

Une approche quantitative de l'usage des items de feedback en français

Laurent Prévot

Travail en collaboration avec Jan Gorisch, Roxane Bertrand et les autres membres du projet CoFee

Séminaire Talep, 10 février 2015, Marseille



Introduction
ooooo

Données
oooooooo

Etudes Forme Fonction
oooooooooooo

Conclusions et ouvertures
oooo

Introduction

Données

Etudes Forme Fonction

Conclusions et ouvertures

Introduction

ooooo

Données

oooooooo

Etudes Forme Fonction

oooooooooooo

Conclusions et ouvertures

oooo

Outline

Introduction

Données

Etudes Forme Fonction

Conclusions et ouvertures

Contexte

Definition

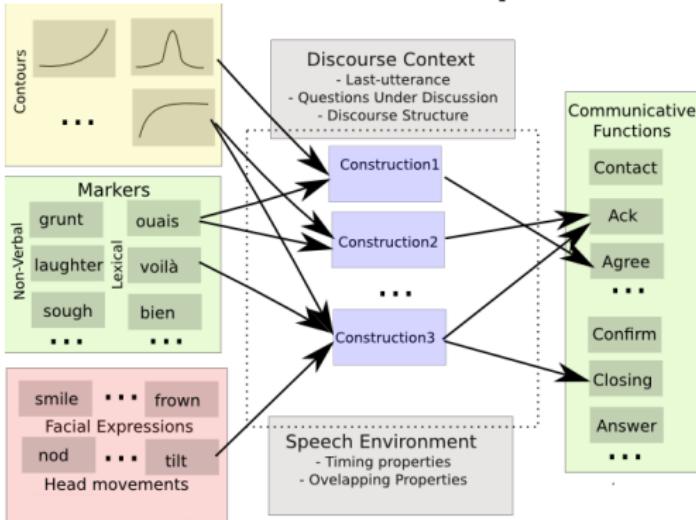
Feedback : Enoncé ayant une fonction évaluative

Back-channel : Enoncé non produit par le locuteur principal

- Enoncés de feedback parmi les plus fréquents en conversation
- Crucial pour les activités communicatives
- Relativement peu étudié d'un point de vue linguistique (pauvreté)
- Pauvreté : Cas intéressant pour l'étude des interfaces entre les domaines linguistiques

Objectifs et modèle

Modèle Schématique



- Modèle ~ Identification automatique des fonctions communicatives

Questions de Recherche

- Compositionalité des domaines vs. nature constructionnelle
- Identification d'une sémantique (sous-spécifiée) pour (certaines) primitives
- Prise en compte précise du contexte (discursif, environnement de parole)
- Impact de la situation de communication et généralisabilité des études formes-fonction

Le feedback

play

Enoncé de Feedback (adapté de [Bunt, 1994])

Un énoncé par lequel son producteur fournit des informations relatives à son traitement des énoncés de son interlocuteur. Cela inclut les informations concernant leur perception, leur interprétation, leur évaluation (accord, surprise,...) ou leur prise en compte effective (satisfaction d'une requête,...).

back-channel (adapté de [Yngve, 1970])

[...] la personne ayant le tour [de parole] et son partenaire sont tous les deux engagés simultanément dans une activité d'écoute et de production de parole. Cela est du à l'existence de ce que j'appelle le *back channel*, sur lequel la personne qui a le tour de parole reçoit des messages courts tels que 'yes', 'uh-huh' sans lâcher ce dernier.

Pourquoi le feedback ?

Central pour les théories de l'interaction langagière

- [Clark, 1996] : un outil central pour le *grounding*
- [Bunt, 1994] : aspect crucial du contrôle du dialogue

Grande fréquence dans le dialogue

- [Stolcke et al., 2000] : ~ 40% des énoncés de SWITCHBOARD peuvent être associé au feedback
- Dans nos données : ~ 34% des IPIUs (MapTask, 36%, CID-DVD, 33%) (11839 / 34355)

→ intéressant méthodologiquement en raison de ses formes limitées

Introduction
ooooo

Données
ooooooooo

Etudes Forme Fonction
oooooooooooo

Conclusions et ouvertures
oooo

Outline

Introduction

Données

Etudes Forme Fonction

Conclusions et ouvertures

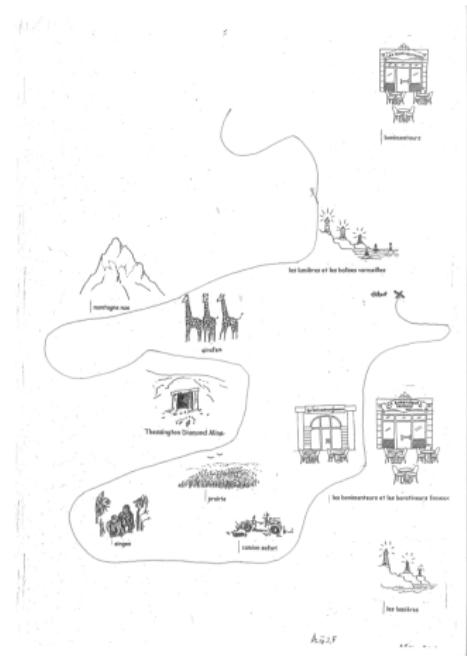
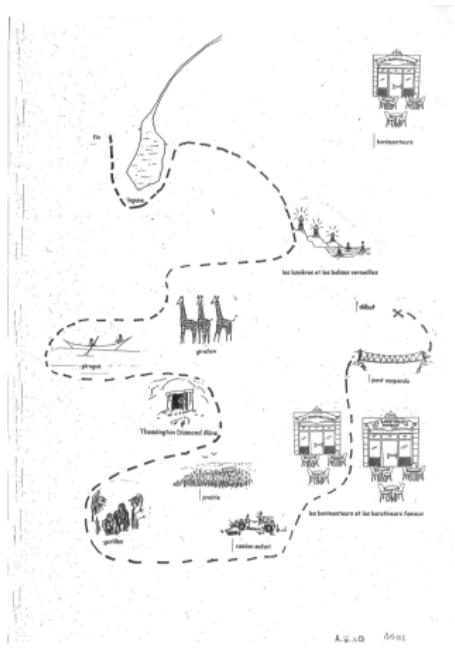
Données Primaires

- **CID** Conversation Interaction Data [Bertrand et al., 2008] ; sldr000027).
- 8 paires de locuteurs ; 1h par paire ; 3 paires avec vidéo.



Données Primaires

- **MTR MapTask Remote** [Bard et al., 2013] ; sldr000732).
- 4 paires de locuteurs ; 8 cartes pour chacune ; 2h audio



Données Primaires

- **MTX MapTask Audio-Visuelle [Gorisch et al., 2014]** ;
sldr000875.
- 5 paires ; 8 cartes chacune ; ~ 2h08 audio et vidéo



Données Primaires

- **DVD** Discussion–negotiation à propos de films (sldr000891).
- 8 paires ; 30 min par paire audio et video.



Unités Inter Pausales (IPU) de Feedback

- IPU feedback Isolée

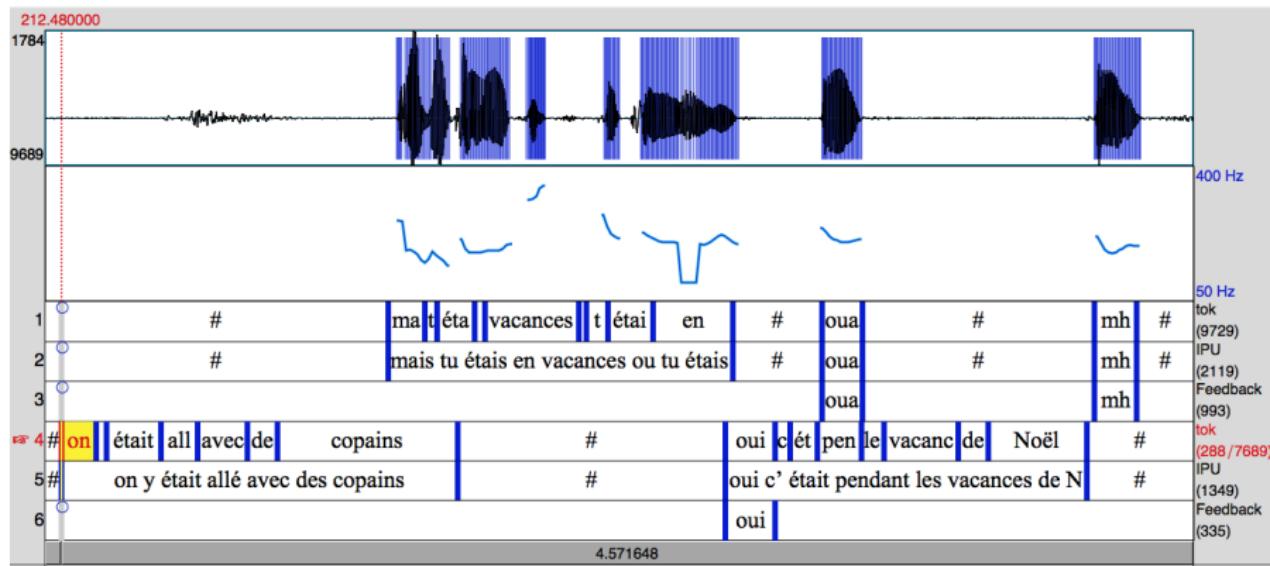
- IPU : Approximation simpliste mais intéressante des énoncés
- Parole bornée par des pauses d'une durée donnée, ici 200ms
- IPU de feedback constituées par un nombre significatif d'items de feedback
 - items de feedback : *mh, ouais, oui, ok, d'accord, voilà, bon, <rire>, non, ah, euh, ben,...*

- Séquence de mots de feedback au début d'une IPU

~~~ 22 000 IPUs candidates

# Exemple IPU candidates - CID

play



## Choix des items

|          |      |
|----------|------|
| ouais    | 2781 |
| mh       | 2321 |
| d'accord | 1082 |
| (rire)   | 920  |
| oui      | 888  |
| euh      | 669  |
| ok       | 632  |
| ah       | 433  |
| voilà    | 360  |
| et       | 360  |

TABLE : Les 10 tokens les plus fréquents parmi les IPUS de 3 tokens ou moins

Suite de la liste : non (319), tu (287), alors (151), bon (150)

# Distribution des IPIs candidates

| corpus   | CID  | DVD  | MTR  | MTX  | Total |
|----------|------|------|------|------|-------|
| complex  | 4740 | 3202 | 2329 | 2104 | 12375 |
| ouais    | 751  | 668  | 531  | 411  | 2361  |
| mh       | 542  | 310  | 73   | 78   | 1003  |
| @        | 571  | 265  | 45   | 79   | 960   |
| *        | 228  | 158  | 297  | 153  | 836   |
| d'accord | 34   | 83   | 195  | 356  | 668   |
| mh+      | 334  | 39   | 245  | 45   | 663   |
| euh      | 220  | 125  | 70   | 135  | 550   |
| oui      | 87   | 64   | 169  | 213  | 533   |
| ok       | 4    | 46   | 129  | 211  | 390   |
| non      | 87   | 90   | 78   | 83   | 338   |
| voilà    | 43   | 40   | 117  | 95   | 295   |
| ah       | 130  | 88   | 18   | 52   | 288   |
| ben      | 63   | 117  | 23   | 21   | 224   |
| ouais+   | 121  | 53   | 24   | 16   | 214   |
| ...      |      |      |      |      |       |
| Total    | 8067 | 5430 | 4405 | 4096 | 21998 |

# Annotation des items de feedback

Pour l'instant seulement pour CID et MTR

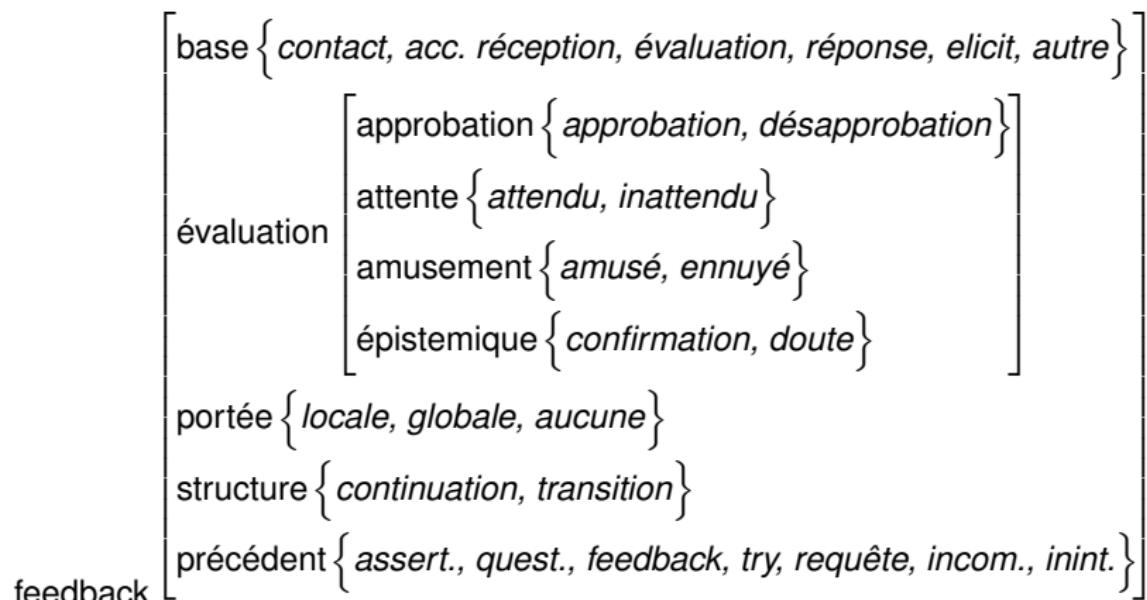
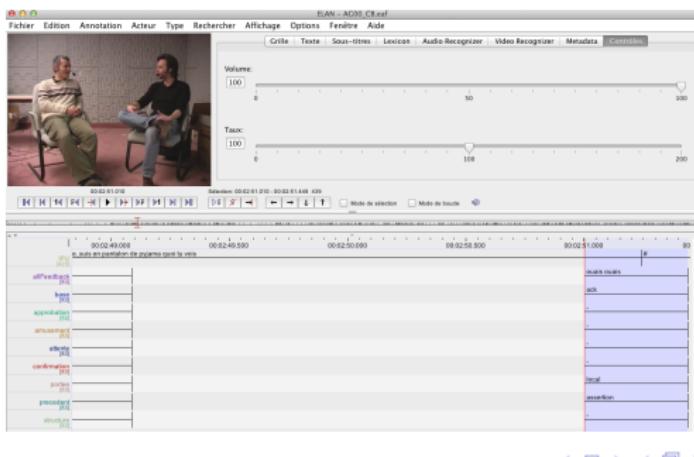


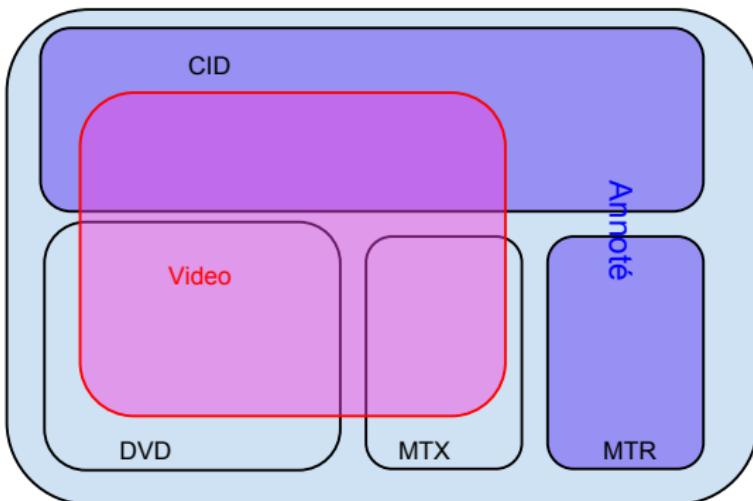
FIGURE : Schéma d'annotation

## Tâche d'annotation

- 3-7 annotateurs semi-naïfs par item
  - $\kappa$  trop bas sauf pour BASE and PRÉCÉDENT et certaines valeurs APPROBATION, INATTENDU, AMUSÉ
  - Référence créée à partir d'un vote entre les annotateurs + seuil (0.66)
    - 4890 instances pour la catégorie BASE
    - 1392 instances pour la catégorie EVALUATION



## Résumé Jeux de données



Introduction  
ooooo

Données  
oooooooo

Etudes Forme Fonction  
oooooooooooo

Conclusions et ouvertures  
oooo

# Outline

Introduction

Données

Etudes Forme Fonction

Conclusions et ouvertures

## Traits Utilisés

- LEX : transcription (complète + simplifiée), longueur, présence de certaines marqueurs (ouais, mhh, ah,...)
- PHON : durée, frontière entre phonème quand possible
- Pos : Pauses avant après, Durée parole avant / après (pour locuteur et interlocuteur), durée normalisée overlap
- CTX : Fonction de l'énoncée précédent annotée
- CTX-AUTO : taille IPU précédente (locuteur et interlocuteur), tokens + bigrammes initiaux et finaux, Nb de "je", de "tu" dans les énoncés
- META : Corpus, Locuteur, Session (MTR), Role
- (ACOU) : Paramètres variés (50+) de Pitch, Intensité, Formants, Apériodicité,...

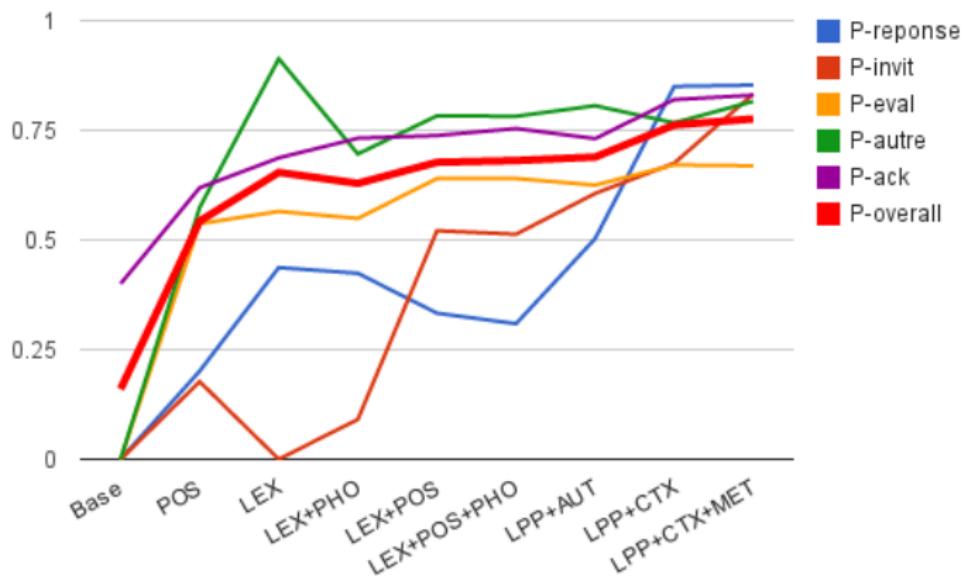
## Fonction à catégoriser

- BASE (4890 instances) : accusé de réception (40%), évaluation (33%), autre (17%), réponse (8%), invitation (2%)
  - EVALUATION (1392 instances) : approbation (42%), amusé (33%), confirmation (18%) , inattendu (7%)

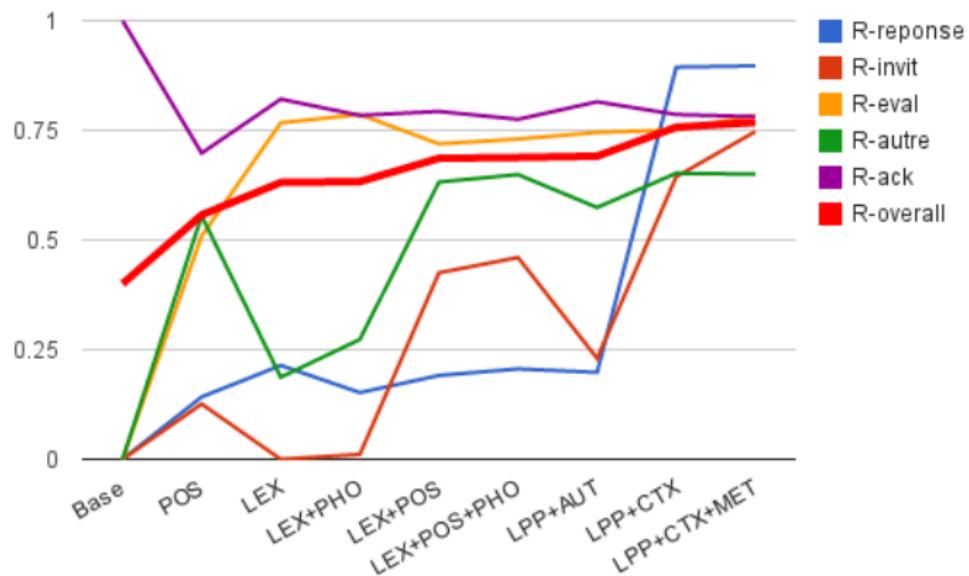
# Méthode de classification

- Exploration d'approches variées avec Weka
- Meilleurs résultats avec les arbres de décisions (J48 dans Weka)
- Fastidieux et impossible de monter en taille
- Explorations avec SCIKIT en cours

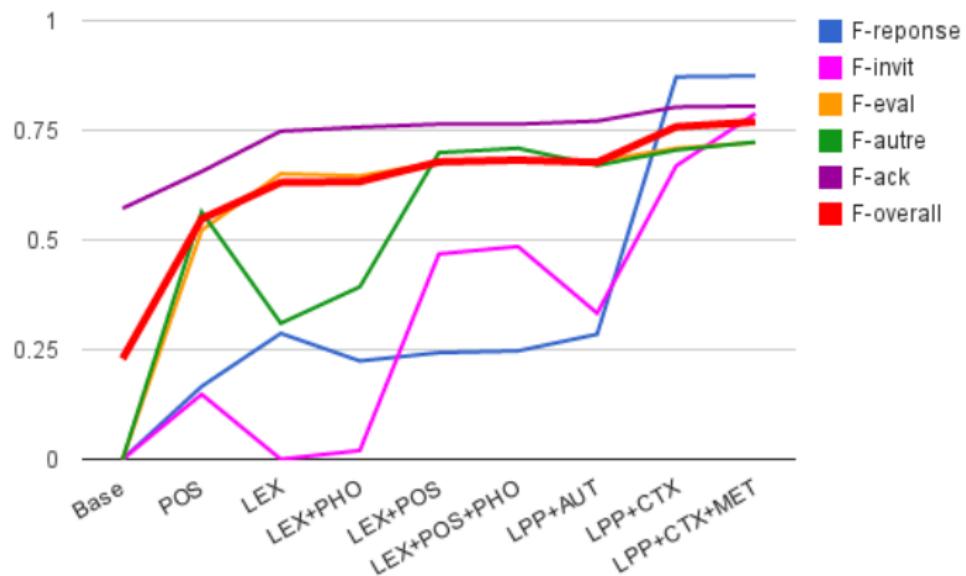
# Precision Classification BASE



# Rappel Classification BASE



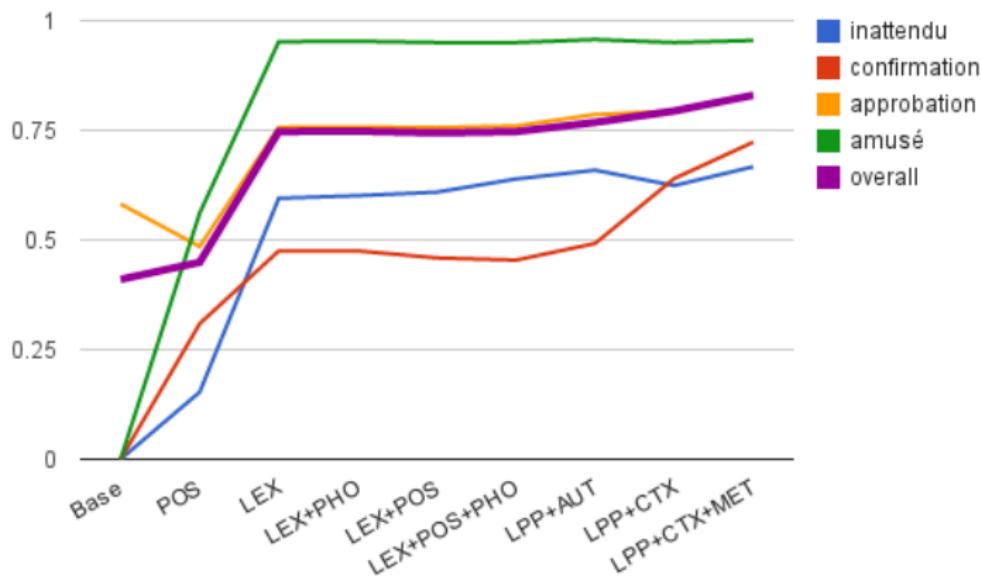
# F-measure Classification BASE



## Résumé Catégorie BASE

- LEX et Pos sont très informatives pour ACK et EVAL
- Pos pour AUTRE : (items complètement isolés)
- CTX fait faire un autre saut (pas Ctx-Auto) pour RÉPONSE et INVIT
- META aide pour le rappel de INVIT (Invit sur-représenté dans MTR)

# F-measure Classif Evaluation



## Résumé Catégorie EVALUATION

- LEX et Pos sont à nouveau les principaux contributeurs
- Pos pour AUTRE : (items complètement isolés)
- CTX utile pour CONFIRMATION (paire essai - confirmation)
- META aide pour toute les catégories (assez fort déséquilibres entre corpus)

# Traits Acoustiques

## Etude préliminaire

| Traits                        | P/R moy. pond. |
|-------------------------------|----------------|
| LEX + POS (opb) + CTX + META  | 0.77 / 0.75    |
| LEX + POS (opb) + CTX         | 0.75 / 0.74    |
| LEX + POS + ACOU + CTX + META | 0.73 / 0.73    |

TABLE : Classification automatique (WEKA) pour la catégorie BASE

- ACOU (automatique) ne joue qu'un rôle marginal
- Mais ~ 10% seulement des cas semblent porter une information intonative pertinente
- Annotation prosodique difficile [D'Imperio et al., 2013]
- Filtrage + passage à la loupe de certaines associations → tests expérimentaux à partir de stimuli des corpus
- Travail plus en profondeur sur le Pitch (extraction Momel)
- Traits acoustiques plus spécifiques / adaptés aux items de feedback ?

## Traits "Méta"

- Meilleur prédicateur de l'alternance oui / ouais  
[Prévot and Gorisch, 2014b] (classe majoritaire par locuteur)
- Idéalement peu / pas d'impact sur nos études formes - fonction
  - OK pour BASE
  - Apport intéressant pour Evaluation mais probablement due à des distributions très déséquilibrées
    - Fonctions ne sont pas réalisées spécifiquement dans un corpus
    - mais certaines fonctions ne sont représentées que dans certains corpus

# Le geste

- Beaucoup de retard pris en raison de problèmes de désynchronisation (Solution dans [Gorisch and Prévot, 2015])
- Pré-segmentation effectuée (MTX, CID)
- Annotation en cours (MTX, CID à peu près terminée)
- Traitements en cours
- Pour MTX 552 des 2998 IPIUs candidates potentiellement associées à un geste (18%)
- Comparaison MTR / MTX : seule différence significative dans les items : *mh* 2 fois plus fréquents dans MTR [Gorisch et al., 2014]

# Cotés données et traitement

- Filtrage des instances avec une classe majoritaire
  - OK pour la question de recherche de CoFee
  - Pas très réaliste dans une optique système de dialogue
  - $\leadsto$  besoin de gérer ces données
- Malgré filtrage, fort bruit résiduel  
[Beigman Klebanov and Beigman, 2009]  $\leadsto$  Besoin d'estimer les meilleurs scores atteignables
- WEKA  $\leadsto$  SCIKIT-LEARN

Introduction  
ooooo

Données  
oooooooo

Etudes Forme Fonction  
oooooooooooo

Conclusions et ouvertures  
oooo

# Outline

Introduction

Données

Etudes Forme Fonction

Conclusions et ouvertures

## Conclusions

- Rôle des modèles formels
    - Nécessaires pour rassembler / intégrer les résultats empiriques dans un modèles riche et cohérent [Prévot and Gorisch, 2014a]
  - Rôle de l'apprentissage automatique
    - Approche plus exploratoire que les méthodes statistiques traditionnelles
    - Plus prédictive, plus proche des applications
  - Place de la variation
    - Besoin d'un modèle robuste aux changements de situations et de locuteurs
    - Rôle des continuums
    - Items non-lexicaux
  - Approche comparative (Mandarin vs. Français)

# Ouvertures

- Données non-transcrites
  - [Neiberg and Truong, 2011] : Modèle spécifique pour l'identification du feedback à partir d'un signal acoustique (et d'un petit lexique)
- Nouveaux corpus
  - Loup-Garou <http://sldr.org/ortolang-000900>
    - Corpus coverbal + physio (avec Thierry Chaminade)
  - Linguistique du dialogue oral spontané

## Ouvertures

- Données non-transcrites
    - [Neiberg and Truong, 2011] : Modèle spécifique pour l'identification du feedback à partir d'un signal acoustique (et d'un petit lexique)
  - Nouveaux corpus
    - Loup-Garou <http://sldr.org/ortolang-000900>
    - Corpus coverbal + physio (avec Thierry Chaminade)



- #### • Linguistique du dialogue oral spontané

# Ouvertures

- Données non-transcrites
  - [Neiberg and Truong, 2011] : Modèle spécifique pour l'identification du feedback à partir d'un signal acoustique (et d'un petit lexique)
- Nouveaux corpus
  - Loup-Garou <http://sldr.org/ortolang-000900>
  - Corpus coverbal + physio (avec Thierry Chaminade)
- Linguistique du dialogue oral spontané
  - Variation
  - Grammaticalisation (Expérience de segmentation à base de maximum d'entropie sur les syllabes, avec Pierre Magistry)

Merci de votre attention !

# References I

-  Bard, E. G., Astésano, C., Turk, A., D'Imperio, M., Nguyen, N., Prévot, L., and Bigi, B. (2013).  
Aix mptask : A (rather) new french resource for prosodic and discourse studies.  
In *Tools and Resources for Analyzing Speech and Prosody (TRASP)*, Aix-en-Provence, France.
-  Beigman Klebanov, B. and Beigman, E. (2009).  
From annotator agreement to noise models.  
*Computational Linguistics*, 35(4) :495–503.
-  Bertrand, R., Blache, P., Espesser, R., Ferré, G., Meunier, C., Priego-Valverde, B., and Rauzy, S. (2008).  
Le cid-corpus of interactional data-annotation et exploitation multimodale de parole conversationnelle.  
*Traitement Automatique des Langues*, 49(3) :1–30.

## References II

-  Bunt, H. (1994).  
Context and dialogue control.  
*Think Quarterly*, 3(1) :19–31.
-  Clark, H. (1996).  
*Using language*.  
Cambridge : Cambridge University Press.
-  D'Imperio, M., Petrone, C., and Prévot, L. (2013).  
Basic prosodic transcription of short french feedback utterances.  
In *Proceedings of Second Advancing Prosodic Transcription for Spoken Language Science and Technology*, Lisbon.
-  Gorisch, J., Astésano, C., Ellen, G. B., Bigi, B., and Prévot, L. (2014).  
Aix map task corpus : The french multimodal corpus of task-oriented dialogue.  
In *Proceedings of the 9th International Conference on Language Resources and Evaluation*.

## References III



Gorisch, J. and Prévot, L. (2015).

Audio synchronisation with a tunnel matrix for time series and dynamic programming.

In *Proceedings of the 40th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brisbane.



Neiberg, D. and Truong, K. P. (2011).

Online detection of vocal listener responses with maximum latency constraints.

In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pages 5836–5839. IEEE.



Prévot, L. and Gorisch, J. (2014a).

Crossing empirical and formal approaches for studying french feedback items.

In *Proceedings of 11th Conference on Logical Engineering and Natural Language Semantics (LENLS)*, Tokyo, Japan.

## References IV

-  Prévot, L. and Gorisch, J. (2014b).  
Verbal feedback : positioning and acoustics of french “ouais” and “oui”.  
In *Proceedings of the 14th Laboratory Phonology Conference*, Tokyo.
-  Stolcke, A., Ries, K., Coccato, N., Shriberg, E., Bates, R., Jurafsky, D., Taylor, P., Martin, R., Ess-Dykema, C., and Meteer, M. (2000).  
Dialogue act modeling for automatic tagging and recognition of conversational speech.  
*Computational linguistics*, 26(3) :339–373.
-  Yngve, V. H. (1970).  
On getting a word in edgewise.  
In *Papers from the sixth regional meeting of the chicago linguistic society*, pages 567–578.