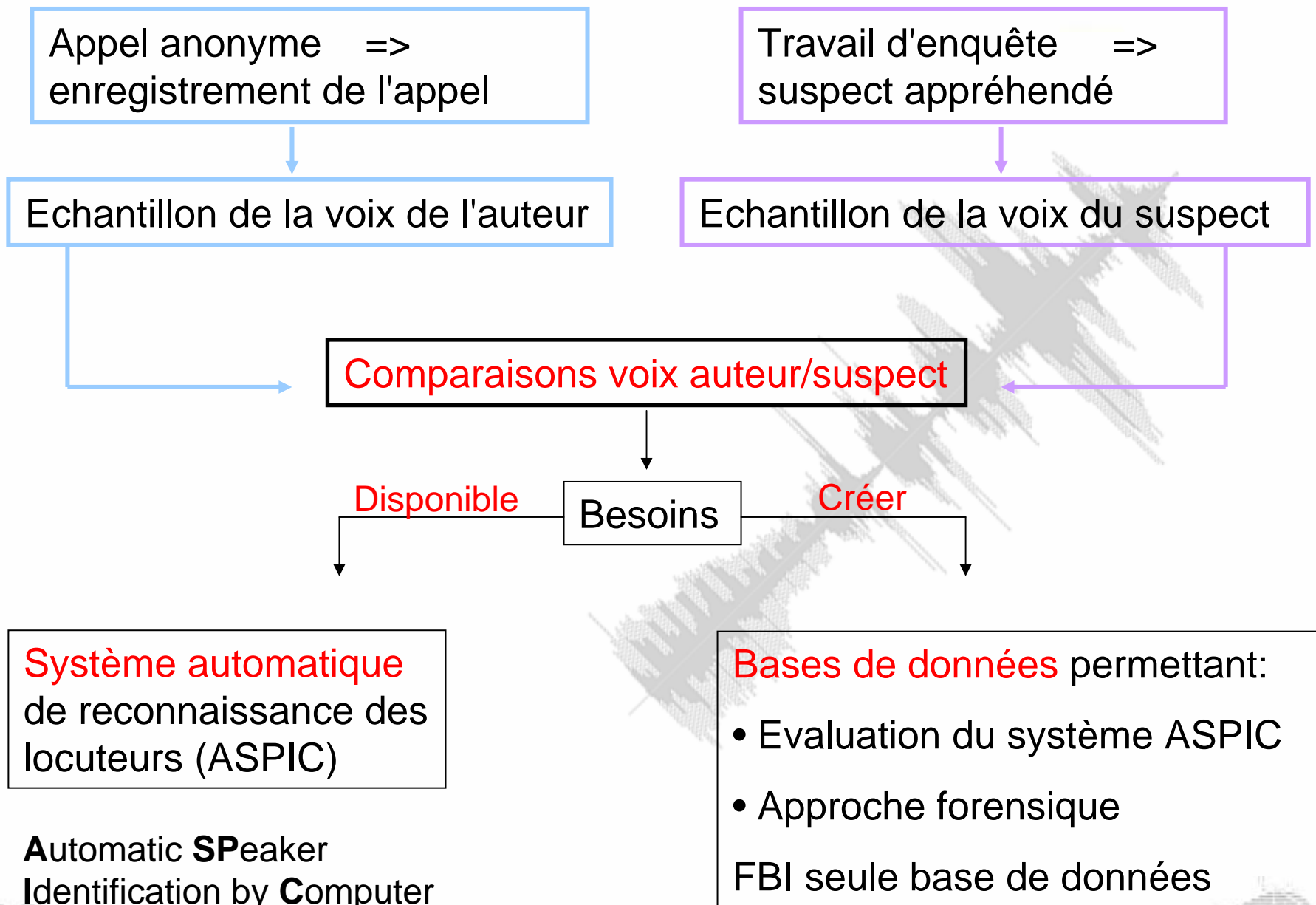


Analyse de l'Influence des Conditions d'Enregistrement dans la Reconnaissance Automatique de Locuteurs en Sciences Forensiques

Séminaire de 4^e année
Philipp Zimmermann



Système automatique de reconnaissance des locuteurs (ASPIC)

Automatic **SP**eaker Identification by **C**omputer

Bases de données permettant:

- Evaluation du système ASPIC
- Approche forensique

FBI seule base de données

- Plan de la présentation
 - Introduction
 - La voix comme donnée biométrique
 - La voix comme indice en sciences forensiques
 - Partie théorique
 - Fonctionnement d'un système automatique
 - Approche forensique de la reconnaissance des locuteurs
 - Partie pratique
 - Présentation de la base de données IPSC03
 - Utilisation de la base de données créée
 - Discussion
 - Conclusion

■ Réflexions sur la voix

- [Bertillon, 1881]
 - Timbre comme élément distinctif, signe difficile à noter
- [Ottolenghi, 1910]
 - Signal complexe, appréhension sous plusieurs facettes nécessaire (Volume, ton, timbre et intonation)
- [Locard, 1932]
 - Dépendance des états physiologique et psychologique sur les conditions de production et de perception du signal de parole

■ La voix comme donnée biométrique

■ La voix

- Vibration des cordes vocales => production d'un bruit de friction ('s') ou de plosion ('t') (phonation)
- Filtrage par le système vocal (articulation)

- Expérience quotidienne => caractéristiques individuelles
=> potentiel d'identification

■ Mais

- Voix caractérisée par un mélange d'individualités anatomiques et comportementales
- Echecs dans la détermination et l'isolement des données de l'unicité et de l'immuabilité

■ La voix comme indice

- En sciences forensiques: trace essentiellement de type enregistrement téléphonique
 - Appels anonymes (art 179 septies CP)
 - Ecoutes téléphoniques (par ex: art 179 octies CP)
- Par enregistrement direct sur support analogique/numérique

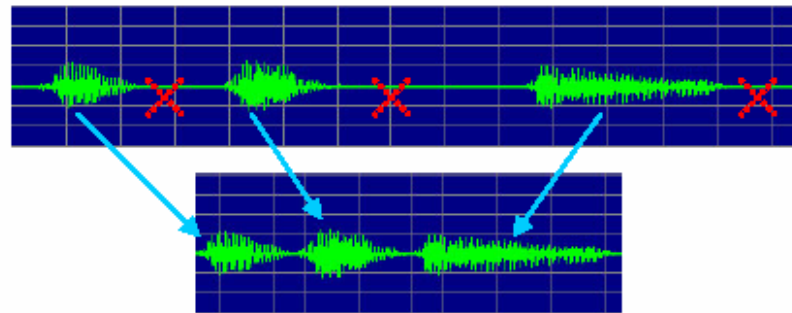
- La voix comme indice => processus et conséquences
 - Acquisition du signal [téléphone, microphone]
 - Transmission et codage [GSM, RTPC]
 - Enregistrement [serveur, enregistreur]

Compréhension de
l'intervention de
chaque élément

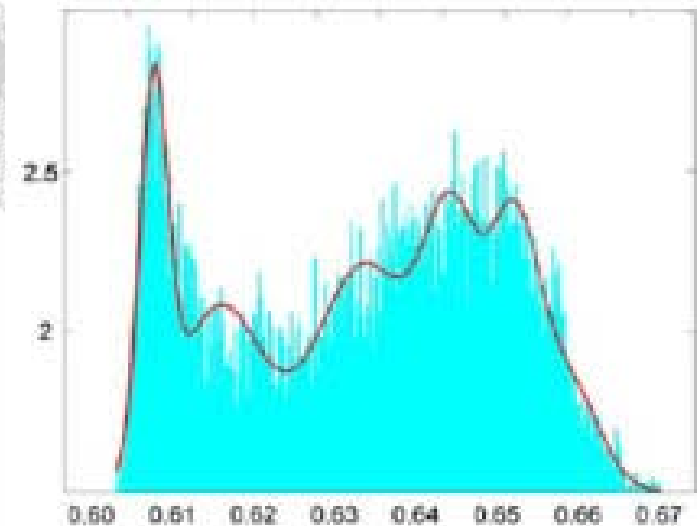
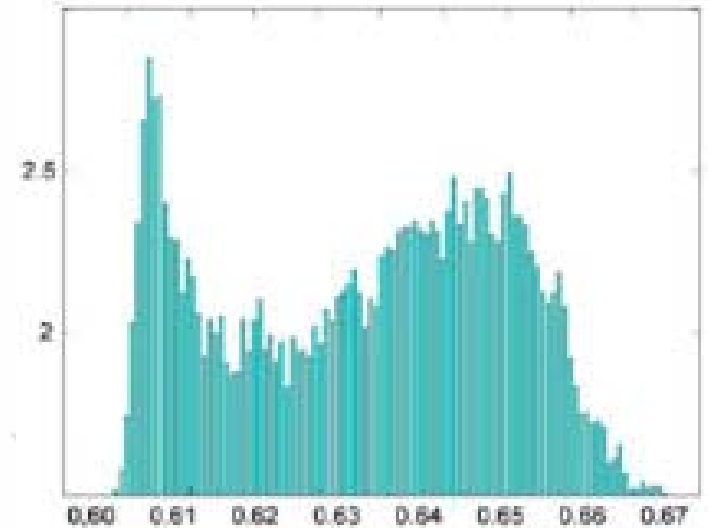


Evaluation de la
valeur probante de
l'indice

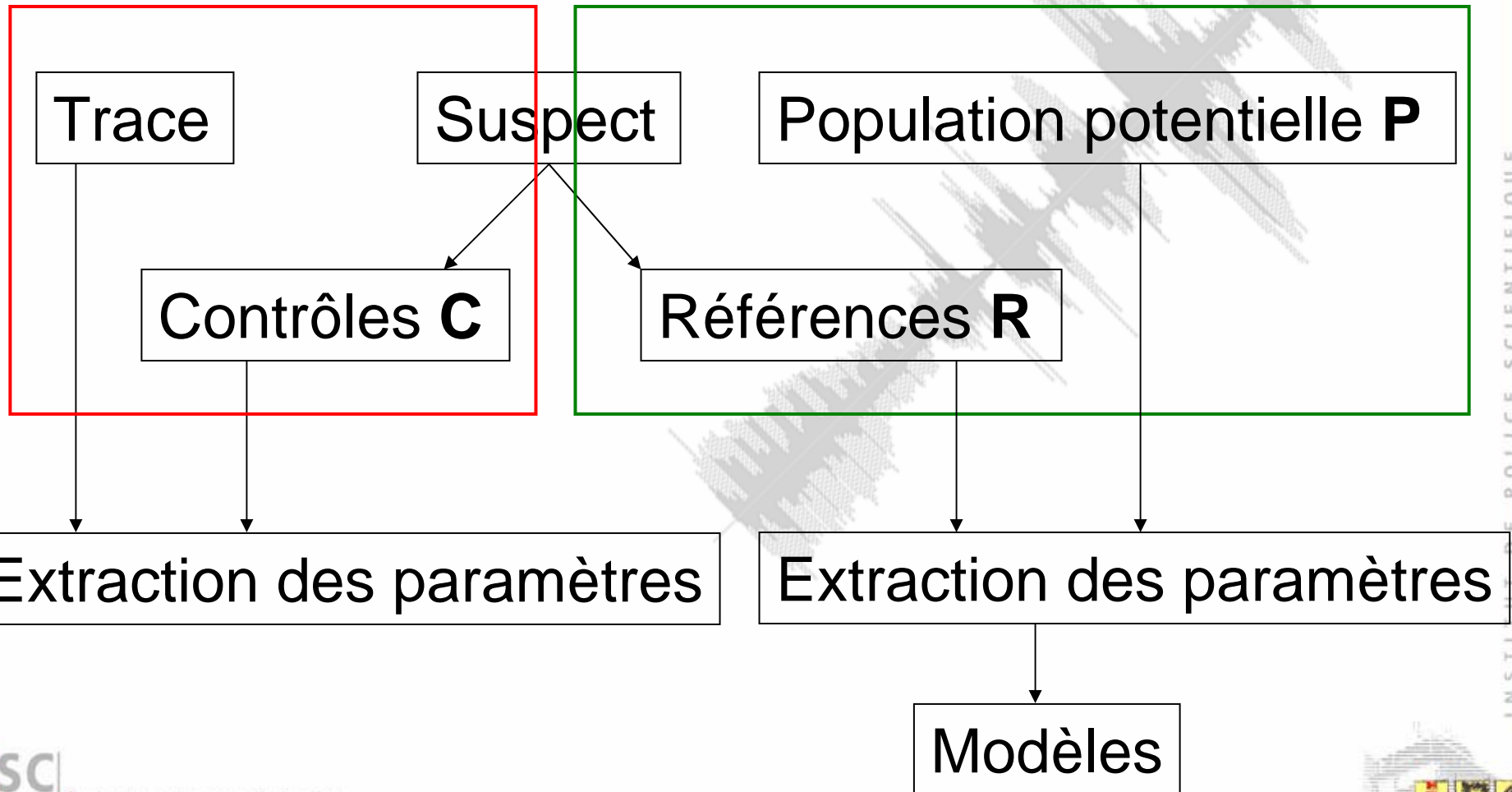
- Système automatique
 - Acquisition du signal
 - Enregistrement du signal
 - Audition
 - trace exploitable? Choix des bases de données
 - Prétraitement du signal
 - Suppression des zones de silence

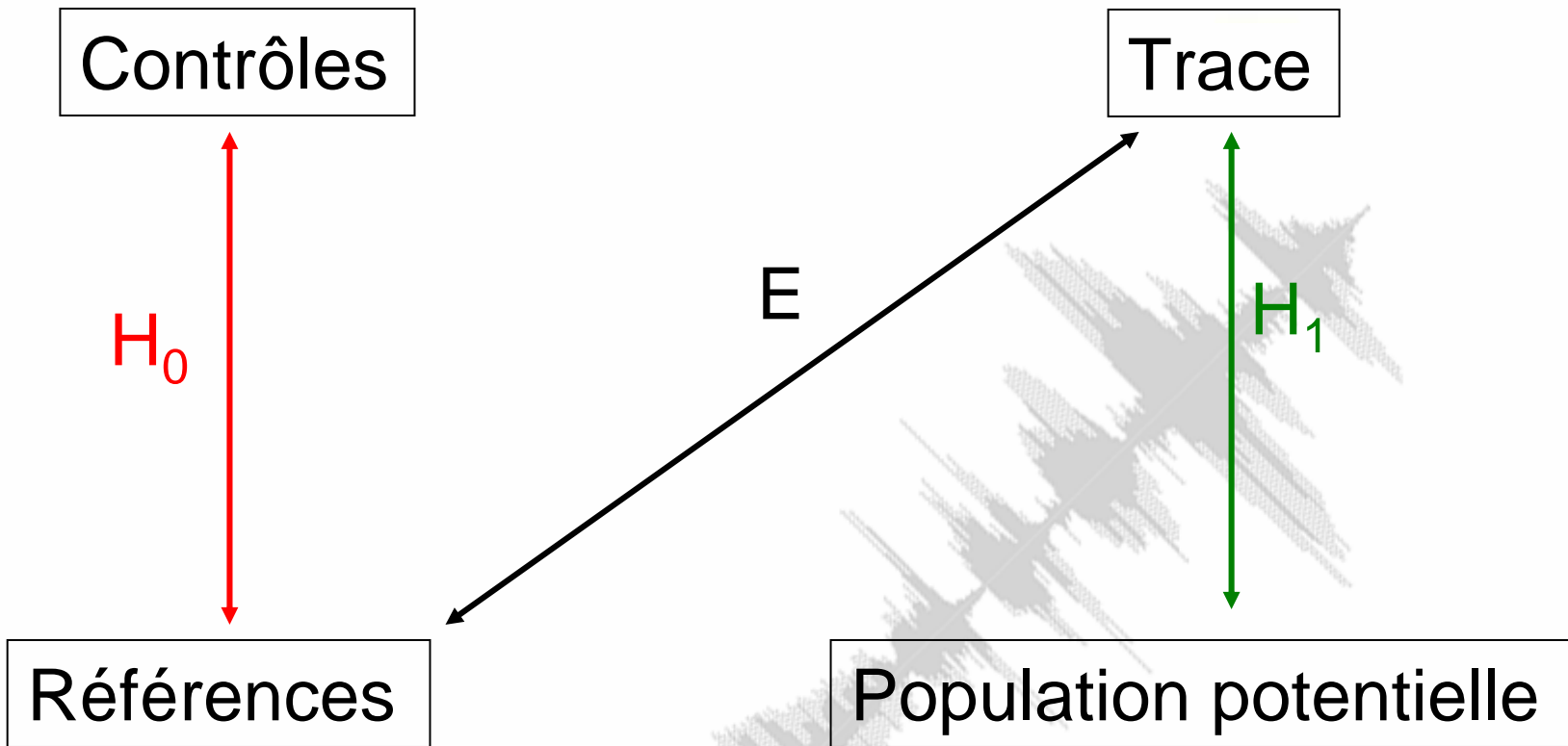


- Système automatique
 - Extraction des paramètres caractéristiques de la voix (méthode RASTA-PLP)
 - Modélisation statistique des paramètres extraits (méthode GMM)



- Comparaisons Trace / Modèle dans un cadre bayésien [Meuwly et Drygajlo, 2001], [Botti et al., 2004]

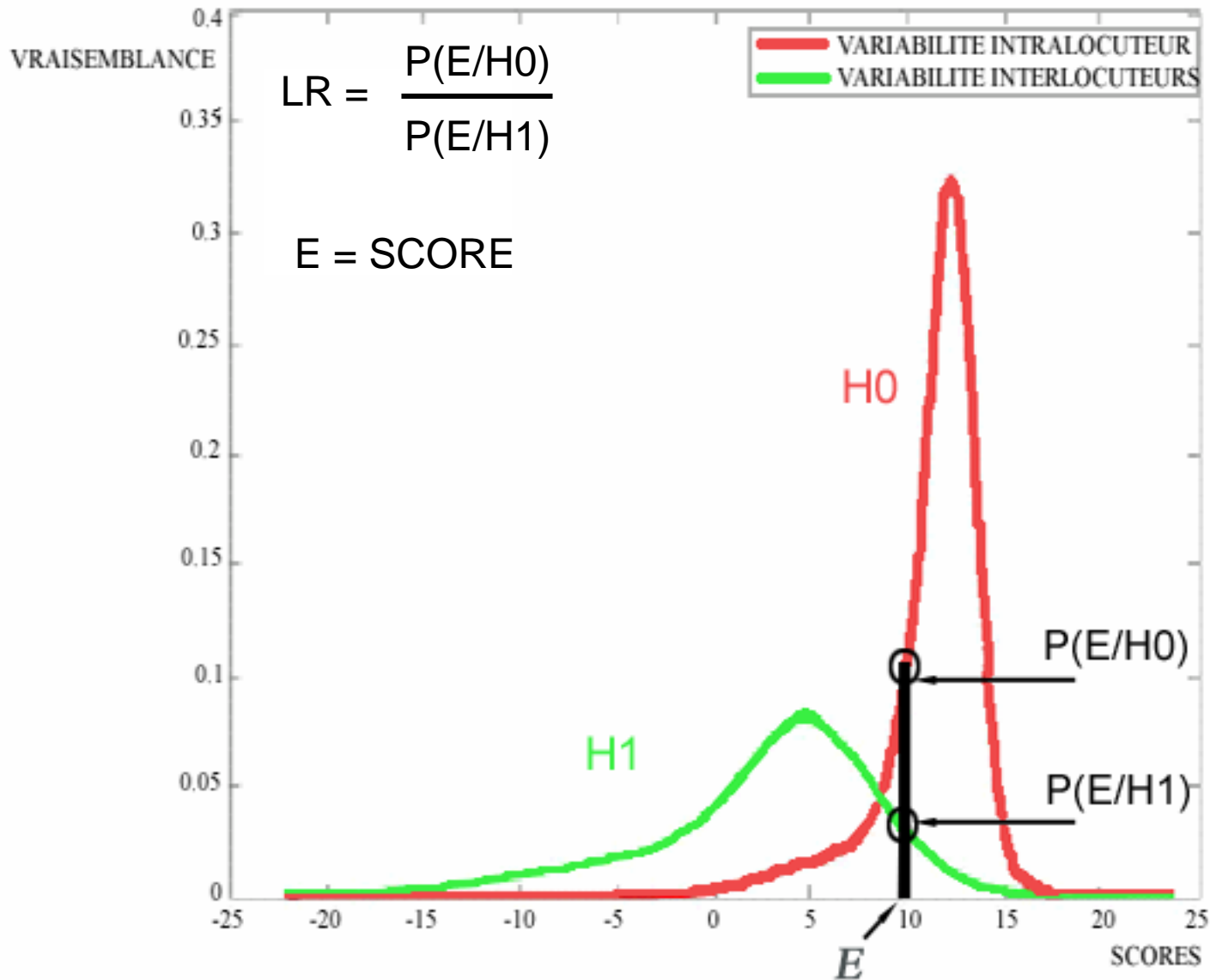




H_0 : le suspect est à l'origine de la trace (variabilité intralocuteur)

H_1 : le suspect n'est pas à l'origine de la trace (variabilité interlocuteur)

E = score; force probante de l'indice



Cas ou H0 est vraie

- Bases de données
 - Travaux réalisés:
 - Meuwly [IPSC01], Rossy [IPSC02], FBI, NIST, MIT, IBM...
 - Problème:
 - Ces bases de données rendent l'analyse de l'influence des conditions d'enregistrement difficile

- Base de données
 - Paramètres qui doivent être pris en compte
 - Taille (stat. représentative)
 - Langue (adaptée à la trace)
 - Conditions (adaptées à la trace)
 - Acquisition/Transmission/Enregistrement
 - Contenu
 - Possibilité de choisir entre des enregistrements de type: **T**race, **C**ontrôle ou **R**éférence

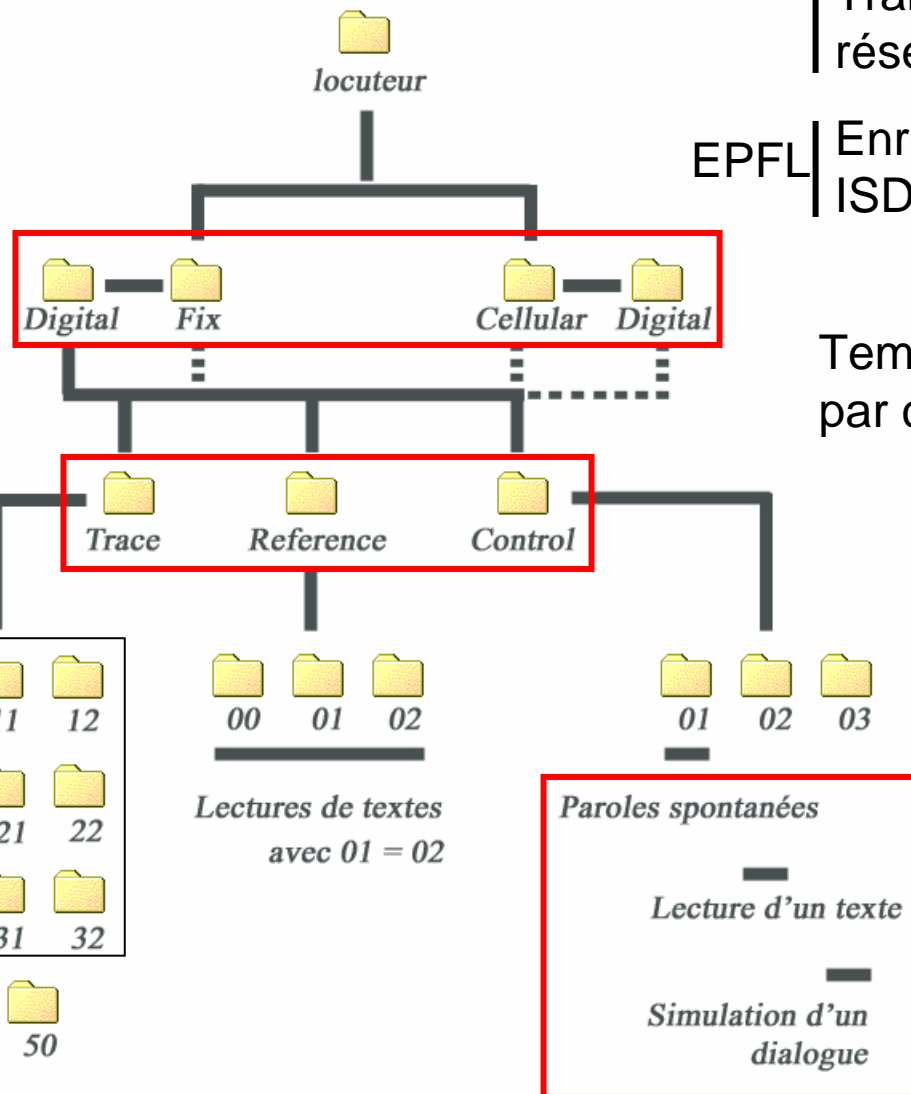
■ Base de données IPSC03

IPS | Acquisition: microphone
Transmission: par réseaux GSM, RPCT

EPFL | Enregistrement: serveur ISDN

Temps d'enregistrement par condition: 10-15 min.

10, 11, 12
20, 21, 22
30, 31, 32
Contenus linguistiques similaires



Lectures de textes

Paroles spontanées

Lectures de textes avec 01 = 02

Paroles spontanées

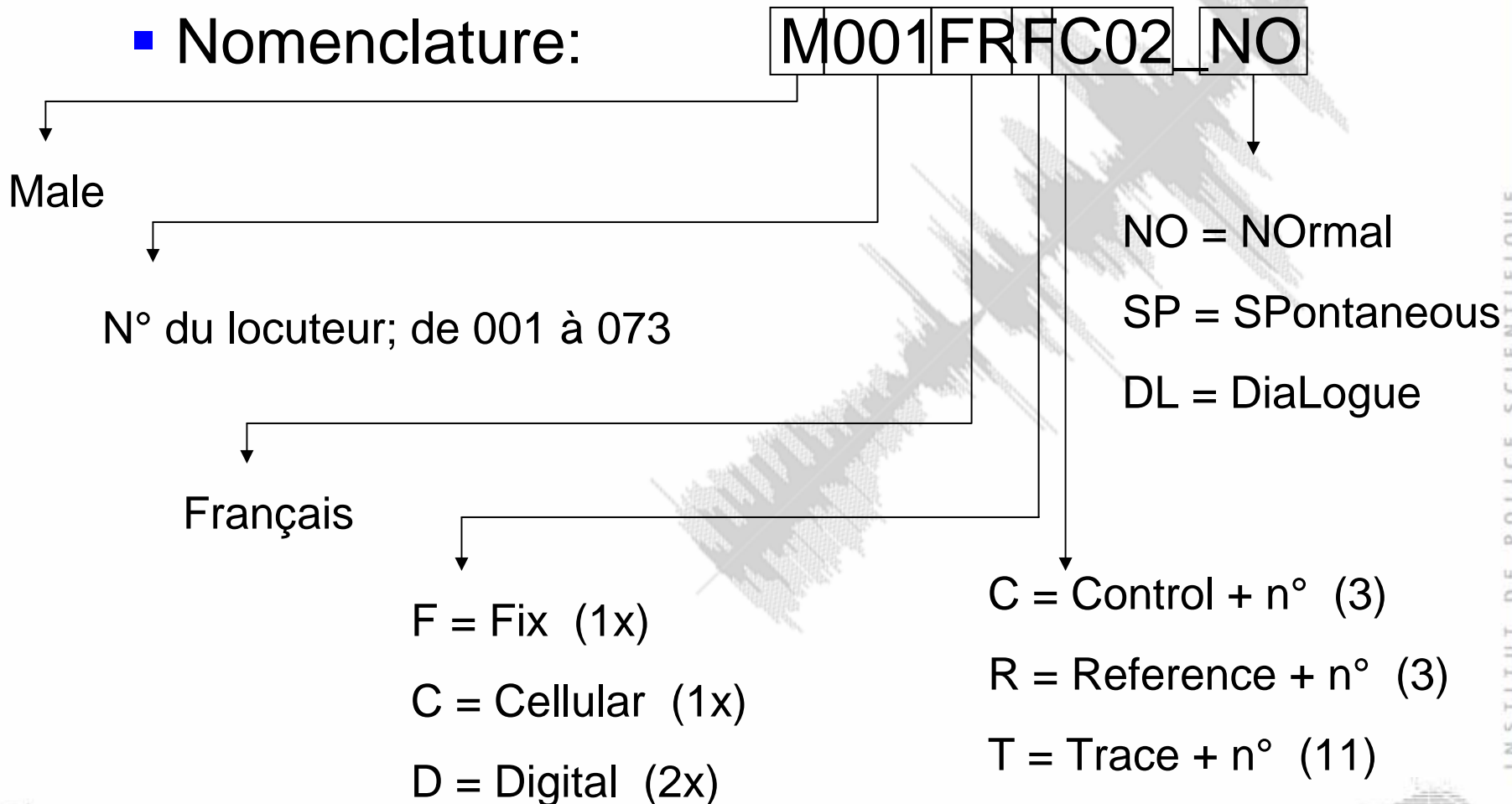
Lecture d'un texte

Simulation d'un dialogue

4798 fichiers > 40 heures

■ Base de données IPSC03

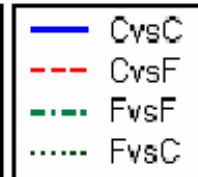
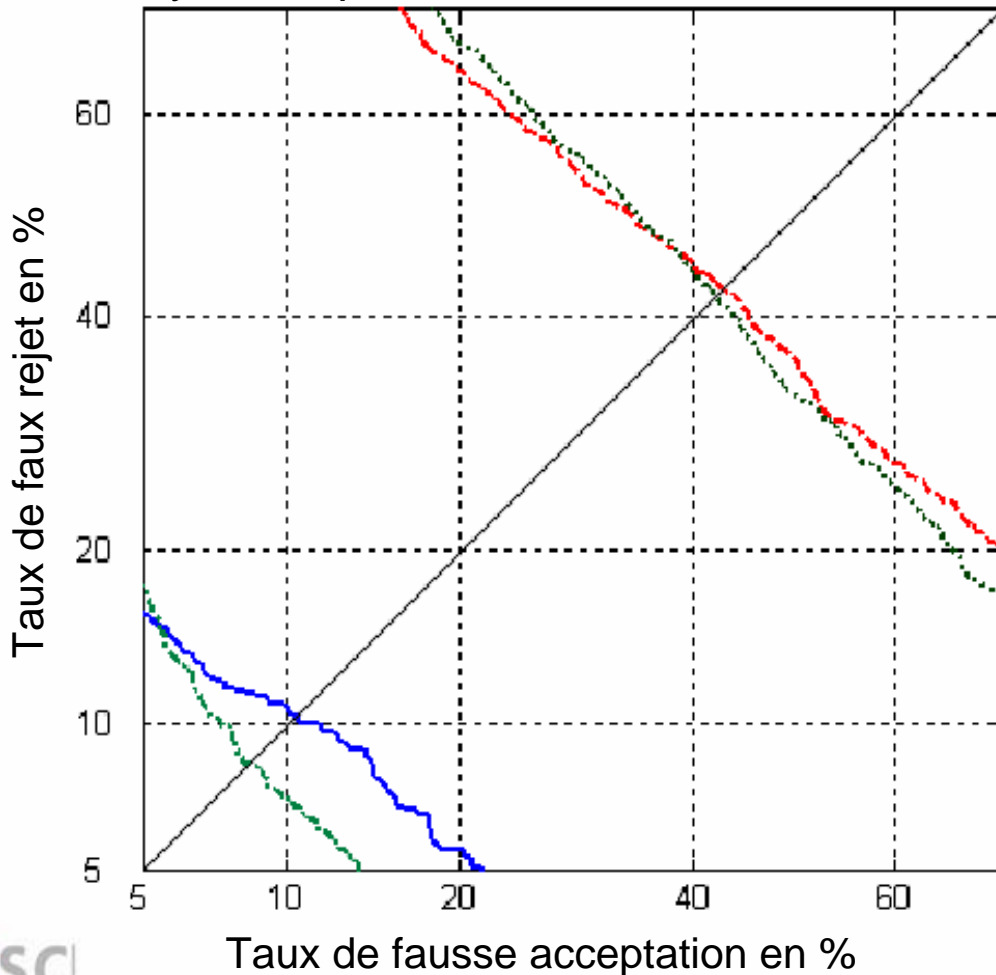
■ Nomenclature:



■ Base de données IPSC03;

Performance du système ASPIC

Représentation de la performance du système par des courbes DET



C = Cellular
F = Fix

| | EER % | DC |
|------|-------|------|
| FvsF | 8.37 | 1.11 |
| CvsC | 10.35 | 0.93 |
| FvsC | 42.00 | 0.12 |
| CvsF | 42.77 | 0.12 |

Comparaisons fichiers *trace* (3) vs fichiers *reference* (3)

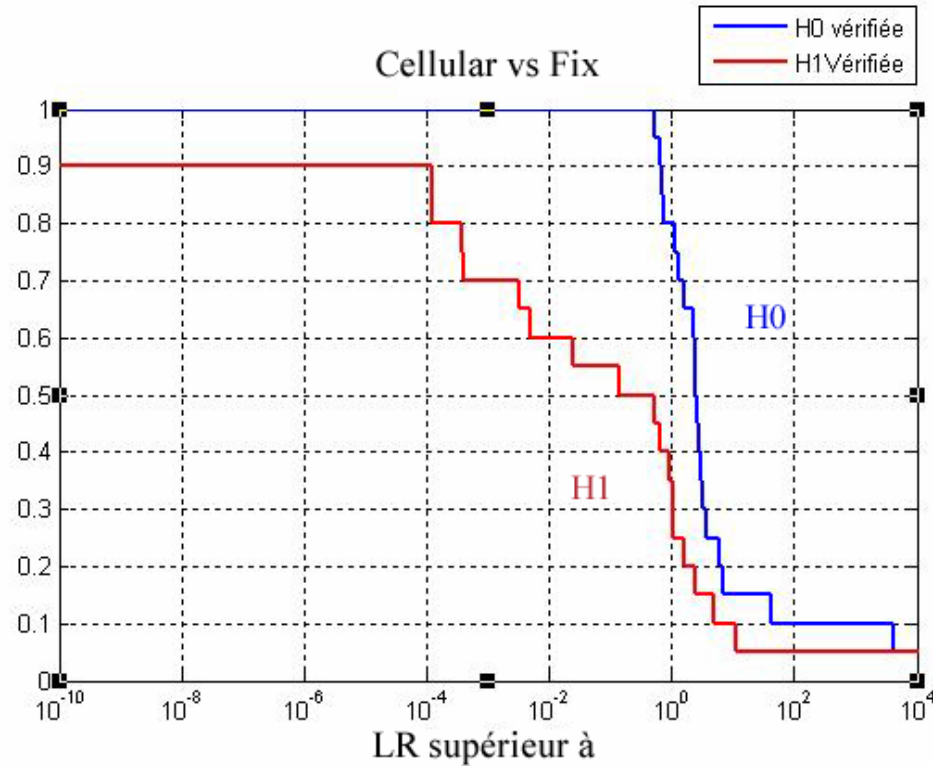
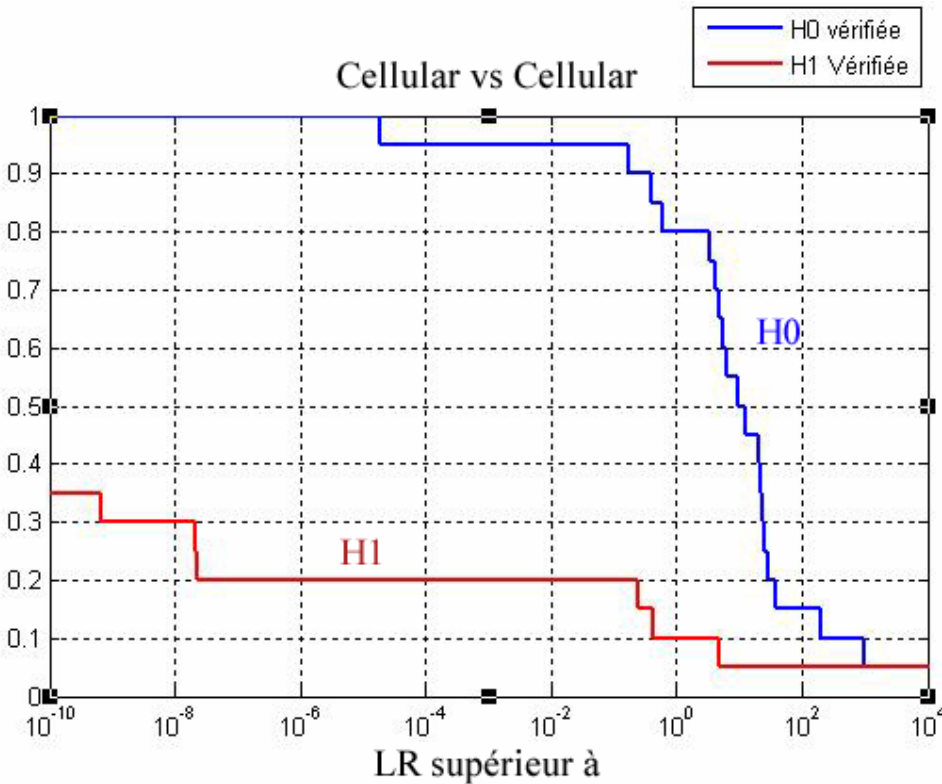
N = 73 locuteurs

■ Base de données IPSC03

Cas simulés

■ Approche forensique

Evolution du LR en fonction de *H0 vérifiée* ou de *H1 vérifiée*



■ 1 locuteur = suspect 72 locuteurs = population potentielle

N=20

■ Base de données IPSC03;

Cas simulés

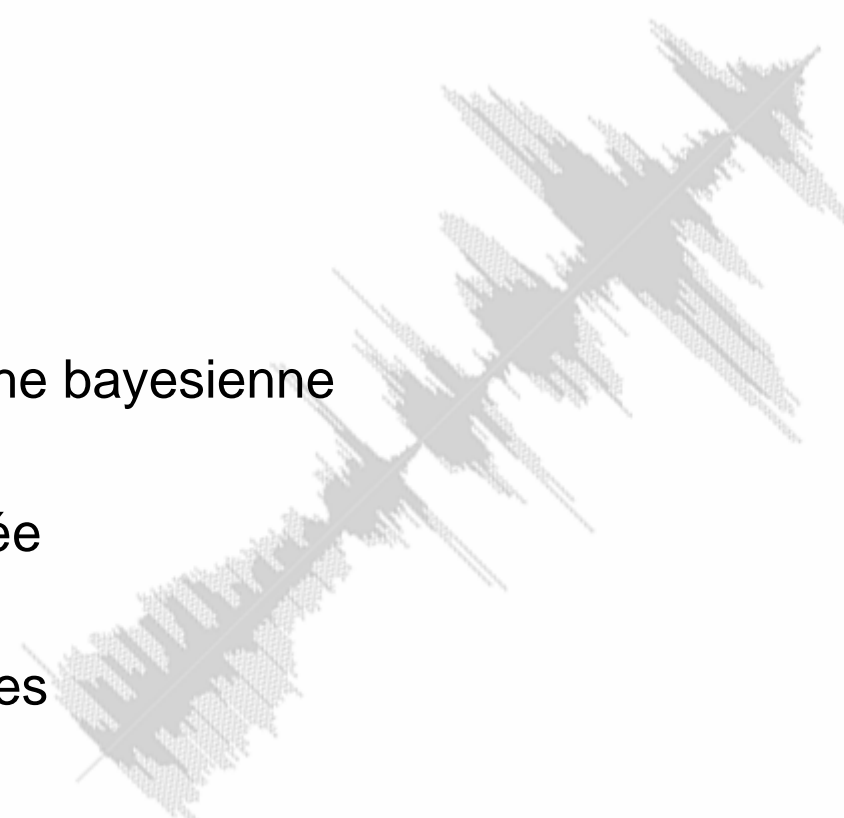
- Utilisation de la base de données pour de la recherche
 - **Arcienega M, Alexander A, Zimmermann P, Drygajlo A**
A Bayesian network approach combining pitch and spectral envelope features to reduce channel mismatch in speaker verification and forensic speaker recognition. Submitted for publication to *Interspeech*; 2005.
 - Evaluation de l'apport des **réseaux bayesiens** face à **l'approche bayésienne « classique »**
 - Modélisation par réseaux bayesiens de la fréquence fondamentale et de l'enveloppe spectrale de la voix (informations complémentaires mais non corrélées)
 - **Amélioration des performances du système**

- Discussion des résultats
 - Influence des conditions d'enregistrement
 - Importante en conditions différentes
 - Influence sur la force de l'indice
 - La maîtrise des paramètres d'influence comme élément de l'interprétation de la valeur de l'indice
 - La base de données a permis de mettre en évidence l'amélioration des résultats en conditions différentes par l'approche par réseaux bayesiens

■ Discussion des résultats

■ Base de données

- Taille substantielle
- Permet une approche bayésienne
- Peut être développée
- Applications multiples



■ Conclusion

■ Base de données créée

■ Un outil est à disposition

■ Utilisable pour:

- Tester les performances d'un système
- Evaluer la force de l'indice dans un cadre bayésien
- Analyser l'influence de la quantité de signal à disposition (trace et suspect)
- Analyser l'influence du contenu linguistique, phonétique
- Analyser l'influence du mode d'élocution
- Analyser l'évolution des résultats en fonction de la taille de la BD
- ...

- Dr Andrzej Drygajlo
 - ITS-EPFL
- Alexander Anil, assistant doctorant
 - ITS - EPFL
- Botti Filippo, assistant doctorant
 - ESC – IPS
- Dessimoz Damien, assistant doctorant
 - ESC - IPS
- Les locuteurs enregistrés