



Laboratoire d'Analyse Informatique de la Parole

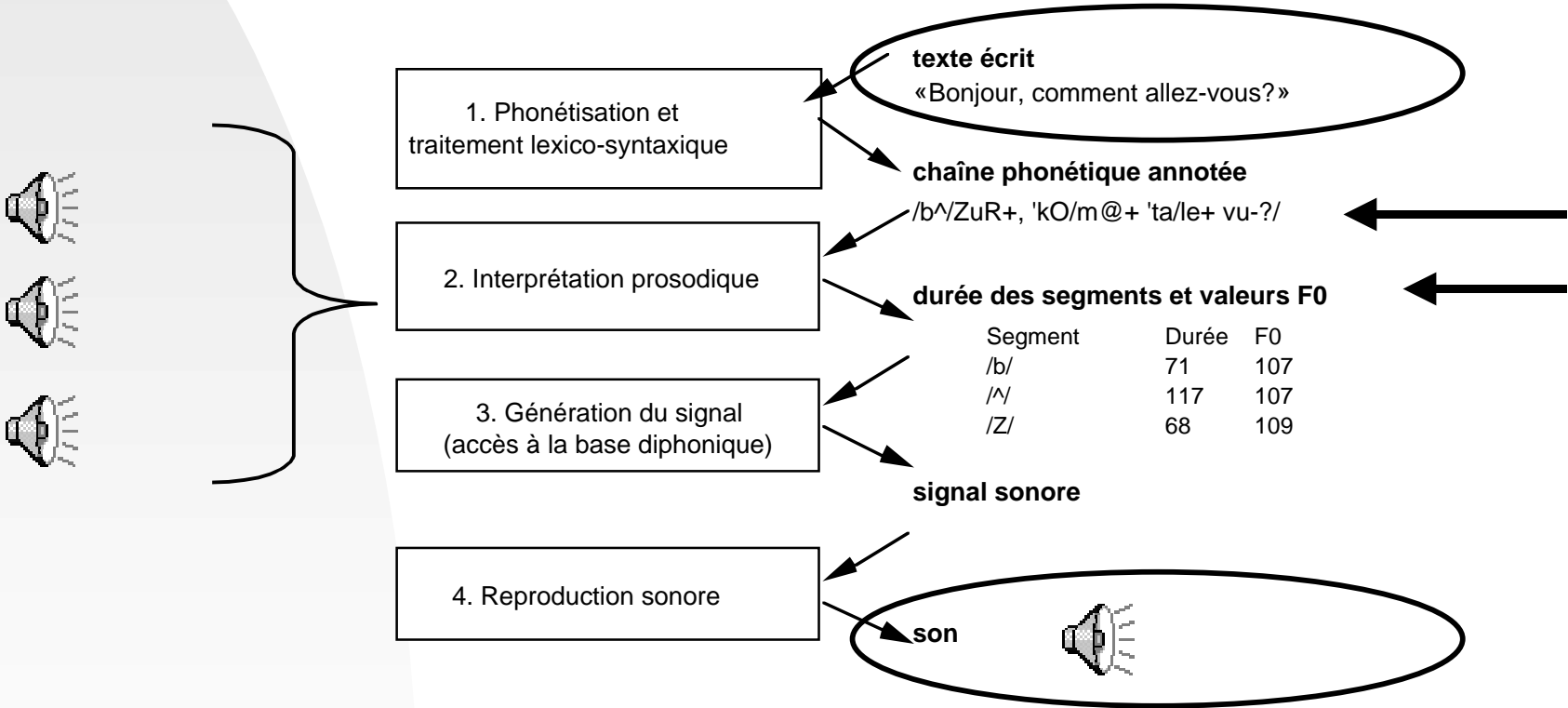
**Simulateur de parole et recherche fondamentale
Exemple en prosodie**

**Brigitte Zellner Keller
LAIP, IMM, Lettres
Université de Lausanne**



Laboratoire d'Analyse Informatique de la Parole

Architecture typique d'un système TTS



http://www.unil.ch/imm/docs/LAIP/LAIPTTS_fr.htm



Laboratoire d'Analyse Informatique de la Parole

SIMULATION
de la structuration temporelle d'un énoncé



1. Groupes de mots

2. Durées

1. Groupes de mots

**Question fondamentale:
Comment des locuteurs organisent leur
communication orale sur l'axe du temps?**

Des paquets d'information organisés



1. Groupes de mots

- **Un simulateur de parole doit imiter la façon spontanée de grouper les mots dans un énoncé.**
- **Les groupes de mots représentent les unités de cohérence sur laquelle se construit l'organisation temporelle de la parole** (rencontre des plans sémantique, lexique, syntaxe, mémoire, etc).

1. Groupes de mots

Comment prédire ces groupes de mots?

- **La structure temporelle de la parole n'est pas toujours congruente avec la syntaxe. Une prédiction basée sur des règles phonosyntaxiques peut donc s'avérer lacunaire .**

Exemple: certains énoncés oraux qui ont du sens posent problème du point de vue des règles de syntaxe (écrite).

=> TEST pour un simulateur de parole

Un bon modèle devrait pouvoir prédire les groupes de mots pour tous types d'énoncés, y compris les énoncés oraux qui ne respectent pas la syntaxe de l'écrit...



1. Groupes de mots

Tester cet énoncé sans ponctuation dans un synthétiseur
(puisqu'il est produit spontanément par un locuteur)

**“Le problème nous ce qui nous fait peur c’est
qu’elle se retrouve dans une structure inadaptée”.**

**1ère ligne: une
structure orale très
fréquente**

ICP



Syntaix



LAIPTTS





1. Groupes de mots

Comment prédire ces blocs d'information?

Hy: Le contrôle du flux d'information est relié à la dimension temporelle (dimension du SN).

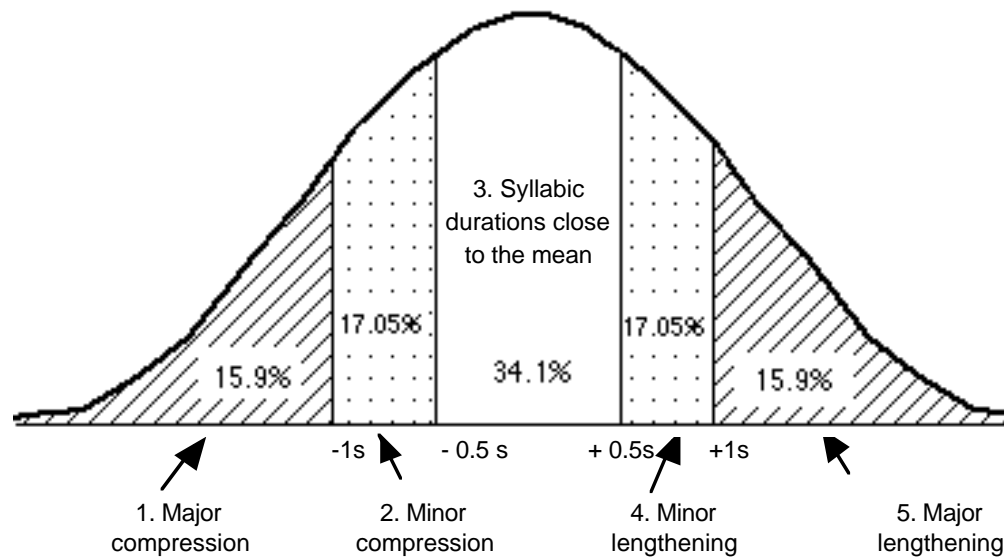
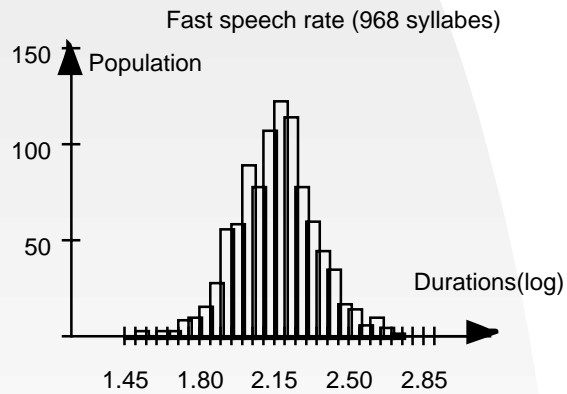
Pour le montrer, études empiriques appuyées sur deux principes de base:

- 1. Etude purement statistique de toutes les durées syllabiques
(pas d'a-priori, ni réduction linguistique)**
- 2. Les durées des unités reflètent la structure temporelle de l'énoncé parlé
=> algorithme qui restitue cette relation**



Laboratoire d'Analyse Informatique de la Parole

1. Groupes de mots

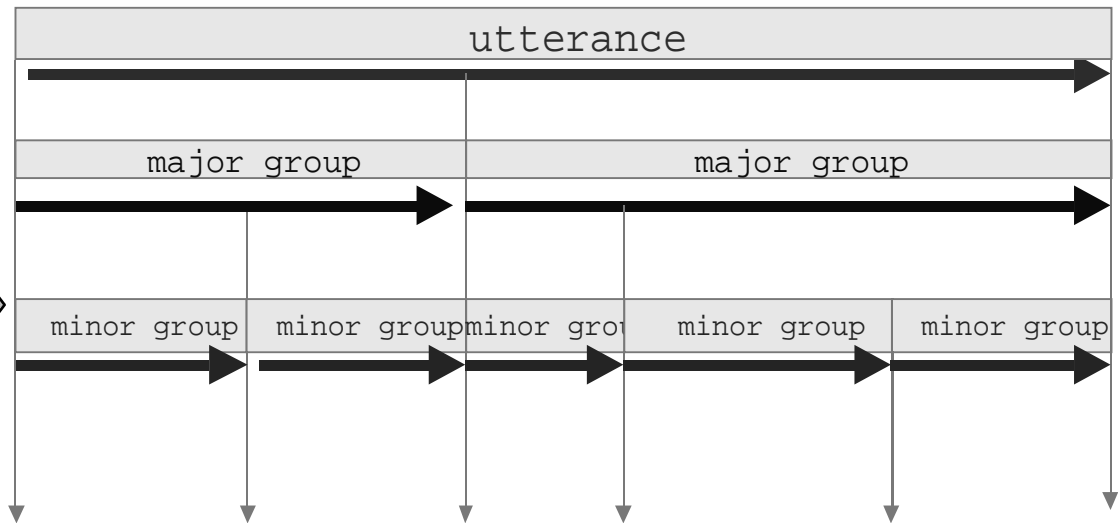
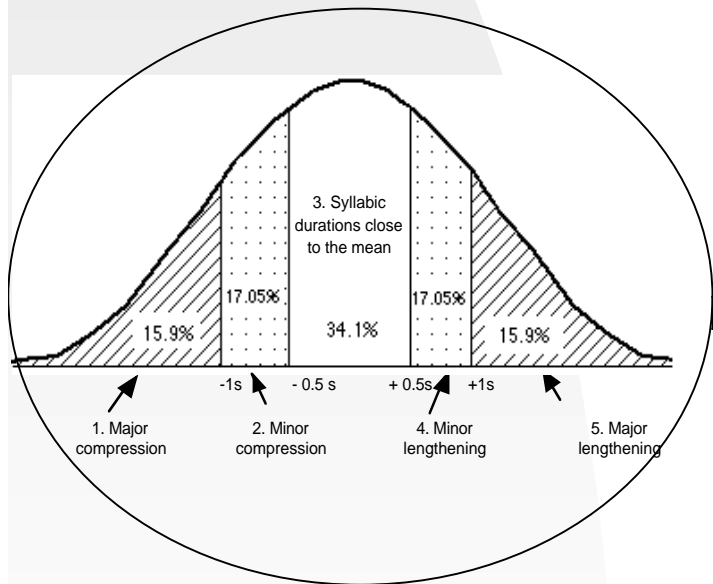


Subdivision des durations syllabiques après normalisation.



1. Groupes de mots

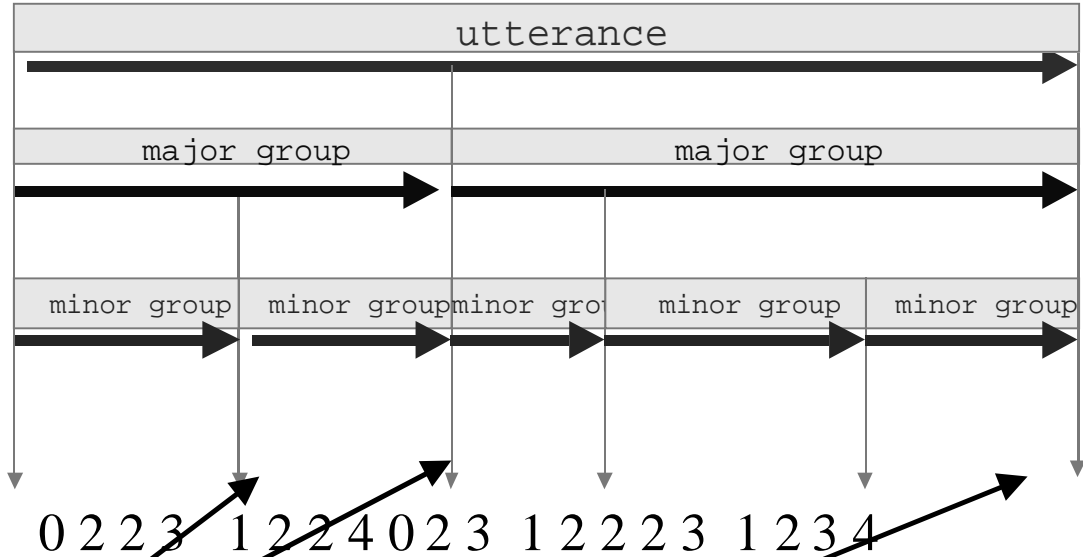
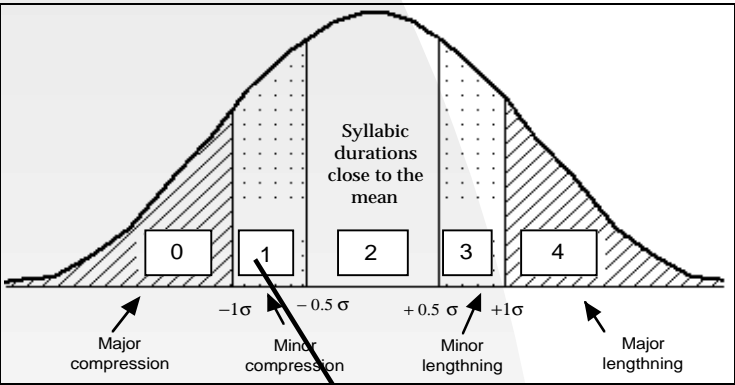
Quelle est la logique qui restitue la relation entre les durées et la structure temporelle?



Hy: des syllabes d'une certaine durée sont associées aux groupes majeurs, et d'autres aux groupes mineurs.



1. Groupes de mots



Des frontières temporelles signalées par les durées des syllabes

Deviation mineure
 compression mineure
 allongement mineur
 -> **Groupe mineur**
 -> début
 -> fin

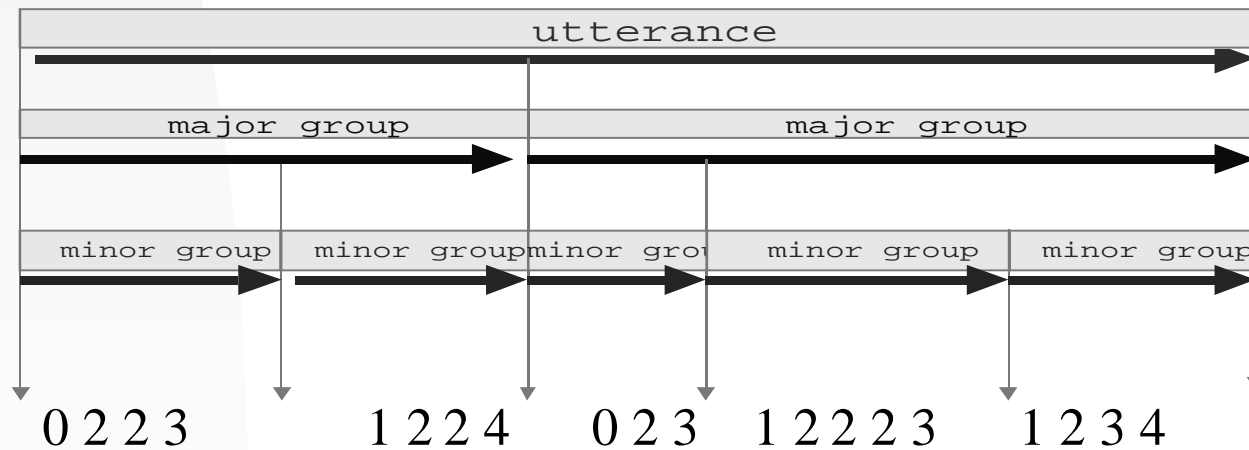
Deviation majeure
 compression majeure
 allongement majeure
 -> **Groupe majeur**
 -> début
 -> fin



1. Groupes de mots

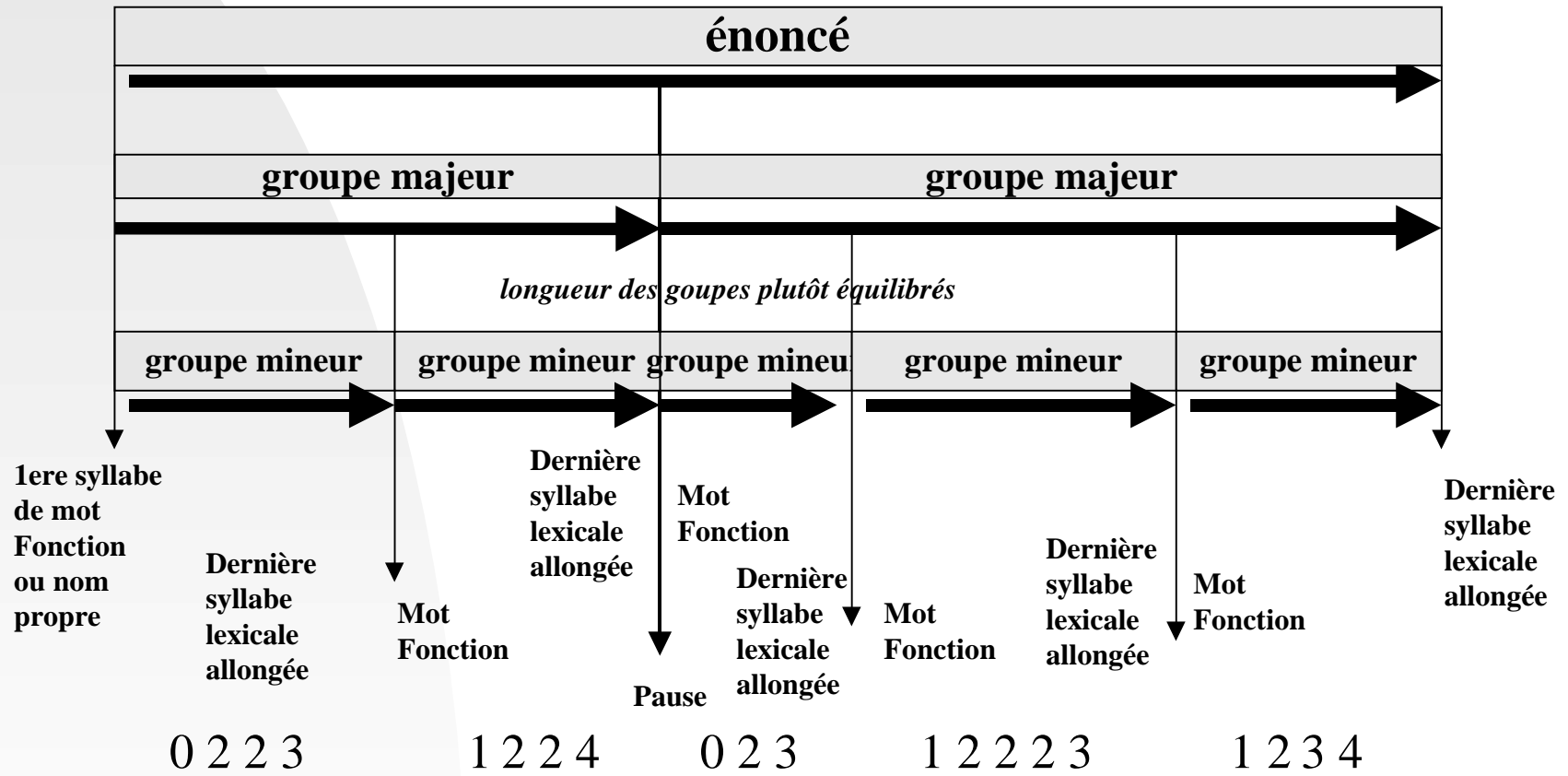
Evidence de relations régulières ($r > 0.90$ pour débits rapide et lent) entre

- *Types de mots (lexical, grammatical) et frontières temporelles (i.e., type de durées)*
- *Pauses et frontières temporelles (i.e., type de durées)*





1. Groupes de mots



L'analyse empirique montre que l'information émerge de blocs d'information mineur et majeur qui sont prédictibles.



Laboratoire d'Analyse Informatique de la Parole

1. Groupes de mots

Etudes empiriques de référence

Keller and Zellner, 1993; 1995; 1998;

Zellner, 1994; 1996; 1998

Confirmation des études psycholinguistiques

Goldman-Eisler, 1968, 1972;

Boomer & Ditman, 1962;

Cook, Smith & Lalljee, 1974;

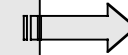
Chafe, 1980,

Grosjean & Deschamps, 1975;

Zellner, 1992.

1. Groupes de mots: défis futurs

**“Le problème nous ce qui nous fait peur c’est
qu’elle se retrouve dans une structure inadaptée”.**

**Elan****peur**

1. Groupes de mots: défis futurs

“Le problème nous ce qui nous fait peur c’est qu’elle se retrouve dans une structure inadaptée”.

Elan



peur

LAIPTTS



problème

Cette structure est prosodiquement contrainte:

- L’insertion d’une pause dans cette structure + un pattern intonatif inversé a pour effet de disloquer la structure attendue.
- On peut ainsi modifier l’attitude du locuteur.

1. Groupes de mots: défis futurs

Version A
sans pause

Version B
pause



1. Ce qui me plait en lui c'est son sens de l'humour

2. Non ça ne me dérange pas

L'introduction de la pause introduit un doute
sur la sincérité du locuteur

The logo for LAIP (Laboratoire d'Analyse Informatique de la Parole) features the letters 'LAIP' in a bold, serif font. The letters are white with a black outline and are set against a dark, textured background that resembles a close-up of a speaker grille or a similar acoustic surface.

Laboratoire d'Analyse Informatique de la Parole

1. Groupes de mots: défis futurs

Non ça ne me dérange pas

Exploration de l'organisation de la parole expressive

- paramètres des valeurs attitudinales
 - paramètres de la parole émotivement marquée
 - paramètres de la parole dialogique
 - paramètres de la parole "en situation"
 - paramètres idiolectaux
- etc...



Laboratoire d'Analyse Informatique de la Parole

**But d'un modèle du rythme de la parole:
SIMULATION
de la structuration temporelle d'un énoncé**



1. Groupes de mots

2. Durées

2. Prédiction des Durées

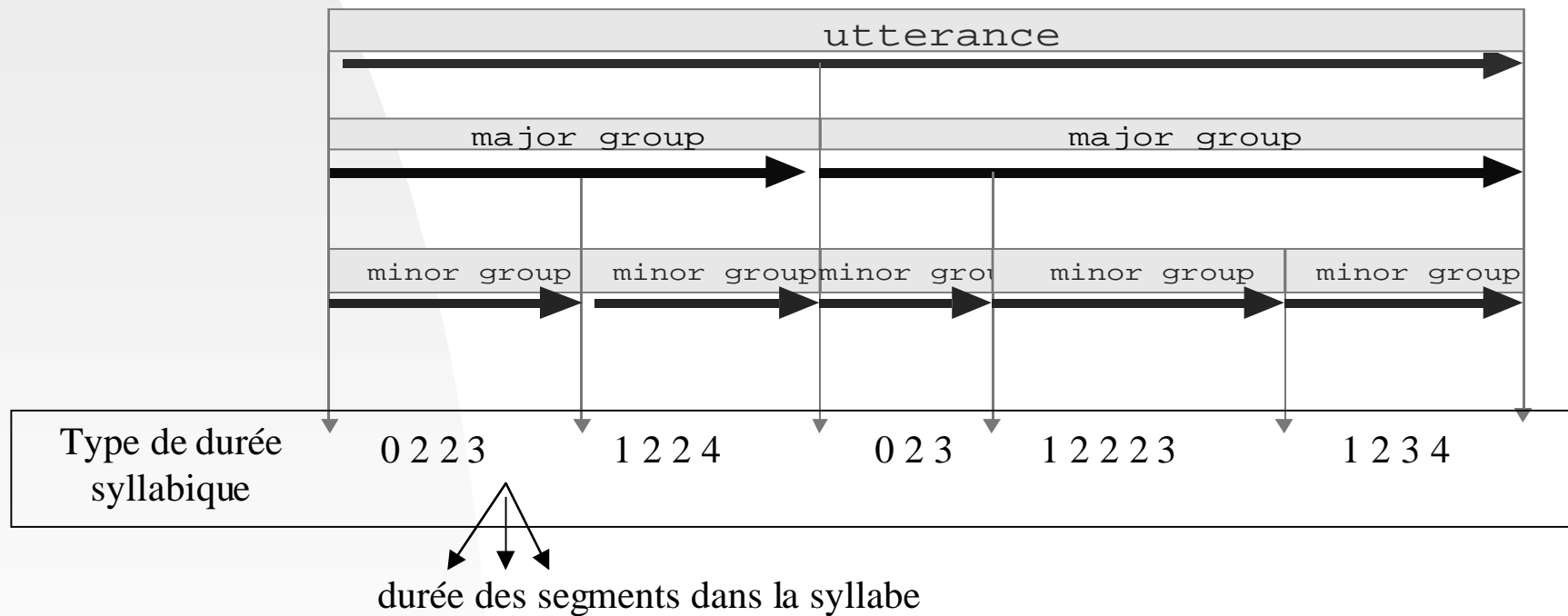
Hy: Le contrôle du flux d'information est relié à la dimension temporelle (dimension du SN).

Conséquence pour les durées des blocs d'information
(les groupes de mots, les pauses, les mots, les syllabes, les phones)

- problème quantitatif (x ms?)
- problème qualitatif (non linéarités)
- problème multi-niveaux (parole multimodale)



2. Prédiction des Durées



Construire un modèle statistique



2.a. Modèle statistique des Durées

• Prédiction des durées (Zellner, 1998)

Corrélation avec données initiales

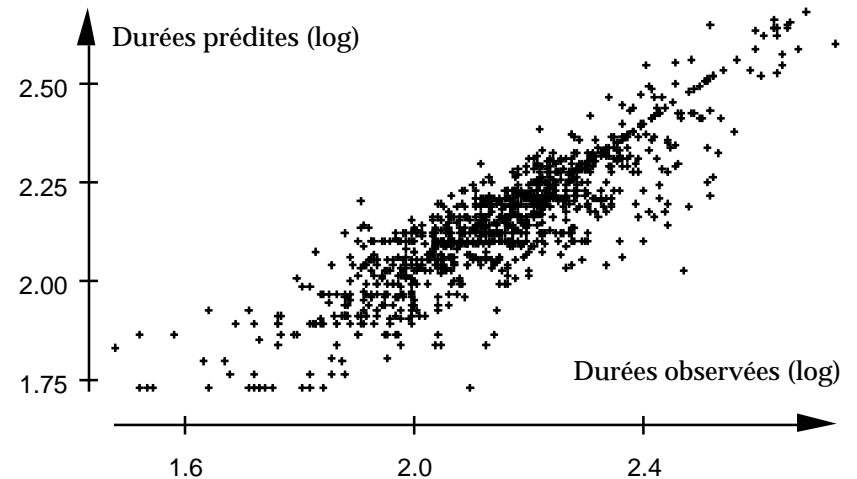
$r = 0.876$ (explained variance: 0.767, $p < 0.0001$)

Predicteurs

- a. Types de Segments* types de syllabes
- b. Frontières temporelles
- c. Nombre de segments dans la syllabe
- d. Presence/absence du schwa 1.63
- e. Mot Grammatical / lexical
- f. Mot Plurisyllabique

Poids dans la prédiction

79.24	%
11.64	%
5.39	%
%	
1.82	%
0.28	%

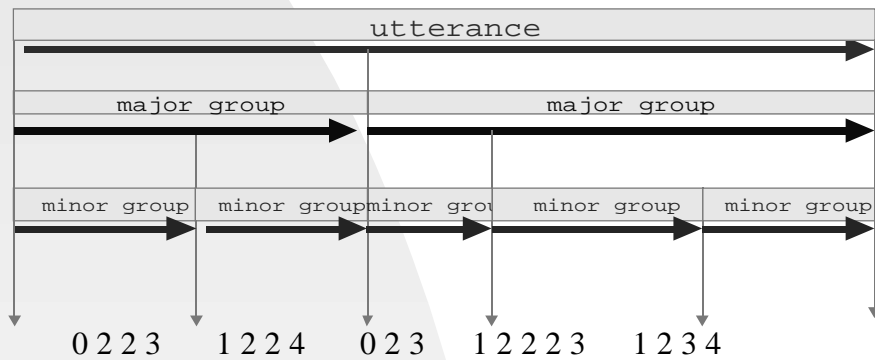


Je me demande si vous savez vraiment parler ou si vous avez encore des choses à apprendre en prosodie, en syntaxe ou en sémantique.



Laboratoire d'Analyse Informatique de la Parole

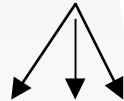
2.b. Ajustement qualitatif



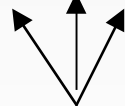
GLM

Durée prédite par le modèle statistique (quantitatif)

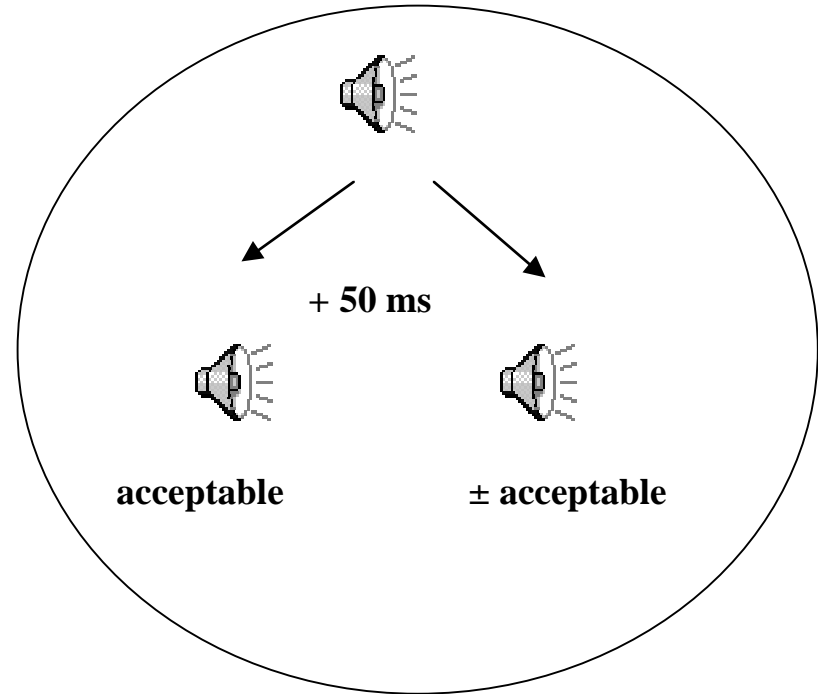
170 ms



Répartition de la durée entre les segments



Durée segmentale "intrinsèque"



Ajustement qualitatif:
structure syllabique





Laboratoire d'Analyse Informatique de la Parole

2.b. Ajustement qualitatif: la composante phonologique

La durée de la syllabe ne se répartit pas de manière uniforme entre ses composantes.

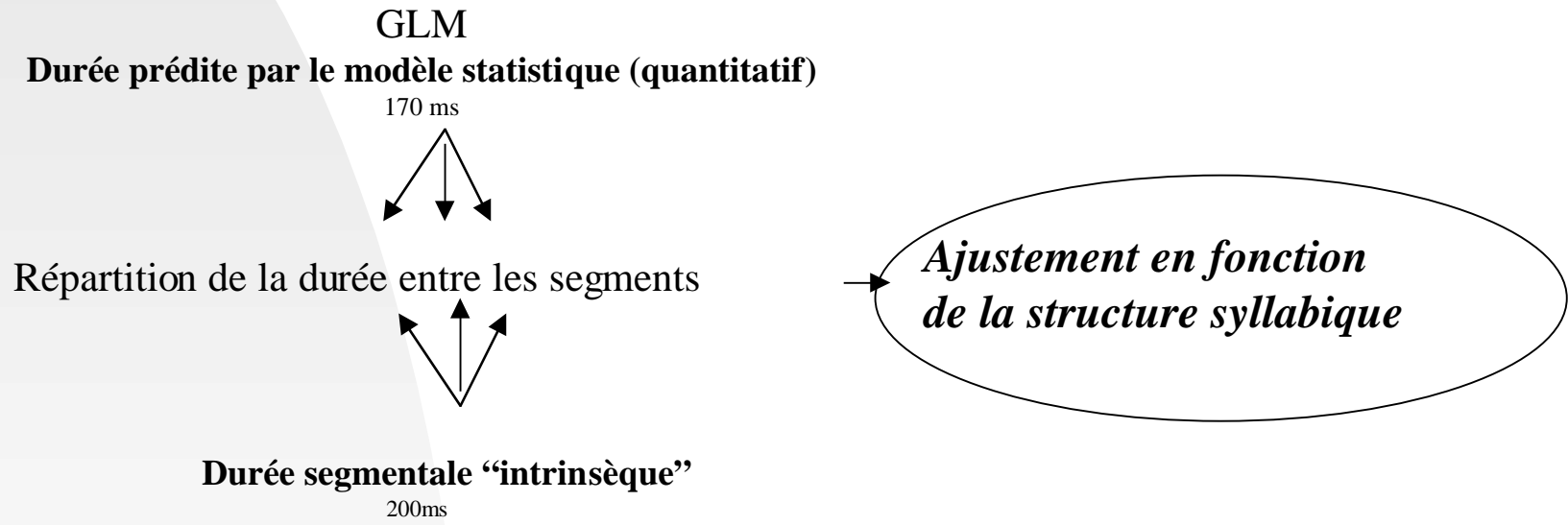
Les coefficients de Pearson montrent que la durée des composantes syllabiques (onset, nucleus, coda) est plus ou moins corrélée au type de syllabe.

Ex. débit rapide

Type	Onset	Nucleus	Coda	N
N		$r = 0.861$		71
ON	$r = 0.733$	$r = 0.597$	—	716
ONC	$r = 0.526$	$r = 0.502$	$r = 0.531$	165
NC	—	$r = 0.749$	$r = 0.603$	16



**2.b. Ajustement qualitatif:
la composante phonologique**



Je me demande si vous savez vraiment parler ou si vous avez encore des choses à apprendre en prosodie, en syntaxe ou en sémantique.



A ce stade, les durées sont prédites sur la base d'un modèle linéaire général qui est ensuite adapté à la structure syllabique

dernière étape:

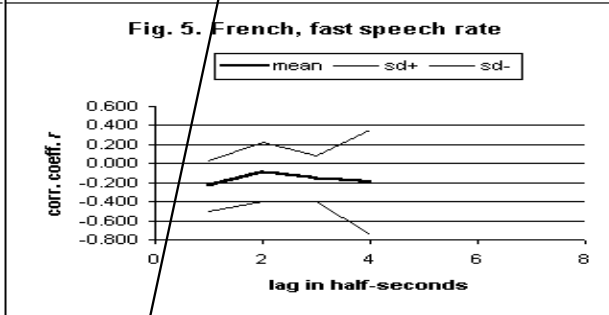
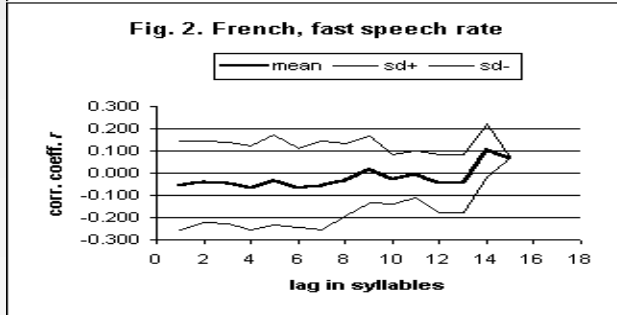
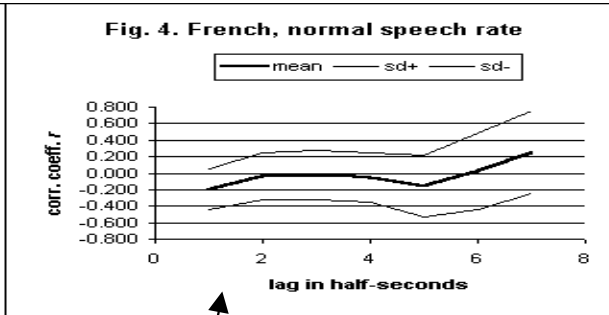
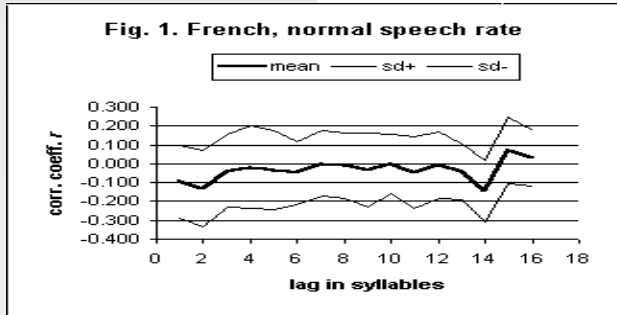
=> effets adjacents entre les durées:

i.e., la contrainte sérielle

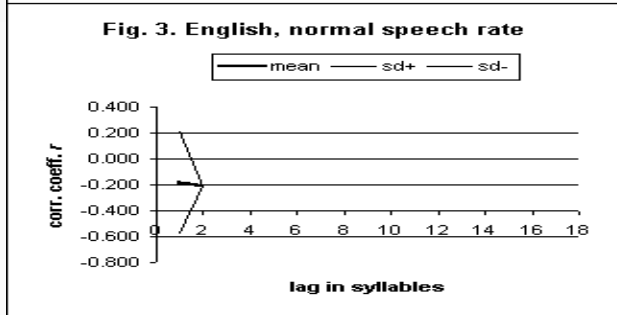


Laboratoire d'Analyse Informatique de la Parole

2.c. La contrainte s rielle



Les r sultats (corr lation n gative   lag 1) ont montr  que dans une fen tre de 500 ms, une dur e br ve a tendance    tre suivie d'une dur e plus longue



French, nrm, lag1, $r = -.197^*$ $N = 389$
 French, fast, lag 1, $r = -.234^*$ $N = 226$



Laboratoire d'Analyse Informatique de la Parole

2.Durées

La reconstruction de la parole lue nous a beaucoup appris du point de vue du rythme temporel:

- structure des unités cohérentes
- aspects quantitatifs, qualitatifs et multi-niveaux des durées.

Références

- Keller, E., & Zellner, B. (1995);**
Keller, E., & Zellner, B. (1996);
Keller E., Zellner Keller, B. & Local, J. (2000).
Zellner, B. (1992).
Zellner, B. (1996)
Zellner, B. (1998.).
Zellner Keller, B. (2001).
Zellner Keller B. (2001).
Zellner Keller, B., Keller, E. (2001).
Zellner Keller, B., & Keller, E. (to appear).



Laboratoire d'Analyse Informatique de la Parole

2.Durées: défis futurs

L'encodage prosodique résulte d'une combinaison de paramètres temporels et intonatifs.

- Comment modéliser cette combinaison pour simuler une parole expressive?
- Comment cette combinaison s'ajuste t-elle selon les besoins de communication? (notion de flexibilité)



Laboratoire d'Analyse Informatique de la Parole

Résumé

*La durée des unités de parole
résulte:*

**d'un bon parseur
d'un modèle statistique multi-
niveaux**

**Les durées sont adaptées à la
structure syllabique**

**Les durées sont ajustées selon les
interactions avec les syllabes
adjacentes**

Fluence

Cohérence

*Simulation
réaliste
du rythme
de la parole
lue*

The logo for the Laboratory of Computer Analysis of Speech (LAIP) features the letters 'LAIP' in a bold, serif font. The letters are white with a black outline and are set against a dark, textured rectangular background.

Laboratoire d'Analyse Informatique de la Parole

La synthèse de la parole offre un outil de simulation puissant pour tous les chercheurs qui s'intéressent à la communication orale. Elle permet par exemple de réaliser la synthèse des connaissances et des méthodes d'exploration en linguistique, en phonétique, en statistique ou en intelligence artificielle, permettant d'ouvrir des champs de compréhension sous des angles nouveaux.

La reconstruction du rythme temporel en était une illustration. Je vous remercie de votre attention.

