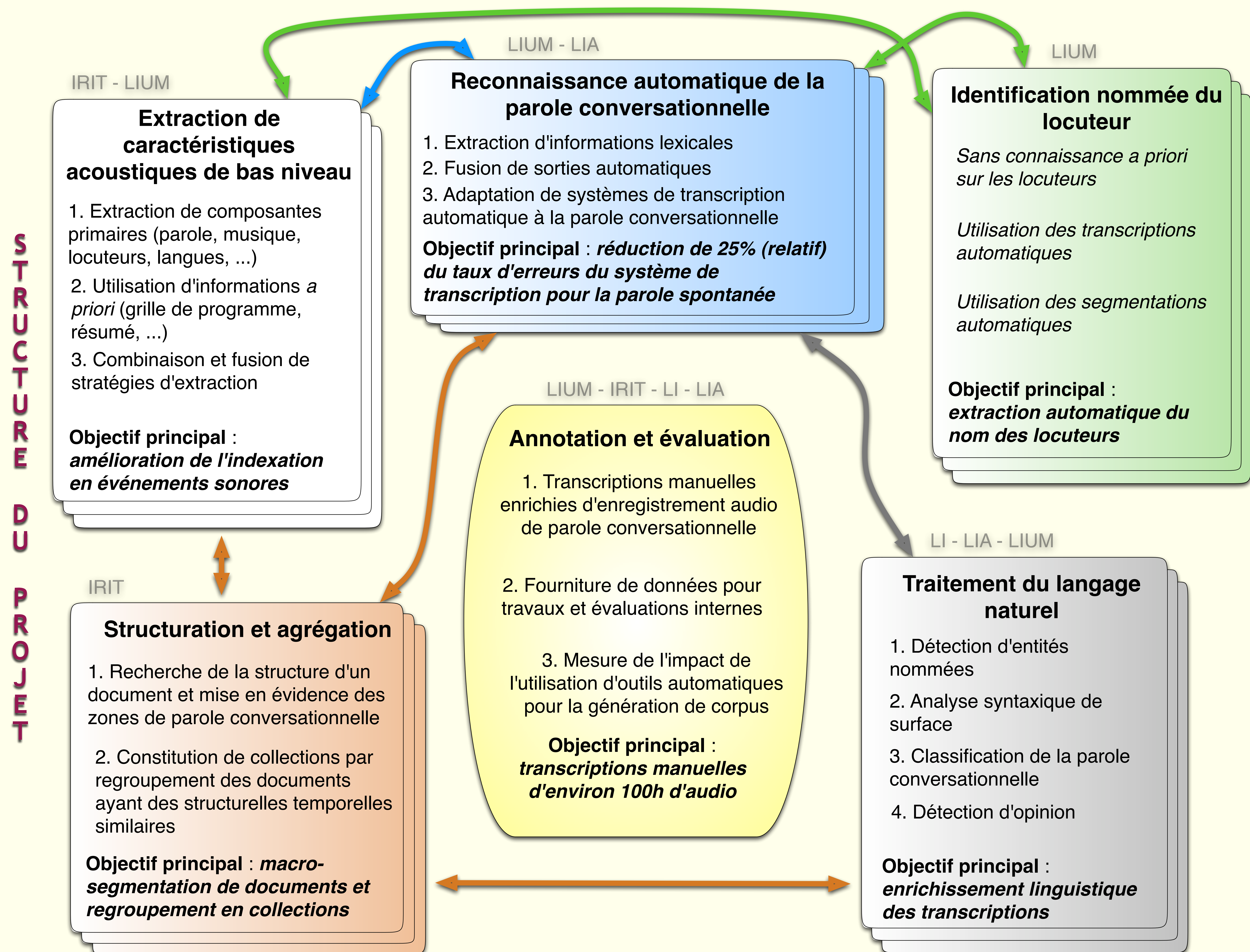


Le projet EPAC concerne le **traitement de données audio non structurées**. Il a pour but de proposer des méthodes d'extraction, d'information et de structuration de documents spécifiques aux données audio, prenant en compte l'ensemble des canaux d'information : segmentation du signal (parole/musique/jingle/...), identification et suivi du locuteur, transcription de parole, détection d'opinion, analyse du discours, interactions conversationnelles, etc. Ces tâches de traitement du signal et de la parole sont en grande partie maîtrisées par les différents partenaires du projet dont la plupart ont participé à la campagne d'évaluation TECHNOLOGUE/EVALDA/ESTER.

Les partenaires disposent d'une base d'environ 2000 heures d'émissions radiophoniques francophones, dont 110h manuellement annotés.

En particulier, le projet EPAC met l'accent sur le traitement de la **parole conversationnelle**.



## RÉSULTATS ET CONSÉQUENCES ATTENDUS

- **Évaluations** internes et évaluations au travers de participations à des **campagnes d'évaluations nationales et internationales** du domaine (de type ESTER ou NIST)
- **Disponibilité sous licence libre des outils** développés durant le projet, spécialisés dans le **traitement de la parole conversationnelle**
- Fourniture d'un nouveau **corpus de transcriptions annotées manuellement de 100h d'émissions radiophoniques**, principalement de la parole conversationnelle
- Fourniture d'un **corpus de transcriptions annotées automatiquement de plus de 1700h d'émissions radiophoniques**
- **Impulsion** dans la communauté 'Parole' francophone
- Développement d'**interactions** entre la communauté de recherche en **linguistique** et la communauté '**traitement automatique** de la parole'