original mais une application de la « loi des trois consonnes » (Grammont, 1894)

Résultats originaux

De plus, dans le cas du français (CATACH, 1984), on a également le problème des homographes hétérophones (mots s'écrivant de la même façon mais se prononçant différemment, comme *portions*). Il suffit de connaître la catégorie grammaticale d'un homographe pour résoudre cette ambiguïté.

Donner des détails factuels qui ont mené au résultat

Ici, l'article aurait dû préciser quel dictionnaire a été utilisé pour vérifier l'affirmation

En fait, il existe quelques contre-exemples :

- *pub*
[pyb] *Il y a trop de pub à la télé*
[pœb] *Que consommer dans un pub irlandais ?*
- *punch*
[pœnf] *On s'est réveillés avec plus de punch et de motivation*
[pɔf] *Remuez bien votre punch à l'aide de la louche*



Citations fidèles

Rapporter fidèlement les contributions des auteurs cités

On peut adapter, simplifier et traduire si on rapporte le contenu fidèlement

Si on cite entre guillemets, signaler toute reformulation

- **passage non cité : (...)**
- **passage modifié ou inséré, entre crochets : [(ici, reformulation par l'auteur de l'article)]**



Structuré

Plan d'un article scientifique

Titre

Résumé

Introduction

Résultats antérieurs

(Parties contenant les résultats originaux)

Conclusion

Références bibliographiques

Certaines conventions peuvent varier suivant les domaines
surtout dans les parties contenant les résultats originaux

Exemple :

- Méthode, modèle formel, hypothèses...
- Expérimentation
- Résultats
- Discussion sur les résultats

Titre
Résumé
Introduction
Résultats antérieurs
(Parties contenant les résultats originaux)
Conclusion
Références bibliographiques

Ici,

- pas de partie séparée sur les résultats antérieurs
- pas de conclusion

Applications de la morpho-phonologie à la production de textes phonétiques

[ici le résumé]

INTRODUCTION

I. PHONETISATION D'UN TEXTE ECRIT

- a) PHONETISATION PAR REGLES
- b) PHONETISATION PAR DICTIONNAIRE PHONETIQUE
- c) PHONETISATION PAR DICTIONNAIRE MORPHO-PHONOLOGIQUE

II. GENERATION AUTOMATIQUE DE TEXTES PHONETIQUES A PARTIR DE REPRESENTATIONS SEMANTIQUES

ANNEXE

BIBLIOGRAPHIE

Introduction

Contexte : toute fin de l'introduction

Nous allons examiner les conséquences de ces faits, dans le cas de la phonétisation de textes écrits, puis dans le cas de la génération de messages oraux à partir de représentations sémantiques. Nous ne parlerons pas de la prosodie et nous nous limiterons aux phénomènes segmentaux.

Il est commode pour le lecteur que la fin de l'introduction annonce le plan de l'article

Ici,

- le plan n'est pas précis
- la dernière phrase ne concerne pas le plan

Résultats antérieurs

Cette partie présente le domaine avant d'exposer les résultats originaux

Ici, les résultats antérieurs ne sont pas dans une partie séparée
Ils sont mélangés à la discussion

Comment opérer la phonétisation à partir du texte ainsi traité ?
Deux solutions se présentent :

a) PHONETISATION PAR REGLES

Ce procédé (DIVAY, GUYOMARD, 1977) modifie progressivement le texte. Une des transformations, par exemple, consiste à remplacer tous les *s* entre voyelles (ex. *rasoir*) par des *z*. Ces règles sont plus compliquées que, par exemple, pour l'espagnol, où 50 règles suffisent à phonétiser des textes avec un taux d'erreur virtuellement nul (SANTOS, NOMBELA, 1982). En général, en français, un son ne correspond pas à un seul graphème, et l'inverse est également vrai.

Toutefois, les difficultés viennent essentiellement du fait que certaines parties de mots se prononcent différemment suivant les mots, comme *-ent* (*présent* / *suivent*), *-tie* (*inertie* / *avertie*), *-tions* (*citons* / *transitions*), *s* entre voyelles (*rasoir* / *parasol*), ... Dans le cadre de la phonétisation par règles, ces ambiguïtés doivent être détectées, puis résolues. Pour cela on a besoin de règles mettant en jeu quelques caractères du contexte (CATACH, 1984) :

"si *s* entre voyelles est précédé de *par* il reste *s*"

(ex. *parasymphatique*, etc). Certains de ces cas sont encore compliqués par l'homographie (*portions*) Il faut des règles pour détecter ces cas particuliers, pour accéder aux informations syntaxiques qui les concernent, et pour utiliser ces informations.

Résultats antérieurs

b) PHONETISATION PAR DICTIONNAIRE PHONETIQUE

L'autre solution est d'accéder à un lexique électronique donnant directement, à partir d'un mot et de sa catégorie grammaticale, sa prononciation (si le mot a plusieurs prononciations, comme *portions*, chacune d'elles est associée à une catégorie grammaticale). Le problème des parties de mots à prononciation variable (comme *s* dans *parasite* et *parasol*) disparaît dès lors. Comme l'analyse syntaxique préalable suppose de toute

Citer les références pertinentes
Ici, référence manquante
Coker, Umeda & Browman (1973)

Résultats antérieurs

II. GENERATION AUTOMATIQUE DE TEXTES PHONETIQUES A PARTIR DE REPRESENTATIONS SEMANTIQUES

La génération automatique de messages oraux à partir de représentations sémantiques repose sur l'utilisation d'un générateur de textes tel que celui présenté par (DANLOS, 1985). La différence avec la génération

Ici, référence manquante
Danlos, Émerard & Sorin, 1985

Conclusion

Contexte : tout dernier paragraphe de l'article

La solution de la fig. 3, caractérisée par la flexion en représentation morpho-phonologique, nous semble la meilleure. En effet, cette représentation optimise la flexion des mots. La mise au féminin des adjectifs cités ci-dessus est uniforme et se fait par l'ajout d'un symbole final, toujours le même ; le radical de chacun des verbes cités est invariable au cours de la conjugaison. D'ailleurs, en pratique, la seule façon de fléchir des mots phonétiques serait de simuler le passage en un code phonologique, la flexion dans ce code et la phonétisation.

La conclusion est l'occasion de résumer l'essentiel de l'article et d'aborder les conséquences ou les perspectives

Ici, les derniers mots concernent spécifiquement une partie de l'article

Références bibliographiques

Les œuvres citées dans le texte doivent être listées dans la bibliographie
Les œuvres listées dans la bibliographie doivent être citées dans le texte

Conclusions

Les résultats d'un chercheur ne sont utiles que s'ils sont lus et cités
Le confort du lecteur et le pouvoir de conviction ont la priorité sur le
plaisir de l'auteur
Savoir écrire, c'est savoir se relire

Eric Laporte

eric.laporte@univ-eiffel.fr

+33 1 60 95 75 52

