

UNIVERSITÉ D'AVIGNON  
ET DES PAYS DE VAUCLUSE

# Campagne d'évaluation ETAPE 2012

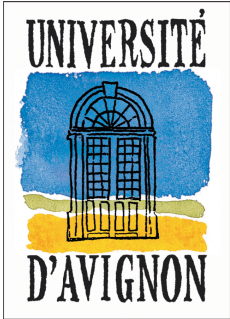
Participation du

## **Laboratoire d'Informatique d'Avignon**

à la tâche de

## **Transcription Automatique de la Parole**

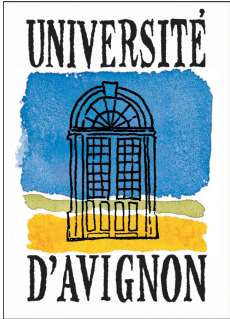
G. Linarès – G. Senay – B. Bigot  
M. Bouallege – E. Ferreira- C. Richier



UNIVERSITÉ D'AVIGNON  
ET DES PAYS DE VAUCLUSE

# Plan

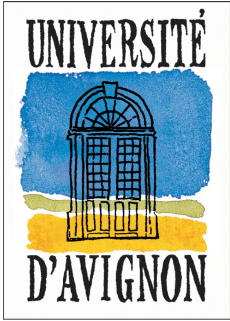
- Introduction
- Présentation du système de RAP
- Soumissions et résultats
- Analyse Factorielle pour la reconnaissance automatique de la parole
- Conclusion



UNIVERSITÉ D'AVIGNON  
ET DES PAYS DE VAUCLUSE

# Introduction

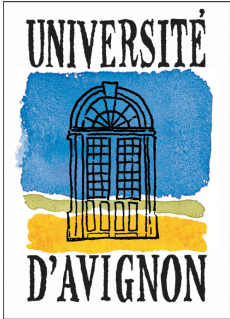
- Objectif
  - Remise à niveau des données du système PERCOL (défi REPERE)
- Participation
  - Mise à jour des ressources du système avec des données récentes
  - Mise en route tardive de ce travail et ressources humaines limitées



UNIVERSITÉ D'AVIGNON  
ET DES PAYS DE VAUCLUSE

# Systeme de reconnaissance de la parole du LIA

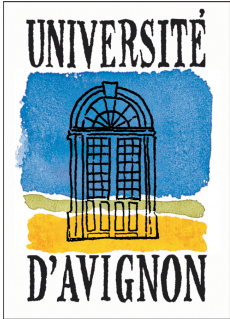
- Speeral
- Algorithme A\*
- Treillis de phonèmes contextualisés
- Décodage sur des triphones cross-word
- Modèles 4-gram
- Adaptation en locuteur en seconde passe



UNIVERSITÉ D'AVIGNON  
ET DES PAYS DE VAUCLUSE

# Modèles de langage

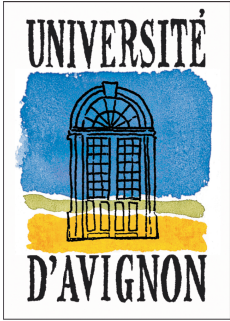
- Ressources :
  - ESTER1&2 (~20Mo), Gigaword (3,5Go), LeMonde (2,6Go), LeSoir (1,2Go), EPAC
  - Dépêches AFP(255 Mo), Wikipédia 2012 (2,2Go)  
News Web Crawling (2,2Go + 620Mo), REPERE
- Augmentation du lexique (ressources récentes)
  - ~ 90.680 entrées ( contre ~85.600)
  - Ajout d'~ 2580 noms propres
- ARPA : 4-gram (~12Go)



UNIVERSITÉ D'AVIGNON  
ET DES PAYS DE VAUCLUSE

# Modèles acoustiques

- Modèles Wide Band : *from scratch*
  - 250 heures de données d'apprentissage
  - Nouveau jeu de questions
- Modèles Narrow Band inchangés
- HMM :
  - 7K états émetteurs
  - Total de 465K composantes gaussiennes

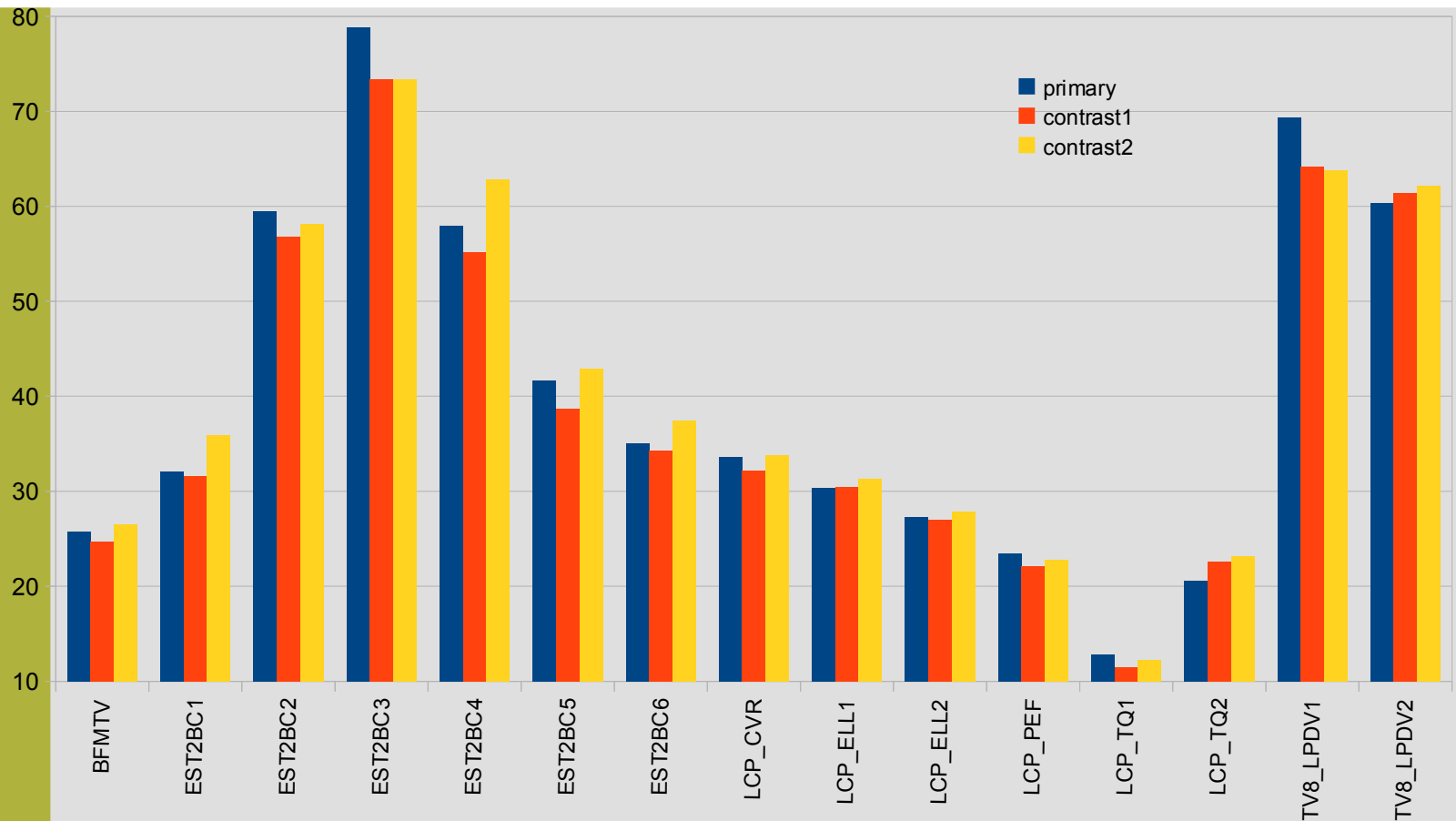


UNIVERSITÉ D'AVIGNON  
ET DES PAYS DE VAUCLUSE

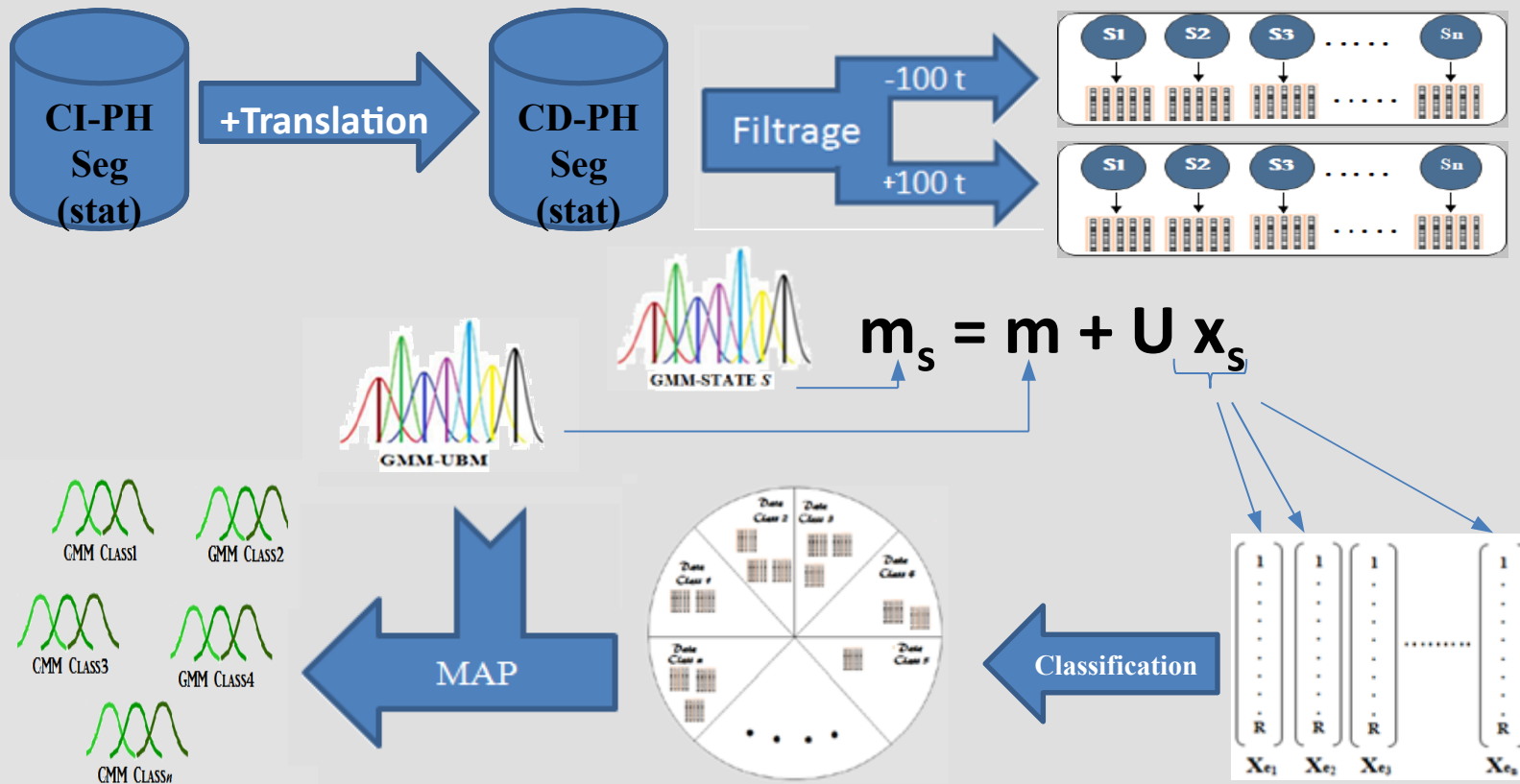
# Soumissions et résultats

- Primaire : SRL LIA + Speeral
- Contraste 1 : SpkDiarization LIUM + Speeral
- Contraste 2 : SpkDiarization LIUM + Speeral + apprentissage de modèles acoustiques par analyse factorielle

# Soumissions et résultats

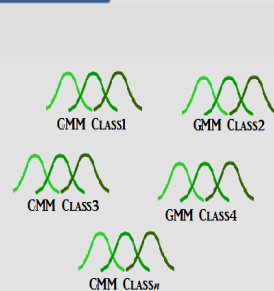
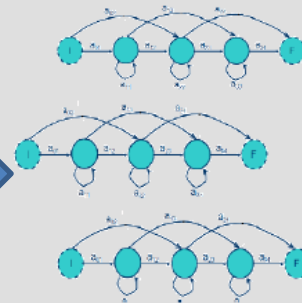
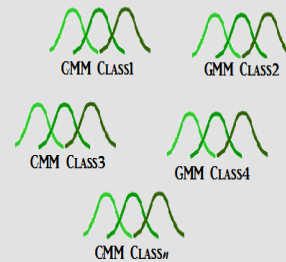
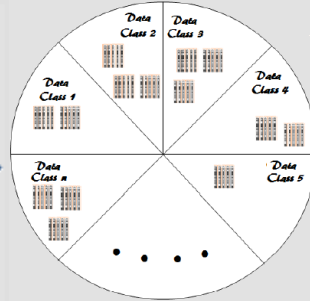
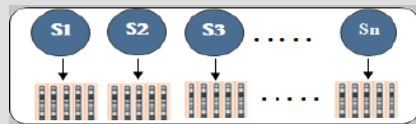






# Analyse Factorielle pour la reconnaissance de parole

-100 t



# Analyse Factorielle pour la reconnaissance de parole

## Test Ester

Station	Baseline (5079)	2672	3730
Inter (5h40mn)	34.88	33.32	32.85
RFI (1h10mn)	18.62	18.06	17.39
tvme (1h)	28.74	26.45	26.02
africa(1h30mn)	32.42	31.16	30.90
Total (9h30)	30.92	29.46	29.03
absolute gain	-	1.46	1.89

# Analyse Factorielle pour la reconnaissance de parole

## Compensation des variabilités sessions

### Apprentissage

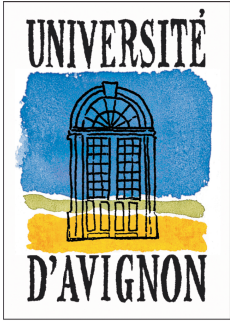
$$m_{observed} = m_{ubm} + Dy_{phoneme} + Ux_{speaker} + Vz_{radio}$$

### Filtrage

$$\hat{t} = t - \sum_{g=1}^M \gamma_g(t) \cdot (\{U \cdot x_{utterance}\}_{[g]} + \{V \cdot z_{utterance}\}_{[g]})$$

### Test Ester

	PASS-1	PASS-2
Baseline	29.6	27.5
Norm-speaker-radio	28.0	26.2



UNIVERSITÉ D'AVIGNON  
ET DES PAYS DE VAUCLUSE

# Conclusion

- Mise à jour des ressources pour les ML et les MA
- Intégration de technologies du LIA dans le contraste 2
- Poursuite du travail entamé dans le cadre de l'évaluation du projet REPERE en 2013