

# IMAGE ET REALITE VIRTUELLE

## Parole et langage

G. Bailly

Sujet d'examen du 26 février 2002 - 2 heures - Cours et documents autorisés

### 1. Synthèse de la parole – 10pts

Soit un système de synthèse opérant par concaténation d'unités stockées

- Donnez des arguments pour invalider le stockage d'allophones (réalisations de phonèmes en contexte)
- Qu'est ce qu'un diphone ? Quelles informations a minima doit-on y stocker ? Peut-on envisager de ne stocker que le signal ?
- Citez d'autres unités candidates
- Quels sont les avantages de la synthèse par diphones par rapport à une synthèse par mots ?
- Quels sont les problèmes majeurs posés par la simple concaténation de mots ? Quels types de variantes de prononciations de mots devrait-on enregistrer ?
- Nous utilisons couramment un vocabulaire de 6000 mots. Etant donné qu'un mot dure en moyenne 500ms, de combien de mémoire doit disposer le système qui stockerait 10 versions de chaque mot, chaque échantillon étant codé sur 16bits et le signal étant échantillonné à 16kHz ? Est-ce envisageable sur la mémoire de travail d'un ordinateur personnel ?

L'entrée d'un système de synthèse à partir du texte est – la plupart du temps - une chaîne de caractères alphanumériques

- Quels sont les informations contenues dans cette chaîne (ne nécessitant pas de gros lexiques) sur lesquelles les divers modules d'un système de synthèse peuvent s'appuyer.
- Quels pré-traitements de cette chaîne peuvent faciliter les traitements linguistiques ultérieurs ?
- Quels problèmes posent la synthèse d'e-mails ?

### 2. Reconnaissance de la parole – 5pts

Soit un système de dictée vocale utilisant un système de reconnaissance des 26 lettres de l'alphabet. Il doit permettre de dicter un texte en épelant chaque lettre des mots.

- Quels sont les « mots » supplémentaires de formatage du texte que l'on doit ajouter afin de pouvoir dicter un texte ?
- Quels seront les groupes de lettres qui engendreront le plus de confusions ?
- Quel sont les avantages et inconvénients d'un tel système de dictée vis à vis d'un système de reconnaissance de mots connectés (a) du point de vue du concepteur ? (b) du point de vue de l'utilisateur ?

### 3. Synthèse visuelle de la parole – 5pts

Soit un système qui prétend permettre aux sourds de comprendre une conversation téléphonique

- Quelles solutions peut-il avoir adopté ?
- Comment peut-on estimer des mouvements à partir du son ?
- Pourquoi un visage parlant aura-t-il des difficultés à assurer un niveau de compréhension suffisant ? Comment peut-on améliorer ce niveau de compréhension ?