

Consortium Corpus Oraux et Multimodaux de HUMA-NUM (TGIR)

Réunion MULTICOM-CORLI du 15 juin 2016



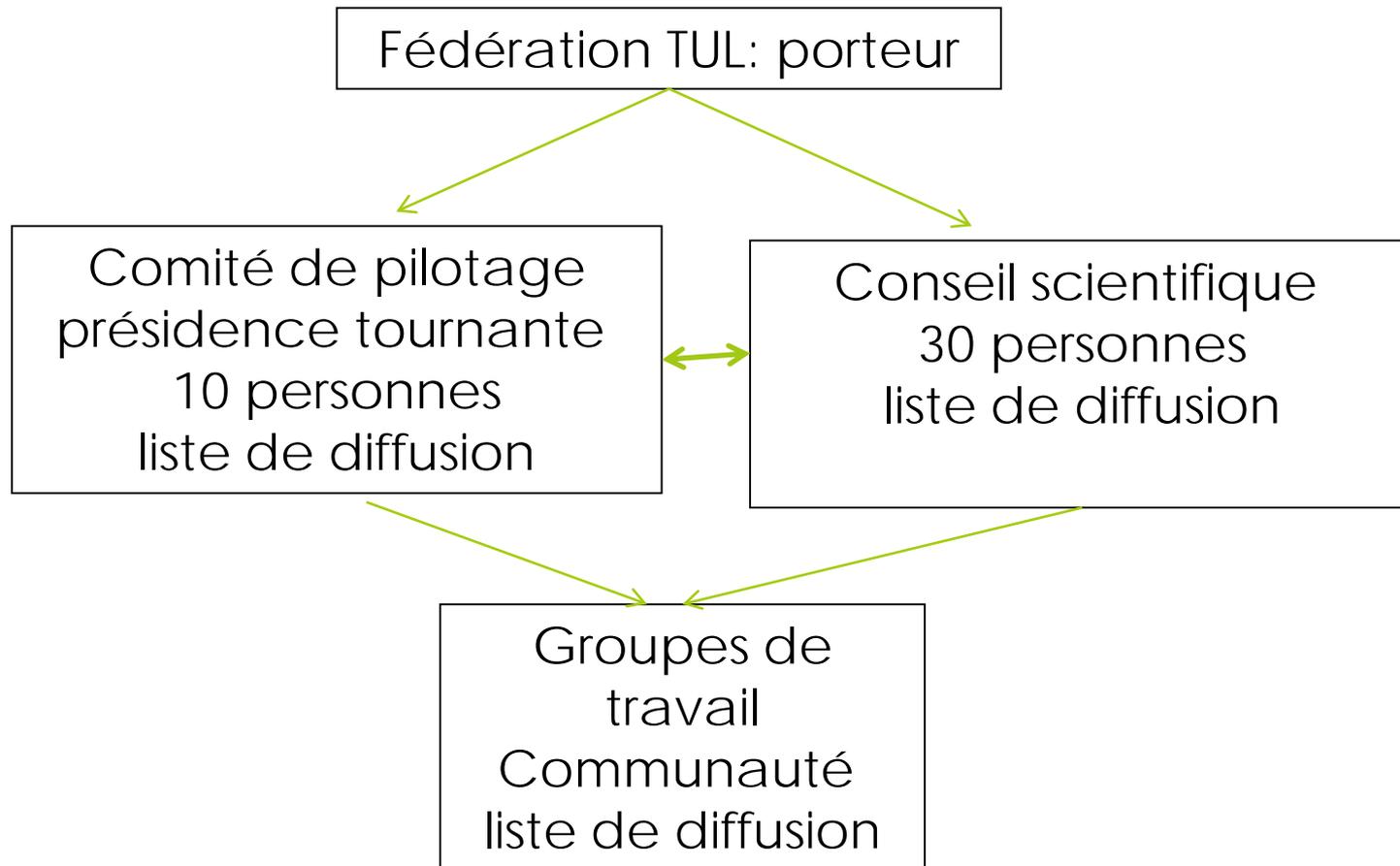
CONSORTIUM CORPUS ORAUX ET MULTIMODAUX

Dominique Boutet

Missions de l'IRCOM

- organiser et accompagner le développement de corpus oraux et multimodaux en linguistique en aidant les chercheurs à s'approprier les outils nécessaires et à développer des standards communs de référence ;
- développer la valorisation, la visibilité et l'accessibilité des fonds existants ;
- améliorer leur mise à disposition, leur mutualisation et leur interopérabilité afin d'intégrer les réseaux internationaux (notamment ERIC-CLARIN) ;
- intégrer la communauté des producteurs et utilisateurs de corpus oraux et multimodaux dans ces pratiques et réflexions.

Organigramme



Les groupes de travail et les actions transversales

- Concertation avec la communauté le 5 octobre 2011 et mise en place de 5 groupes de travail +1 en fin 2013
 - Corpus linguistiques et finalités scientifiques
 - Interopérabilité
 - Corpus multilingues et corpus plurilingues (arrêt en 2012)
 - Multimodalité et modalité visuo-gestuelle
 - Questions juridiques, droits des personnes et des producteurs de corpus
 - Stockage sécurisé et mutualisation des corpus multimodaux (2013 >)

- Actions transversales
 - Supports de communication, aide à la finalisation de corpus et métadonnées

Les formations et soutiens

- Formations : enregistrement et Annotation de corpus : 20 journées de formation IRCOM
 - CLAN (2014 et 2015)
 - ELAN (2013, 2014 et 2015)
 - Vidéo et audio (2012, 2013 et 2015)
 - Formation SPPAS (2015)
 - Formation TEI (2012)

- Soutiens à des actions de formation
 - Ecole thématique de Biarritz « Annotation des données langagières » (septembre 2011)
 - Conférence de Lou Burnard à la BnF ("Encoder l'oral en TEI" - 10 mai 2012)
 - Ecole d'automne d'Aussois sur le document numérique et la gestion de projet (octobre 2012)

Les formations et soutiens

- Soutiens à des journées d'étude
 - Journée d'étude sur *Perspectives croisées sur l'analyse multimodale de corpus oraux*, le 24 mai 2013 (synthèse sur le site)
 - Colloque CREAGEST du 23 novembre 2013 (synthèse sur le site)
 - Colloque Multimodalités, nouvelles perspectives (6 juin 2014)
 - Colloque GESPIN (septembre 2015)
 - Colloque CorpAGEst/CLARe (décembre 2015)
- Aide à la finalisation de corpus multimodaux
 - Subventions accordées en 2013 et 2014 conjointement avec ORTOLANG
 - Aides techniques

Groupe de travail : GT4 Multimodalité

- Groupe Multimodalité et modalité visuo-gestuelle
 - Réflexions autour des besoins émergeant des corpus associés à trois thèmes :
 - l'étude de la multimodalité dans l'utilisation des langues vocales et signées ;
 - l'étude des langues des signes ;
 - l'étude de la gestualité.
 - Identification des unités et individus concernés par ces thèmes sur le plan national.
 - Identification des corpus disponibles sur le plan national
 - Définitions pour le glossaire en relation avec les thèmes du groupe

Groupe de travail : GT4 Multimodalité

- Groupe Multimodalité et modalité visuo-gestuelle
 - Journée d'études internationale *Multimodalité — perspectives interdisciplinaires : nouvelles questions et nouvelles méthodologies* (6 juin 2014)
 - Inventaire des besoins en formation et mise en place d'un programme 2012-2015 de la production de corpus à sa diffusion
 - Formations 2012 (2 jours) :
 - « Prises de vue et de son » ; « Transfert, traitement et montage des données numériques » (25 personnes)
 - Formations 2013 (6 jours) :
 - « Prises de vue et de son » ; « Transfert, traitement et montage des données numériques » (23 personnes)
 - ELAN, Annotations pour les corpus multimodaux (18 débutants et 11 avancés)

Perspectives : GT4 Multimodalité

- Groupe Multimodalité et modalité visuo-gestuelle
 - Autres formations (16 jours)
 - 2014 : « Transcription/Annotation » avec CLAN (Poitiers) ;
« Transcription/Annotation » avec ELAN (Poitiers)
 - 2015 : « Archivage et Diffusion » avec le groupe de travail 6 ; « Prises de vue et de son » ; « Transfert, traitement et montage des données numériques » (déjà donnée en 2012 et 2013), « Analyser les données multi-annotées avec SPPAS » (2 journées)
 - Proposer les formations validées aux écoles doctorales
 - Traduction et mise à disposition des supports pédagogiques auprès de l'ERIC-DARIAH.
- A faire....
 - Continuation et élargissement du recensement :
 - Corpus, besoins, tutoriels, métadonnées, normes de diffusion, manuels de codage...
 - Journée d'étude sur « Nouvelles méthodes de captation pour la constitution de corpus » automne 2015
 - Elaborer des guides de « bonnes pratiques » liées à l'étude de la multimodalité et de la modalité visuo-gestuelle