

**OBSERVER LES LANGAGIERS EN ACTION :
EXEMPLE DU PROJET CRISTAL ET NOUVELLES
COLLABORATIONS**

Unige-UZH Joint Seed Meeting

20.06.2019

aurelie.picton@unige.ch

APERÇU GLOBAL

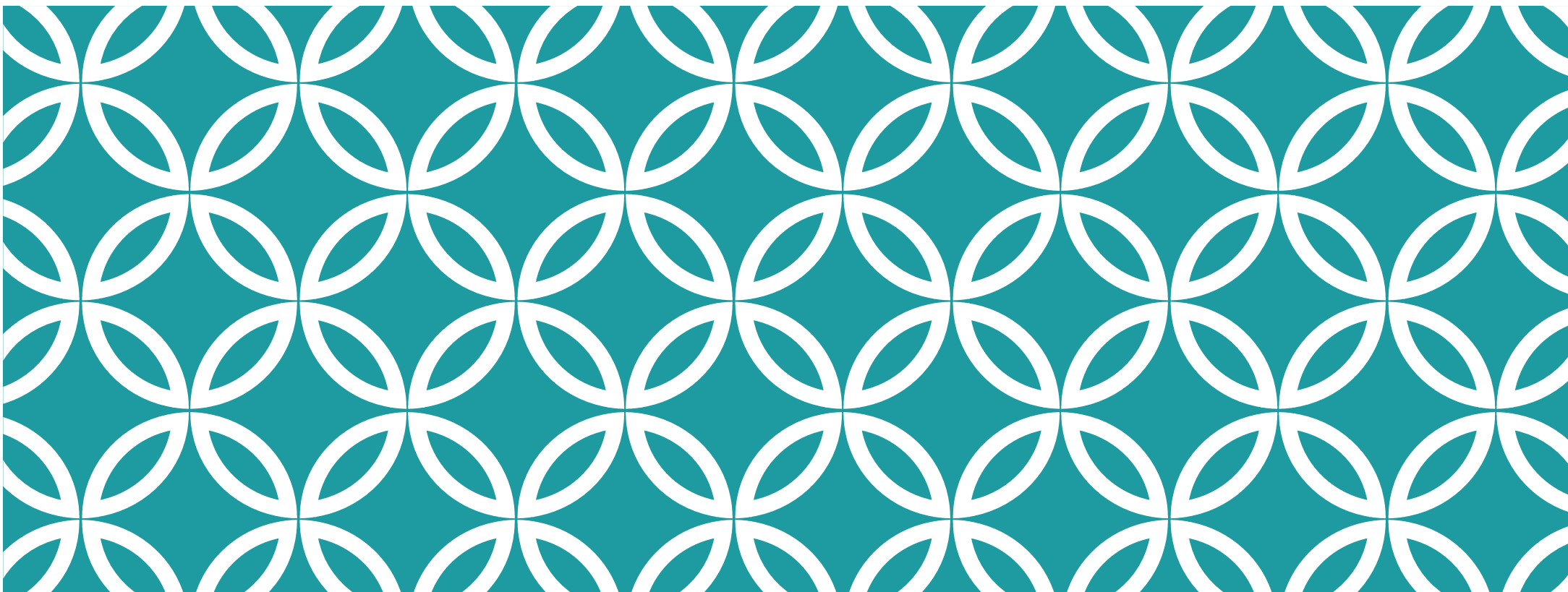
PARTIE 1 – EXEMPLE DU PROJET CRISTAL

1. Contexte du projet
2. Synthèse des expérimentations
3. Exemples de résultats

PARTIE 2 – COLLABORATION AVEC CLLE

1. CLLE-ERSS
2. CLLE-LTC
3. Plateforme CCU

BILAN



EXEMPLE DU PROJET CRISTAL

PARTIE 1



Contextes
Riches en connaissances
pour la TrAduction
terminoLogique

CRISTAL: Knowledge-Rich Contexts for Terminological Translation

Projet de 3, financement Agence Nationale pour la Recherche [ANR-12-CORD-0020]

Début novembre 2012 – fin : janvier 2016

4 partenaires :

- Laboratoire d'informatique de Nantes-Atlantique, France (LINA)
- Cognition, Langue, Langage, Ergonomie (équipe ERSS, UMR5263)
- Faculté de traduction et d'interprétation (Genève)
- Entreprise Lingua et Machina



Contextes
Riches en connaissances
pour la Traduction
terminologique

Projet global

- retrieving semi-automatically Knowledge-rich contexts (KRCs), by retrieving them from bilingual comparable corpora and integrating them into CAT tools (such as translation memories)
- assessing the use and usefulness of this type of data for translators, while translating

Tâche 5

- Proposition d'un protocole d'observation et d'évaluation
- Amélie Josselin-Leray (CLLE-ERSS), Aurélie Picton (FTI), Emmanuel Planas (LINA-UCO)

1. CONTEXTE DU PROJET

Mandat dans le projet

- évaluer finement l'apport de ces contextes riches en connaissances dans une situation de traduction
 - Quels sont les types de contextes les plus utiles aux traductrices et traducteurs ?
 - Quel est le gain obtenu en termes de qualité de traduction par l'apport des contextes riches en connaissances ?
 - Quelle interface mettre en place ?

1. CONTEXTE DU PROJET

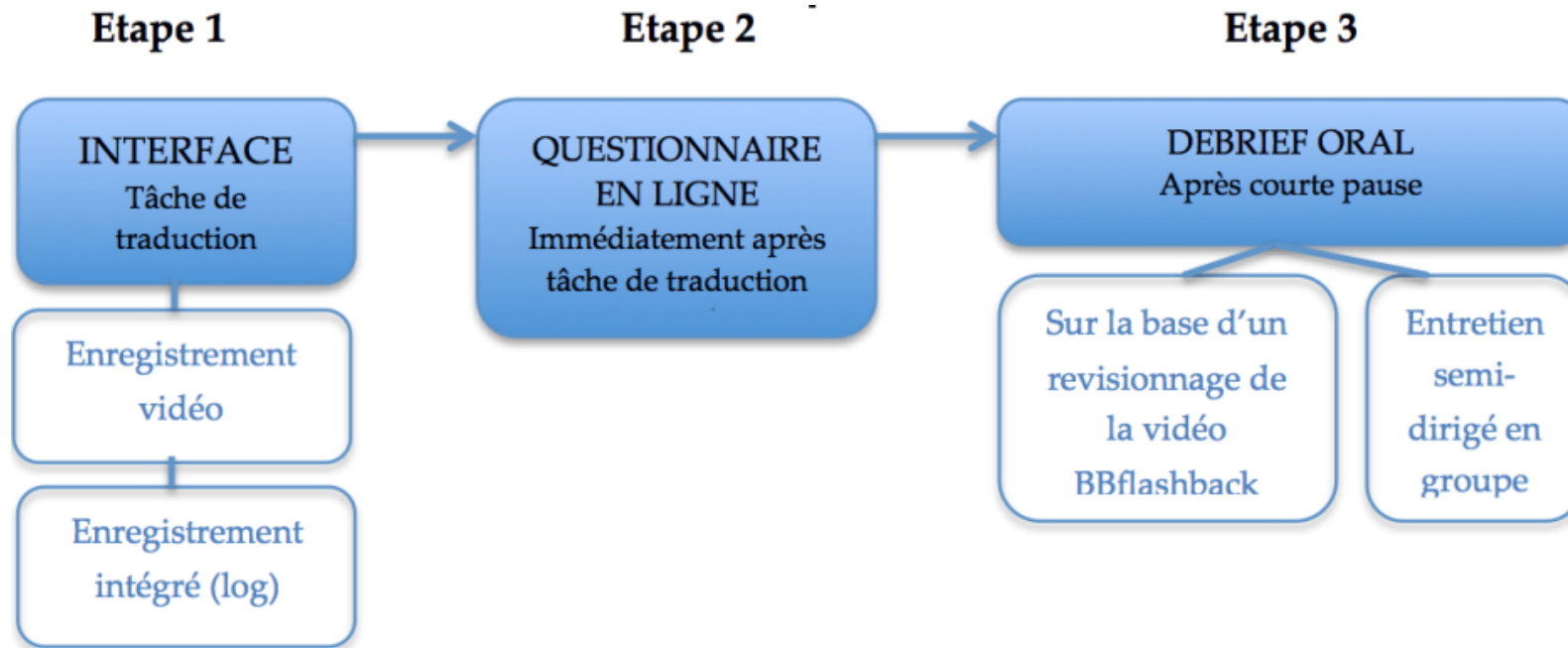
Proposition d'un protocole

1. En phase avec les travaux actuels dans le domaine de la traduction	• Qui prenne en compte l'activité du traducteur dans une tâche complexe de traduction • Triangulation des résultats
2. Écologique	• Qui respecte l'environnement de travail des traductrices et traducteurs
3. Innovant	• Données innovantes • Grand nombre de participant.es

2. SYNTHÈSE DES EXPÉRIMENTATIONS

	Lieu	Date	Participants	Objectifs
		2013-2014		Proposition d'un protocole
Pré-test	Genève	2 décembre 2013		Ajustement du protocole général
Expérimentation sur Argos	Angers	10 mars 2014	Étudiant.es de M2 (42)	Validation de l'interface prototype et recueil de premiers résultats
	Toulouse	17 mars 2014		
		2014-2015		Adaptation et transfert sur Libellex
Expérimentation sur Libellex	Angers	24 mars 2015	11 étudiant.es M2	Évaluation du transfert prototype -> Libellex
		13 avril 2015	8 professionnel.les	Évaluation de l'intégration des CRC (LINA/ERSS) Validation auprès de professionnels

2. PROTOCOLE GLOBAL



2.1

File ARGOS 1.3 CRISAL

CINDER CONES

Basalt cinder cones are the most common kind of volcano worldwide. They are also some of the smallest volcanoes. A typical eruption goes through two stages. The first is called the *fountaining stage*. When the basalt magma first breaks out at the surface, the dissolved gases bubble off vigorously enough to carry blobs of magma into the air with them. The blobs may rise up 2,000 feet or more. During their flight through the air, the blobs cool to

LES CONES DE SCORIES

Les cônes de scories basaltiques représentent le type le plus répandu de volcan dans le monde. Une éruption classique s'effectue en deux phases.

Contextes Dictionnaire Spécialisé Robert et Collins GDT Termium

Source Cible

cone

Ok?	CRC
	que nous pouvons observer. Elle présente, suivant les cas, des cratères, des cônes, des dômes, des coulées, des dépôts de cendres, plus ou moins nombreux et de dimensions très variables.
***	Il y a au contraire une forte probabilité pour que s'édifie un cône de scories basaltiques, accompagné d'une coulée de lave peu étendue : ce serait plutôt une aubaine pour le tourisme.
	La grande fissure se ferma partiellement et l'émission de laves se localisa en quelques points le long de la fissure : il se forma là de petits cônes de cendres et de

Note Pad: utilisez-moi pour vos notes !!

BILAI

Libellex 4 - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages ScrapBook Outils Aide

Libellex x Libellex 4 x

Pas de proxy

demo.libellex.fr/Lbx/TT?sjobId=12062

Infos aLIRE WebMails TAO-2015 Sites IPLV LINA CETIM Sites EP BÉPO Confs Now CRISTAL Doodles Google 日本語

Affichage des balises Remplacement automatique

100%MT

SourceText.html

1	CINDER CONES	CÔNES DE SCORIES	✓
en-GB — 2/20 Ctrl+0 Basalt cinder cones are the most common kind of volcano worldwide.		fr-FR Basalt calcinées cones sont les plus commune genre de volcan dans le monde entier.	
3	They are also some of the smallest volcanoes.		✗
4	A typical eruption goes through two stages.	Typique éruption va par deux étapes.	✓

en-GB → fr-FR

cinder

Mémoire de traduction Résultats de la recherche

Ctrl+1 Basalt calcinées cones sont les plus commune genre de volcan dans le monde entier. MT

Ctrl+2 — —

Ctrl+3 — —

Ctrl+4 — —

Ctrl+5 — —

CRC

Puu Oo is a **cinder** and spatter cone created by fountaining eruptions from 1983-1986.

The tephra accumulates near the central vent and builds a **cinder cone**.

Instead, the volcanoes might erupt lava and then, tens of thousands of years later, create a **cinder cone**.

On the floor of the channel sprawled a field of ice and snow blackened by **cinders** and split by crevasses that gaped white in the enveloping mists.

As we crunch along the floor of black

2. PROTOCOLE

Objet			Durée approx.
	1. Phase d'entraînement	Prise en main et simulation des étapes demandées	15'
Étape 1	2. Expérimentation	a. Lancement interface et BBFlashback	1h-1h15
		b. Tâche de traduction	
		c. Arrêt interface, enregistrement : sauvegarde et dépôt	
Étape 2	3. Questionnaire en ligne	Réponse au questionnaire, individuellement	20'-25'
Étape 3	4. Débriefing	groupes généraux + 3 sur vidéos	30'-45'

2. PROTOCOLE GLOBAL

Etape 1	Etape 2	Etape 3
- choix des CRC et à quel moment de la tâche de traduction - choix des équivalents, traduction finale - consultation des ressources (non CRC) et à quels moments de la tâche - durée de la traduction - notes prises - enregistrement vidéo intégral - changements de sous-fenêtres	- informations générales - retours sur la tâche ○ principales difficultés de traduction ○ estimation de la difficulté de traduction des termes ○ jugement sur qualités et limite des CRC - appréciation de l'interface et du contexte global	- précision sur des choix de CRC - précision sur des choix de termes - précisions des réponses (fermées) données dans questionnaire en ligne - explicitation d'actions enregistrées

2. PROTOCOLE GLOBAL

Etape 1	Etape 2	Etape 3
Données quantitatives et qualitatives	Données quantitatives et commentaires	Données qualitatives
- enregistrements vidéo - enregistrement de l'activité des touches et de l'écran (key-log de BbFlashback et log de l'interface) - enregistrement du changement de sous-fenêtre - post-traitement des logs (tableaux 1a à 2b) - fichier texte de la traduction finale	- réponses à questionnaire : graphiques, statistiques, données exportables vers un tableur (ex. : Excel) ou un logiciel d'analyse des données (ex. : SPSS)	- enregistrements audio/vidéo des débriefs - prise de notes lors des débriefs

3. RÉSULTATS : VALIDATION

1. Protocole général

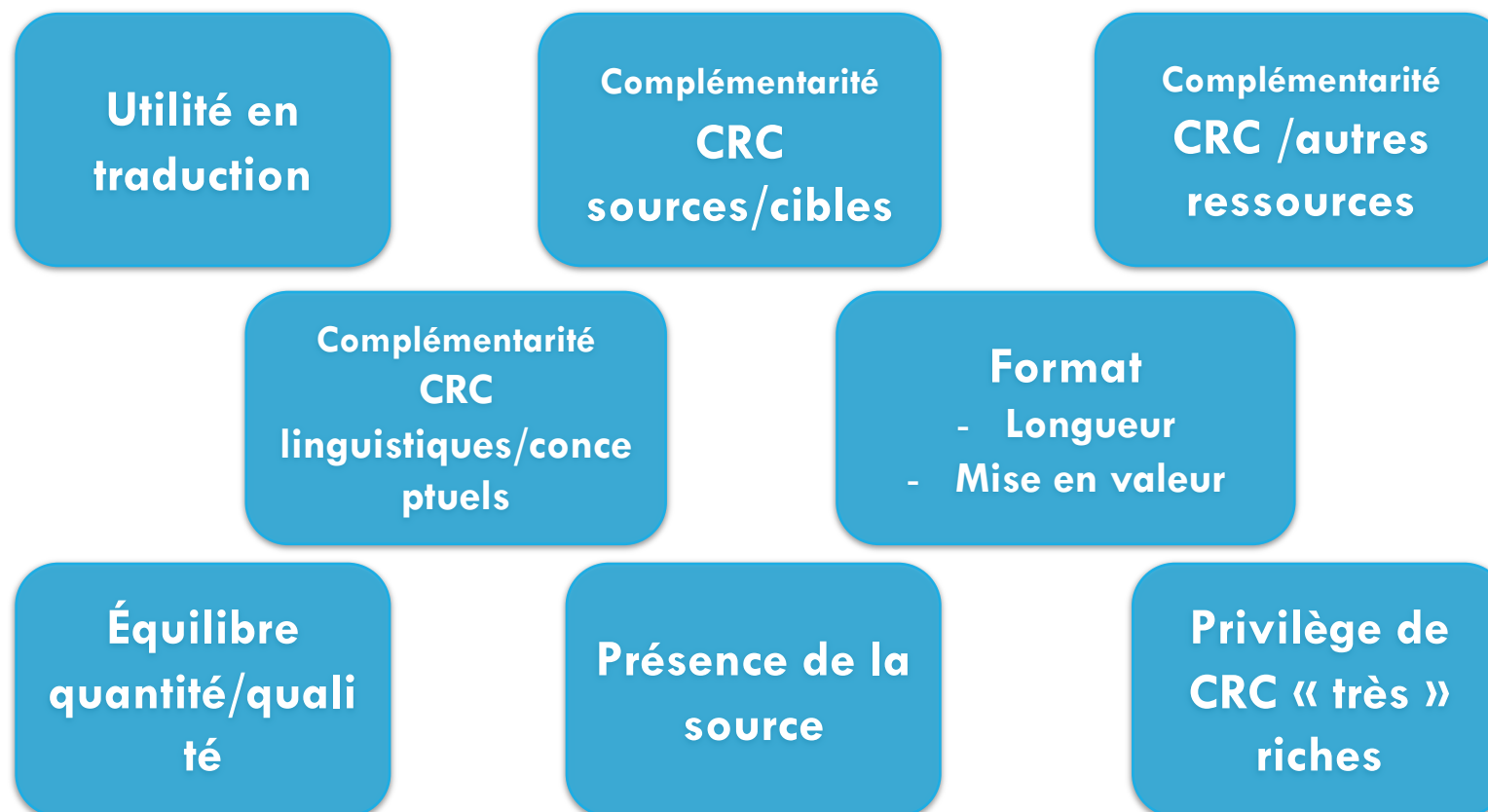
- Type de tâche : traduction de texte court en contexte écologique
- Déroulé de l'expérimentation
 - Timing
 - Sources choisies
 - Complémentarité logs, vidéos, questionnaires, entretiens

3. RÉSULTATS : VALIDATION

2. Ergonomie de l'interface

- Double fenêtre
- Intégration de toutes les données
- Plusieurs types de données
- Raccourcis claviers
- Personnalisation

3. RESULTATS : DÉFINITION DE CRC & TRADUCTION



3. RESULTATS : EN SUSPENS

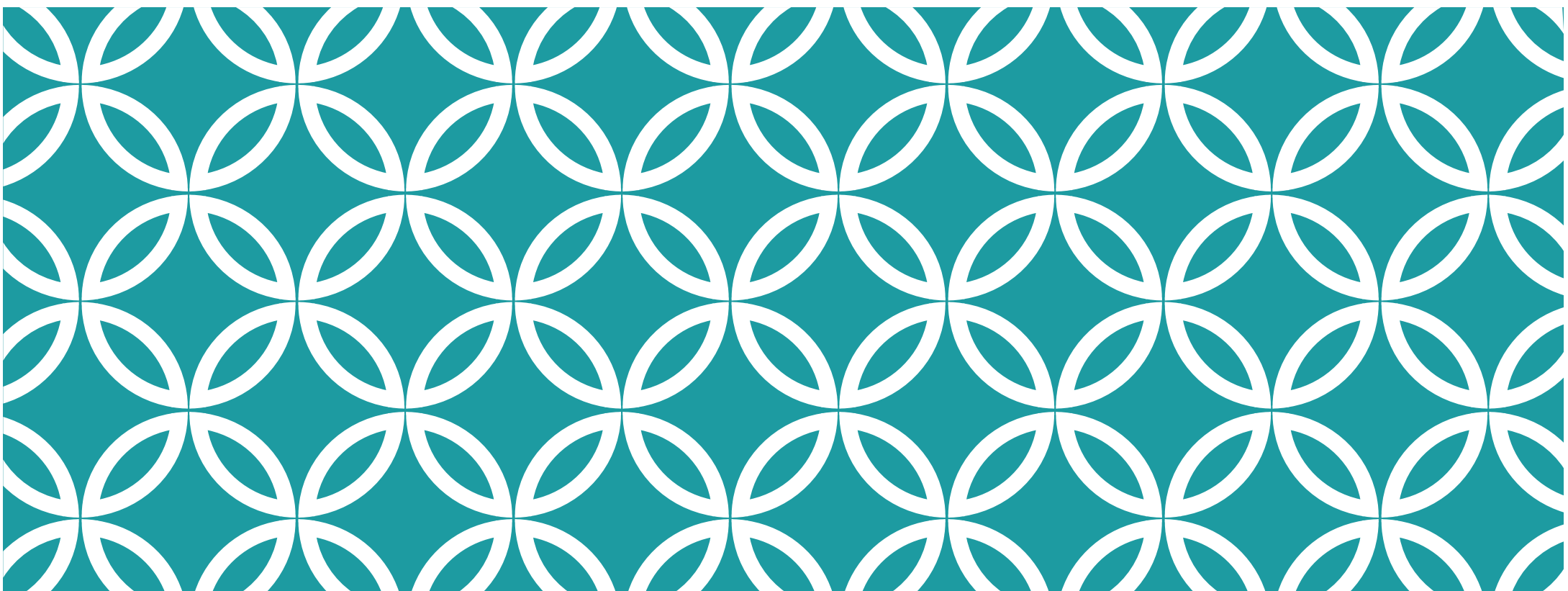
1. Différences professionnel.les vs. étudiant.es

2. CRC

- CRC extraits automatiquement à affiner et valider (Libellex)
- Utilisation et caractérisation des CRC
 - Résultats à continuer d'explorer encore (toutes expé.)
 - Erreurs de logs et enregistrements (Libellex)
 - Autres types de CRC non testés (cause-effet)
 - Extraction et caractérisation automatique des CRC à affiner

PERSPECTIVES IDENTIFIÉES DANS CET AXE

- Comparaison de différents publics de traductrices et traducteurs (expert.e-non expert.e)
- Potentiel **pédagogique** du revisionnage vidéo (cf. travaux à Zurich de G. Massey) :
 - « Cognitive research on translation competence appears to have resulted, at least in the literature concerned with translation pedagogy, in a growing shift towards process-orientation in the training of translators. »
- Exploration de l'utilité des contextes “très riches” ou “high-density knowledge rich contexts) (León-Araúz & Reimerink 2019)
- Intégration de l'eye-tracking (oculométrie)



COLLABORATION AVEC CLLE

Partie 2

CLLE : PRÉSENTATION GÉNÉRALE

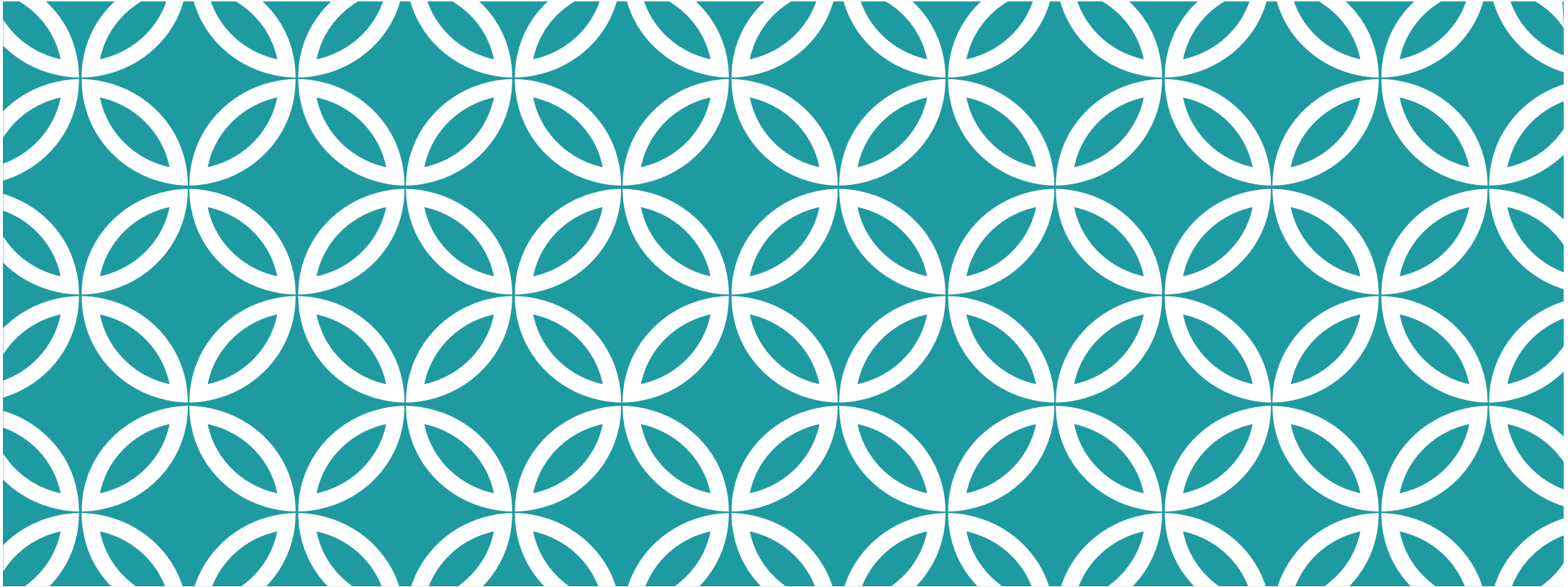
Cognition, Langues, Langage, Ergonomie

Laboratoire **pluridisciplinaire** et polyvalent en sciences cognitives, qui mène des recherches qui quadrillent un vaste périmètre des domaines de la **linguistique** et de la **psychologie**.

Les recherches menées dans le laboratoire s'articulent autour de deux grandes équipes de recherche comportant chacune quatre axes.

1) L'équipe **ERSS** (*Equipe de Recherche en Syntaxe et Sémantique*) travaille plus particulièrement des thématiques relevant de la linguistique (phonologie, morphologie, syntaxe, sémantique, discours, TAL, didactique des langues, psycholinguistique),

2) L'équipe **LTC** (*Laboratoire Travail et Cognition*) couvre de nombreux champs de la psychologie (ergonomie cognitive, cognition sociale, développement du langage et de la communication, neurosciences).



2.1 CLLE-ERSS

1. AXE S'CALADIS – SÉMANTIQUE ET CORPUS SPÉCIALISÉS

Amélie JOSSELIN-LERAY, MCF, Département des Etudes du Monde Anglophone

- Co-directrice du CeTIM (Centre de Traduction, Interprétation et Médiation Linguistique), UFR Langues
- Visiting Researcher à Dublin City University (janv-juillet 2019), équipe SALIS :
 - Approfondissement des connaissances dans les champs suivants : *études empiriques, eye-tracking, traduction automatique, post-édition, interaction homme-machine*
- Co-responsable du Master TIM (appartenant au réseau EMT)
- Présence de cours de **sous-titrage** en Master 1 et Master 2
 - anglais / **allemand** / français figurant dans les combinaisons de langues possibles
 - Sous-titrage **malentendant** (parcours spécialisé anglais / français / **LSF**)
 - **Elise Leroy (MCF LSF)**

1. AXE S'CALADIS – SÉMANTIQUE ET CORPUS SPÉCIALISÉS

Amélie JOSSELIN-LERAY

- Thèmes principaux de recherche
 - Linguistique de corpus
 - Lexicographie
 - Terminologie
 - Technologies de la traduction
 - Dimension empirique : Utilisation des dictionnaires et des bases de données terminologiques pendant le processus de traduction; adaptation des outils aux besoins de l'utilisateur (en particulier, le traducteur)
 - Travail en cours sur les informations contextuelles dans les outils à la disposition des langagiers (dictionnaires, bases de données terminologiques, logiciels de TAO)

1. AXE S'CALADIS – SÉMANTIQUE ET CORPUS SPÉCIALISÉS

Amélie JOSSELIN-LERAY

Axes et intérêts dans le projet :

- Poste de travail des sous-titreur/s/titreuses :
 - interaction et outils (logiciels de sous-titrage, outils de TAO, traduction automatique)
 - Adaptation de l'outil aux besoins des professionnel-le
 - Recherche documentaire (dictionnaires, bases de données terminologiques, Internet) pendant le processus de sous-titrage
- Problématique de la prise en compte du contexte dans le cas du sous-titrage
- Analyse de stratégies de résolution de problèmes de traduction

- **Méthodologie** : études empiriques
 - Observation des sous-titreur/s/titreuses « en train de sous-titrer »
 - Conditions écologiques
 - Enregistrement de l'activité à l'écran, verbalisation post-tâche de sous-titrage, questionnaires
 - Recours à l'oculométrie / eye-tracking (peu de travaux existants sur la question; ils concernent surtout la réception des sous-titres cf. Hvelplund 2017)

2. AXE CARTEL

Cécile FABRE, PR Sciences du Langage; Mai HO-DAC, MCF Sciences du Langage

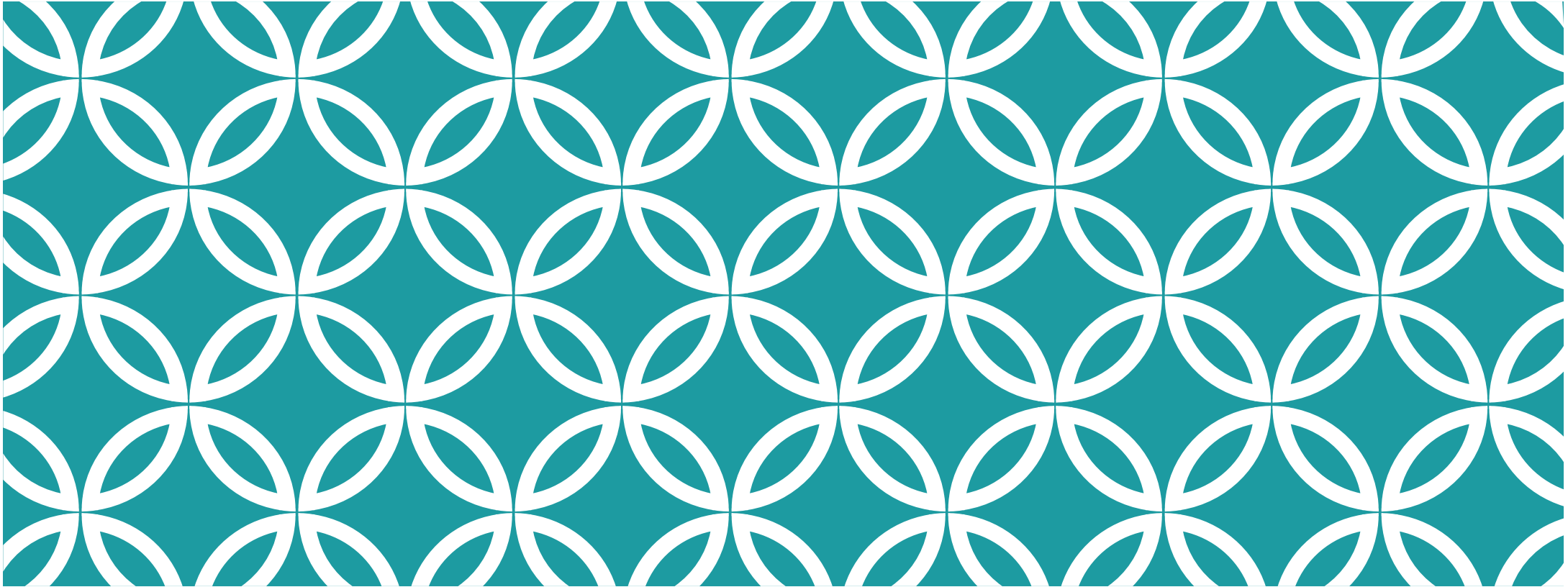
(1) Une thématique concerne la **caractérisation et la classification automatique de corpus**

- Travail essentiellement axé sur l'écrit, mais début de recherche sur les corpus oraux
 - Mai Ho-Dac
 - Projet ORFEO :
 - <https://www.projet-orfeo.fr/> (Outils et Recherches sur le Français Ecrit et oral)
- **Intérêt pour le projet** : caractérisation automatique d'un corpus de sous-titres en termes d'indicateurs linguistiques variés (par comparaison avec d'autres corpus textuels)

2. AXE CARTEL

(2) Travail **exploratoire** de modeste envergure sur **les sous-titres** (partenariat entreprise Archean Labs; projet tutoré de Master)

- *Objectifs du projet* : utilisation des sous-titres comme source de matériel linguistique authentique pour l'apprentissage des langues étrangères (FLE) : exemples de français ordinaire et spontané pour illustration de points grammaticaux
- *Matériel* : corpus de sous-titres OPUS openSubtitles (corpus de 3 M de mots (455 films))
 - Prétraitement des données : nettoyage et annotation automatiques (parsing avec l'analyseur Talismane)
 - Recensement des points grammaticaux à extraire (à partir du Cadre Européen Commun de Référence pour les Langues)
 - Extraction à l'aide de patrons morphosyntaxiques. Traitement des temps verbaux, des connecteurs, de la négation, etc.
 - Evaluation des performances de l'extracteur sur la base d'annotations manuelles de 316 phrases uniquement sur les temps verbaux
 - Première évaluation qualitative des difficultés liées au traitement automatique de ces données bruitées



2.2 CLLE-LTC

AXE 1 : ERGONOMIE COGNITIVE

Céline LEMERCIER (MCF en psychologie cognitive et ergonomie)

- Processus attentionnels dans le traitements de sources multiples

Loïc CAROUX (MCF en psychologie cognitive et ergonomie)

- Interface Homme Machine
- Travail en interdisciplinarité

Radouane EL YACOUBI (MCF)

- Multilinguisme

AXE 1 : ERGONOMIE COGNITIVE

- Publications pertinentes : sélection

Lemercier, C., Tardieu, J., Gaillard, P., Paubel, P. 2017 Regarder sans voir : impact de conversations téléphoniques courtes sur l'attention visuelle en conduite simulée, Looked but failed to see: Effect of short phone conversations on visual attention in simulated driving Recherche Transports Sécurité 2016 01-02 p. 59-65

Terrier, P., Lemercier, C., Quaireau, C. 2005 Decision processes and recognition memory in an information retrieval task: there is more than one memory trace. PSYCHE: An Interdisciplinary Journal of Research on Consciousness 11 (1) p. 1

Boyer, S., Paubel, P., Ruiz, R., El Yagoubi, R., daurat, A. 2018 Human Voice as a Measure of Mental Load Level Journal of Speech, Language, and Hearing Research 6111 p. 2722-2734

Caroux, L. 2016 Propriétés visuelles de bas niveau des interfaces de jeu vidéo et expertise des joueurs, Low-level visual features of video game interfaces and player expertise Actes de la 28ième conférence francophone sur l'Interaction Homme-Machine Fribourg p. 126-133

AXE 2 : CONTEXTE SOCIAL ET RÉGULATION DE LA COGNITION

Claudine MELAN (PR en neurosciences)

- Charge cognitive
- Travail posté, charge de travail, vigilance, cognition et état fonctionnel

- Publications pertinentes : sélection

Edith, G., Mélan, C. 2013 INTERACTIONS BETWEEN COGNITIVE LOAD FACTORS ON WORKING MEMORY PERFORMANCE IN LABORATORY AND FIELD STUDIES
Working memory: developmental differences, component processes and improvement mechanisms

CASCINO, N., Mélan, C., Galy, E. 2016 Étude des perceptions de l'environnement de travail au cours du poste chez des travailleurs postés : intérêt d'une évaluation subjective et temporellement située de la charge de travail
Le travail humain 79(2) p. 97-124

AXE 3 : RECHERCHER, APPRENDRE, DÉCIDER

Franck AMADIEU (PR en psychologie cognitive et ergonomie)

Julie LEMARIE (MCF HDR en psychologie)

- Traitement de sources multiples et charge cognitive
- Recherches analysant les activités des apprenants comme les comportements de lecture, la navigation dans les hypertextes, la construction de cartes de concepts, ou encore les mouvements oculaires, afin d'identifier les processus cognitifs sous-jacents. Ces recherches s'inscrivent dans le champ de la psychologie cognitive et de l'ergonomie des Interactions Homme-Machine en contribuant à l'amélioration des outils et des documents pour l'apprentissage et la recherche d'information (Amadiéu)
- Processus mentaux impliqués dans la compréhension et l'apprentissage à partir de documents, multimédia ou non. (Lemarié)

AXE 3 : RECHERCHER, APPRENDRE, DÉCIDER

Florence BARA (MCF, psychologie des apprentissages et du développement)

- Production de l'écrit
- Compréhension des processus cognitifs mis en jeu dans les apprentissages scolaires (mémoire, langage, geste moteur)
- Utilisation de méthodes quantitatives (analyses statistiques) en complément de méthodes qualitatives (entretiens, vidéos)

Aline CHEVALIER (PR en psychologie cognitive et ergonomique)

- Recherche d'information dans des sources multiples
- Recherche d'information au sein de documents électroniques, e.g. Internet, (rôle du vieillissement, de l'expertise et de la complexité des tâches à effectuer)
- Conception de documents électroniques. Quelles activités cognitives et difficultés cognitives les concepteurs rencontrent-ils

PUBLICATIONS PERTINENTES : SÉLECTION

- Publications pertinentes : sélection

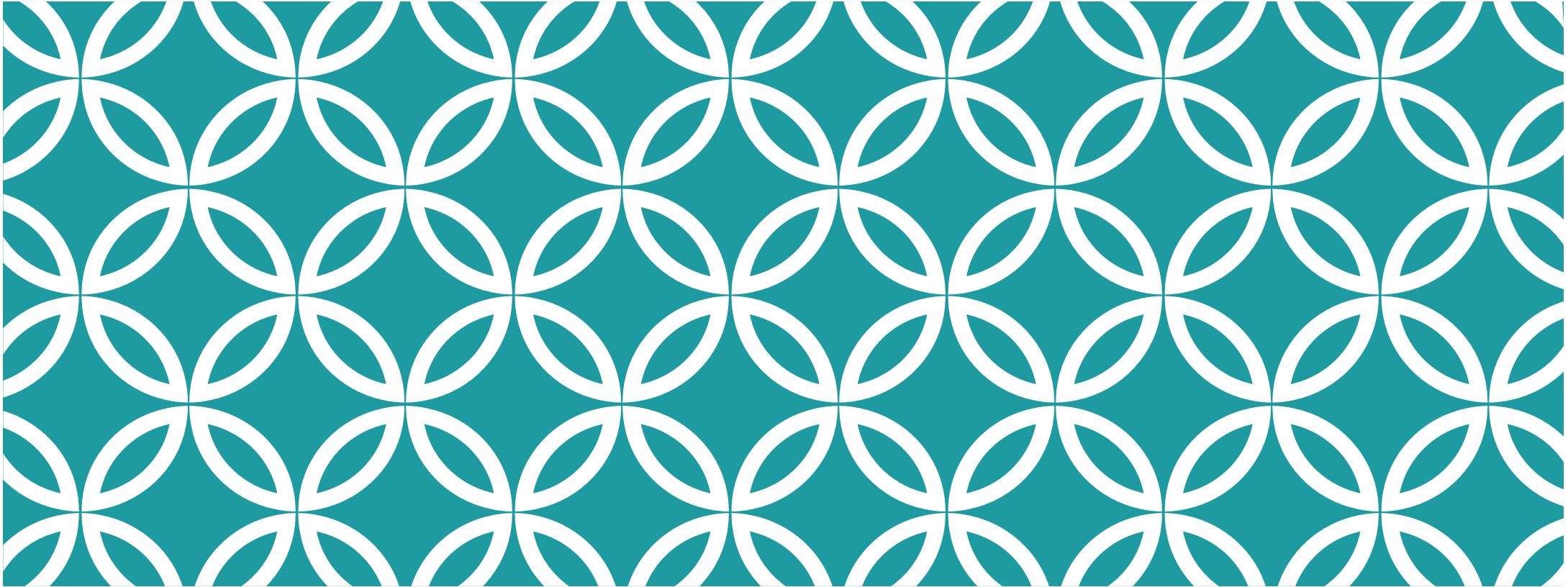
Sanchiz, M., **Amadiou**, F., Wai-Tat, F., **Chevalier**, A. 2019 Does pre-activating domain knowledge foster elaborated online information search strategies? *Applied Ergonomics* 75p. 201-213

Monchaux, S., **Amadiou**, F., **Chevalier**, A., Mariné, C. 2015 Query strategies during information searching: Effects of prior domain knowledge and complexity of the information problems to be solved *Information processing & management* 515p. 557 - 569

Lombard, J., **Amadiou**, F., Bråten, I., van de Leemput, C. 2018 Reading Multiple Documents on Tablet: Effects of Applications and Strategic Guidance on Performance and Acceptance *LCT 2018: Learning and Collaboration Technologies. Design, Development and Technological Innovation* p. 157-169 Springer

Amadiou, F., Salmerón, L., Cegarra, J., Paubel, P., **Lemarié**, J., **Chevalier**, A. 2015 Learning from Concept Mapping and Hypertext: **An Eye Tracking Study**. *Journal of Educational Technology and Society* 184p. 100-112

Lemarié, J., Eyrolle, H., Cellier, J. 2008 The segmented presentation of visually structured texts: Effects on text comprehension *Computers in Human Behavior* 243p. 888 - 902



2.3 LA PLATEFORME CCU

LA PLATEFORME CCU

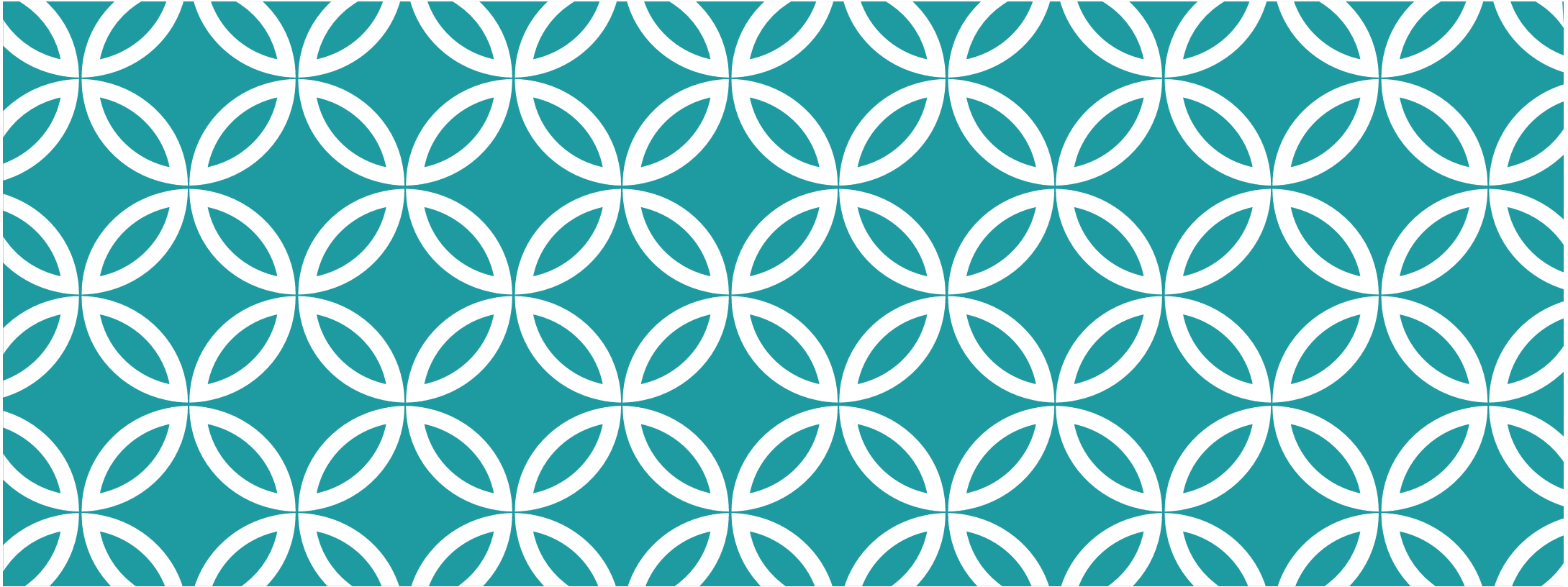
Les chercheurs du laboratoire CLLE ont accès à la plateforme technologique **Cognition, Comportements et Usages** ([CCU](#)) dédiée à l'étude du comportement cognitif humain dans diverses situations à travers les réponses physiologiques et psychologiques de participants confrontés à différentes situations

<http://mshs.univ-toulouse.fr/spip.php?article106> (19.06.2019)

- La plateforme (450 m²) est hébergée, depuis son ouverture en janvier 2015, par la [MSHS-T](#) à la Maison de la Recherche de l'université Toulouse Jean-Jaurès.
- Les chercheurs responsables des 9 plateaux de la plateforme [CCU](#) sont issus du laboratoire CLLE et de la Maison des Sciences de l'Homme et de la Société de Toulouse.

EYE-TRACKING / OCULOMETRIE

- Oculomètre : trois enregistreurs de mouvements oculaires (un modèle EYELINK 1000, un modèle SMI RED 250, et un modèle ASL 504)
 - traçabilité des prises d'informations visuelles (saccades, points de fixation, diamètre pupillaires, ...)
 - tâches épurées ou des tâches complexes (conduite automobile, lecture, ...)
- Un oculomètre portatif (SMI Eye Tracking Glasses)
- Recherches sont variées actuellement : neurophysiologie de la perception des visages, psycholinguistique textuelle, étude du jugement et de la décision, étude de la mesure de l'intelligence fluide, perception et vieillissement
- Plateau impliqué et financé dans plusieurs projets (ANR, Convention ENAC/UTM, Bourse EADS, FEDER)



BILAN

BILAN

- 1. Wie sehen der aktuelle, technologisierte Arbeitsplatz von Untertitlern und die Produktionsabläufe von Untertitelungsprojekten aus?
 - FTI (Aurélie Picton)
 - CLLE-ERSS (Amélie Josselin-Leray)
 - CLLE-LTC (mais partenaires précis à définir sur la base des discussions et intérêts)

Pistes possibles (générales)

- Mise en place d'un protocole
- Analyse des ressources et besoins linguistiques « en action »
- Analyse des différents types d'interactions (avec ressources, personnes, etc.)
- Charges mentale, multitasking et/ou autres en fonction des éléments présentés et intérêts

BILAN

- 2. Welche sprechsprachlichen Merkmale weisen Untertitel im Vergleich zu audiovisuellen Originalproduktionen einerseits, zu spontaner Mündlichkeit andererseits auf
 - Intérêt exprimé par :
 - FTI (Aurélie Picton)
 - CLLE-ERSS (Cécile Fabre et Mai Ho-Dac; Amélie Josselin-Leray)

Pistes possibles (générales)

- Cf. pistes sur oralité présentées par CLLE-ERSS
- Analyse des processus de créativité lexicale et de variation entre productions originales et spontanées
- Caractérisation et résolution de difficultés de traduction liées à l'oral

BILAN

- 3. Welche Erwartungen an Untertitel in Bezug auf ihre sprachliche Gestaltung haben Hörende einerseits, Hörbehinderte andererseits?
 - Intérêt potentiel avec :
 - CLLE-ERSS

Pistes possibles (générales)

➤ Collaboration avec CETIM (cursus LSF)

CONTACTS POUR LA SUITE SUR CES ASPECTS

- FTI : Aurélie Picton, aurelie.picton@unige.ch
- CLLE-ERSS : Amélie Josselin-Leray, amelie.josselin@univ-tlse2.fr
- CLLE-LTC : Franck Amadieu, amadieu@univ-tlse2.fr
- Point langue : FR, EN, DE (Amélie Josselin-Leray)