

Faciliter l'accès aux échanges en ligne et leur analyse, le cas de ViCoDiLi

Philippe TEUTSCH, Francis BANGOU, Charlotte DEJEAN-THIRCUIR
[(LIUM, Le Mans)]

■ **RÉSUMÉ** : Cet article s'intéresse aux traces issues d'échanges entre participants (étudiants et tuteurs) d'une formation en ligne. Il étudie la double question de l'accès à ces traces et de leur manipulation par les chercheurs analystes. Cette réflexion a été guidée par la nécessité de restituer l'accès à un corpus issu d'une formation dont la plateforme d'origine n'était plus disponible. Ce "défi" a été relevé à travers la conception de ViCoDiLi, outil destiné à la Visualisation de Contenus de Discussions en Ligne qui permet aux chercheurs d'accéder aux traces d'échanges (forum, clavardage et courriel) d'une session de formation et de les examiner selon différentes modalités d'affichage. Le projet se situe dans une démarche d'instrumentation des activités du chercheur. L'objectif est de proposer un modèle d'accès aux données textuelles de corpus, ainsi que des outils d'aide à l'analyse de ces corpus. Cet article présente les questionnements, à la fois théoriques, techniques et méthodologiques soulevés par la conception d'un tel outil.

■ **MOTS CLÉS** : échanges en ligne, manipulation de corpus textuels, analyse qualitative, anonymisation, diffusion de corpus.

■ **ABSTRACT** : This article focuses on the traces of interactions between students and tutors during an online training. We examine both, the issues of access to these traces as well as their manipulation by the researcher. This reflection was guided by the need to access the data of an online course that became unavailable at some point. This challenge resulted in the development of a tool called Visualizing the Content of Online Interactions (ViCoDiLi) that enables both the access to the formatted traces of online interactions (forum, chat and email), and their examination with diverse display features. The project aimed at providing tools for the researcher's activities. The objective is to provide both a model of data access and tools to assist the data corpus analysis. This article highlights the theoretical, technical and methodological issues raised by the development of such a tool.

■ **MOTS CLÉS** : Online Interactions, Manipulation of Textual Data, Qualitative Analysis, Anonymization, Data Corpus Dissemination.

- [1. Introduction](#)
- [2. Contexte de recherche](#)
- [3. Analyse qualitative des interactions en ligne](#)
- [4. Présentation de l'outil ViCoDiLi](#)
- [5. Conclusion](#)
- [BIBLIOGRAPHIE](#)

1. Introduction

Les environnements numériques utilisés pour la mise en œuvre de dispositifs de formation ouverte et à distance (FOAD) gèrent un grand nombre de données produites par l'activité des participants. Les chercheurs peuvent ainsi accéder aux contributions issues des activités d'apprentissage, ou encore aux messages échangés grâce aux outils de communication utilisés. Ces traces numériques conservées sur les ordinateurs intéressent les chercheurs dans leur démarche d'observation et de compréhension des situations d'enseignement-apprentissage provoquées par les dispositifs médiatisés. L'analyse des traces d'activité permet par exemple d'améliorer l'utilisation et l'appropriation des dispositifs par les utilisateurs

(Tselios et al., 2008), (Ollagnier-Beldame et Mille, 2008), d'étudier l'usage des dispositifs dans une démarche de ré-ingénierie (Choquet et al., 2007), de les interroger dans le cadre d'une analyse didactique (Lamy, 2006) ou encore de les caractériser dans le cadre d'une formation de formateurs.

L'étude des échanges verbaux entre les acteurs de ces environnements médiatisés de formation constitue un champ de recherche particulièrement porteur (Bruillard, 2006), (Henri et al., 2007)¹. Cette thématique intéresse plusieurs domaines de référence : la communication médiatisée par ordinateur (CMO) qui étudie la spécificité des interactions en ligne selon les outils utilisés (Preece et Maloney-Krichmar, 2005), (Herring, 2004), la communication pédagogique médiatisée par ordinateur (Peraya, 2005), (Mangenot, 2006), le domaine de l'apprentissage collaboratif assisté par ordinateur (ou CSCL : Computer Supported Collaborative Learning) où les technologies de l'information et de la communication sont envisagées comme des moyens de favoriser les apprentissages dans l'interaction (Teasley et Roschelle, 1993), (Dillenbourg et al., 1996), l'EIAH à travers l'étude des pratiques en relation avec les outils qui les instrumentent (George et Bothorel, 2006). Les principales disciplines concernées par ces recherches sont les sciences du langage, la didactique des disciplines, la sociologie, les sciences cognitives, les sciences de l'information et de la communication et l'informatique.

Pour toutes ces recherches sur les interactions en ligne en contexte de formation, il est nécessaire de travailler avec un ou plusieurs corpus issus de dispositifs de formation réels, avec de vrais enjeux et responsabilités pour les participants (Bruillard, 2006). Cette exigence de travailler sur des données issues de formations authentiques soulève la question de la préservation et de la diffusion de ces données car les plateformes utilisées pour l'accueil des formations en ligne n'offrent qu'une diffusion limitée des traces d'activité et ne permettent pas de garantir leur pérennité. Or, la démarche de diffusion, de mutualisation et d'interopérabilisation fait partie des recommandations méthodologiques de la communauté (Henri et Charlier, 2005) et concerne autant les corpus eux-mêmes que les outils d'analyse et les résultats obtenus (Bruillard, 2006). (De Wever et al., 2006) par exemple, recommandent que les instruments d'analyse soient comparés entre eux. La question de la construction d'instruments valides préoccupe les chercheurs intéressés à développer des méthodes éprouvées pour l'analyse des interactions médiatisées (Bruillard, 2006).

Ces recommandations ont été prises en compte au sein du projet Odil² dans le cadre de la mise en place de deux actions de recherche parallèles : le projet Mulce³, piloté par le LIFC et cherchant à définir un protocole de diffusion de corpus et le projet ViCoDiLi⁴, piloté par le LIUM (Laboratoire d'Informatique de l'Université du Maine). Ce dernier a abouti à l'élaboration d'un outil de Visualisation de Contenus de Discussions en Ligne permettant aux chercheurs d'accéder aux échanges d'une session de formation, hors de leur plateforme d'origine, et de les examiner selon différentes modalités d'affichage. Motivée par le besoin d'accéder aux données d'un corpus d'apprentissage en ligne, la création d'un tel outil soulève des questions liées à l'instrumentation des activités d'analyse de ce type de corpus, au format d'échange des données, à leur structuration et à leur manipulation.

Cet article aborde ces différentes questions avant de présenter l'outil lui-même. La première partie porte sur le contexte général de la recherche. Nous présentons ensuite les spécificités de l'analyse des interactions en ligne afin de rendre compte des besoins des chercheurs. Enfin, la dernière partie porte sur la présentation, la description et un premier bilan de l'outil ViCoDiLi.

2. Contexte de recherche

La conception de ViCoDiLi a été guidée par des questions d'accès aux données d'un corpus d'apprentissage. Cette partie présente les origines, la problématique et les objectifs du projet ainsi que la méthodologie de conception utilisée.

Origines du projet : retrouver un corpus perdu

Les plateformes de formation en ligne (FEL) ne sont pas conçues pour que les données (notamment de type échanges) soient conservées à long terme et puissent ainsi constituer un matériel d'étude. Même si ces données peuvent être exploitées en tant que corpus pendant un temps par les chercheurs qui ont eux-mêmes mis en place des formations ou des modules d'enseignement en ligne, les données risquent de disparaître. Leur accès devient alors dépendant des techniciens de l'informatique, des possibilités

d'extraction normalisées des plateformes, des formats de structuration interne des données (bases de données interrogeables ou exportables) et de la possibilité de retrouver une interface de consultation de ces données.

La question de la réutilisation des corpus s'est posée dans le cadre du projet Odil à l'occasion de la perte puis de la restauration d'un des corpus du projet. Les différents corpus concernés étaient tous dépendants de l'environnement informatique support à la formation les ayant produits. Pour consulter un corpus et effectuer leur tâche d'analyse, les chercheurs utilisaient leur identifiant d'enseignant et les modalités de consultation de la plateforme, ce qui suppose que la session de formation concernée par l'analyse soit toujours accessible et que la plateforme soit toujours opérationnelle. Un de ces corpus est devenu complètement inaccessible lorsque l'organisme de formation a arrêté l'exploitation de la plateforme hôte. Il s'agit du corpus Simuligne, issu d'une formation à distance pour l'apprentissage du FLE (Français Langue Étrangère) médiatisée en 2001 sur la plateforme WebCT, dans le cadre du projet ICOGAD (Reffay et al., 2002). Au cours de cette formation, les apprenants ont participé à une activité de simulation globale⁵ en ligne. Le scénario invitait les apprenants à une production collaborative dans la langue cible. La situation de communication créée par le dispositif Simuligne était authentique, et le mode de communication textuel et asynchrone. La collaboration développée au sein de cette activité a donné lieu à de nombreuses interactions entre les étudiants et entre tuteur et étudiants pour organiser, négocier, décider et finalement produire ensemble. La simulation a rassemblé 40 adultes anglophones en formation continue, 10 étudiants natifs francophones et 4 tuteurs (1 par groupe). Elle s'est déroulée sur 10 semaines et a produit plus de 12 000 interventions à travers les outils de forum, courriel et clavardage.

Après l'arrêt de la licence d'exploitation de la plateforme hôte (WebCT), les chercheurs disposaient d'une sauvegarde numérique de l'ensemble des contributions produites par la formation mais n'avaient aucun moyen technique de consulter ou de manipuler ces données réduites à une arborescence de fichiers texte. L'enjeu technique et scientifique lié à la restauration du corpus Simuligne était de permettre aux propriétaires d'un ensemble de données gérées par une plateforme "disparue" de retrouver ces données sous forme de corpus structuré et homogène présentant les échanges de façon assez proche de leur mode de structuration d'origine. C'est cet objectif qui a été poursuivi et atteint à travers la conception de l'outil ViCoDiLi, celle-ci s'inscrivant dans un questionnement sur la mise en place d'un modèle de structuration des données indépendant de la plateforme d'origine.

Problématique

Dans un article proposant un état des lieux des recherches dans le domaine de la formation en ligne (Henri et Charlier, 2005), Henri et Charlier remarquent que certains résultats d'expérimentation de formations en ligne sont présentés comme étant encourageants alors que d'autres sont plutôt mitigés, voire décevants, quant à l'efficacité de ces environnements en termes d'apprentissage. Pour expliquer ces disparités dans les résultats d'analyse, les auteurs soulignent que les recherches concernées sont menées dans des contextes différents, peu comparables, avec des méthodes d'investigations et d'analyse également variées. Pour pouvoir partager des résultats issus d'analyses de corpus d'échanges en ligne, il semble donc nécessaire de rendre ces corpus accessibles à d'autres chercheurs que ceux qui les ont créés ainsi que de proposer des outils spécifiques à ce travail d'analyse et susceptibles d'être appliqués à différents types de corpus.

La conception de l'outil ViCoDiLi qui a, entre autres, permis la restauration du corpus Simuligne repose sur l'hypothèse qu'une relative harmonisation des modalités de structuration, de visualisation et de manipulation de ces données peut faciliter le travail d'analyse sur les corpus. L'intérêt a priori est, par exemple, de pouvoir dérouler un même protocole d'analyse sur plusieurs corpus, ou de comparer des corpus issus d'un même scénario pédagogique et joué dans des contextes différents : participants, calendrier ou environnement technologique. L'objectif de cette recherche est alors de montrer qu'il est possible de permettre un accès homogène à des corpus hétérogènes.

Les questionnements soulevés par la réalisation de cet outil ont porté sur la préservation, la structuration et les modalités de manipulation des données, ainsi que sur leur anonymisation, condition indispensable à leur diffusion.

Structurer des corpus d'échanges en ligne hétérogènes

On peut considérer le corpus d'échanges comme un ensemble de données que l'on cherche à extraire de leur environnement de production afin de les mettre à disposition des chercheurs. Chercher à rendre les données indépendantes de leur environnement d'origine conduit à s'interroger sur les propositions de normalisation des traces d'activité sur plateforme de téléformation et de structuration des corpus de traces. Trois questions se posent alors : définir la granularité des objets d'étude, identifier les outils de communication concernés et définir un format de référence permettant une manipulation homogène des données.

Les méthodologies de recherche et les techniques liées à l'étude de corpus (transcription, annotation, codage et interprétation) interrogent en premier lieu l'identification et le grain de l'unité d'analyse de référence (Rourke et al., 2001). S'agit-il de l'ensemble des discussions, du fil de discussion, du message, d'une partie du message, d'une phrase, d'une expression, ou d'un marqueur linguistique ? L'unité d'analyse retenue le plus souvent dans les recherches sur les interactions en ligne est le message. Ce dernier peut être considéré comme une intervention qui, combinée à d'autres interventions, constitue un échange. Le terme "échange" est ici utilisé dans son sens technique, en référence au domaine de l'analyse séquentielle des interactions (Charaudeau et Maingueneau, 2002). Il renvoie à l'unité dialogale minimale, constituée d'au moins deux interventions produites par deux locuteurs, l'une initiative, l'autre réactive.

Les dispositifs de formation en ligne s'appuient sur différents outils de communication permettant les échanges entre participants, principalement forum et clavardage. La variété de ces outils qui instrumentent les échanges en ligne est à prendre en considération du point de vue de l'analyse. Les plateformes de formation en ligne actuellement disponibles (WebCT, QuickPlace, Dokeos, Caroline et Moodle dans le cas du projet Odil) possèdent différents modes de présentation des données (Mangenot, 2008) (dans ce numéro). Ainsi, les fils de discussion sont plus ou moins structurés : il n'y a pas de structuration des discussions sur la plateforme Esprit, tandis que Moodle propose un outil de structuration performant qui permet, d'abord aux participants, puis à l'analyste, d'effectuer des recherches par liste, par structure et par historique. Les interfaces de présentation et de manipulation de ces données (affichage, création, reprise du message auquel on répond, indentations) sont souvent complexes et hétérogènes.

Les structurations internes des données sont, elles aussi, hétérogènes. Certaines plateformes utilisent des fichiers structurés spécifiques (WebCT), d'autres des bases de données propriétaires ou standardisées (MySQL pour Moodle). Quelques-unes permettent d'exporter les traces d'activité dans un format d'échange normalisé (XML). Ce constat d'hétérogénéité des plateformes et des formats de traces d'échanges qu'elles produisent nécessite de définir un format de référence permettant à la fois d'uniformiser la description des interventions d'un point de vue informatique et d'optimiser la manipulation de ces données du point de vue de l'analyste. L'outil ViCoDiLi s'appuie sur un tel format de référence pour offrir des fonctionnalités de manipulation homogènes d'un corpus à l'autre.

Manipuler des interventions en contexte

La participation aux activités médiatisées constitue "une activité située, finalisée et cadrée par un contexte" (Henri et Charlier, 2005). La connaissance de ce contexte et des différentes dimensions concernées est globalement nécessaire à la compréhension des contributions aux activités proposées et des échanges en ligne.

On peut constater que les échanges s'inscrivent dans un contexte à deux dimensions : contexte général du dispositif de formation et contexte spécifique de la session de formation. Le dispositif décrit le cadre général de la formation : objectifs, activités, environnement de travail médiatisé, modalités d'accompagnement. La session est une instance du dispositif pour un groupe de participants (apprenants et accompagnateurs) donnés et sur une période donnée (calendrier, planning). Le corpus d'échange est directement lié à ce contexte spécifique. La conception de ViCoDiLi a été guidée par la prise en compte de ce contexte spécifique à la session de formation.

Anonymiser les données

D'un point de vue éthique, le caractère authentique, personnel et social des corpus d'interactions en ligne interroge les conditions de diffusion, d'exploitation et d'affichage de ces corpus. Pour des questions de protection des personnes, l'accès aux contributions des acteurs d'une session de formation ne peut se faire qu'à condition de ne pas pouvoir reconnaître les personnes concernées (Reffay et Teutsch, 2007).

L'extrait suivant illustre ce besoin d'anonymisation des corpus préalablement à leur diffusion, ne serait-ce que pour en citer un extrait.

"Nous avons constitué un corpus de textes (courriels) écrits par une personne [...] âgée d'une quarantaine d'années. [...] Ce corpus est constitué de 205 courriels issus de sa correspondance personnelle. [...] Tous les noms propres ont été changés afin de rendre ce corpus anonyme et utilisable par d'autres. Nous avons ensuite procédé à la suppression des entêtes et des signatures des courriels". (Boissière et Schadle, 2006) (p. 4).

Nous avons cependant annoncé précédemment que la compréhension et l'analyse du corpus nécessitent de connaître le contexte de production des échanges en ligne. La problématique de l'anonymisation est donc de préserver la perception du contexte tout en protégeant l'identité de la personne physique (Reffay et Teutsch, 2007).

Démarche de conception

Le double objectif de réaliser un environnement informatique soutenant l'activité de l'analyste tout en élaborant un modèle des données à analyser pose en soi une difficulté méthodologique. Notre proposition d'interface d'exploration homogène pour différents corpus issus de différentes plateformes risque de modifier le point de vue des analystes sur la notion même de corpus. Cette situation d'évolution technologique est caractéristique des travaux en Interaction Humain-Machine (IHM) qui, nécessairement pluridisciplinaires, cherchent à augmenter la performance du couple système-utilisateur dans le cadre de la conception d'environnements interactifs.

Le concept le plus important en IHM est probablement celui de conception centrée sur l'utilisateur (Norman, 1988), qui prend en compte la perspective de l'utilisateur dans le processus de conception de la nouvelle technologie. Dans ce cadre, la rencontre entre disciplines permet de multiplier les perspectives et d'avoir une représentation la plus large possible de l'activité considérée pour la création d'un artefact (Mackay et Fayard, 1997). L'usager étant au centre du processus de conception, Rabardel (Rabardel, 2005) définit la co-adaptation comme un phénomène croisé d'instrumentation de l'utilisateur par l'artefact proposé et d'instrumentalisation de ce même artefact à travers la mise en place de schèmes d'utilisation que l'utilisateur développe dans la situation d'usage réel.

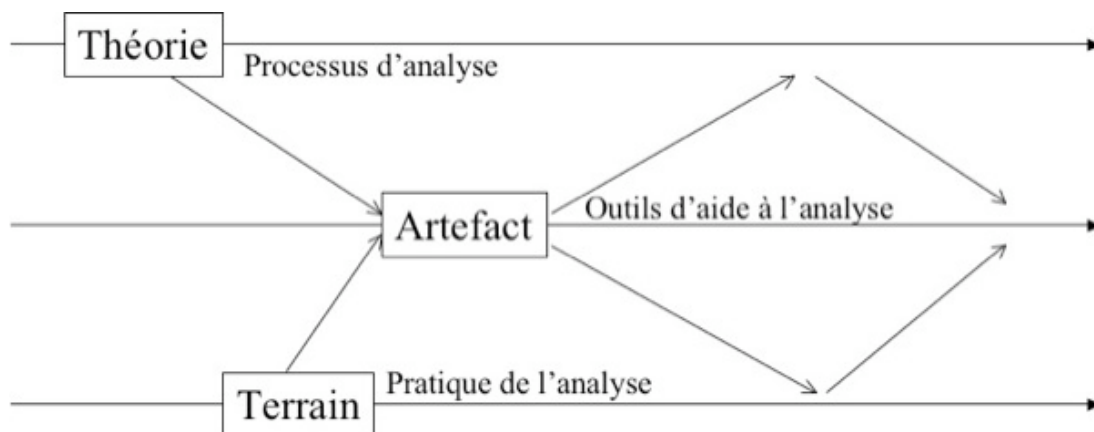


Figure 1, Principe d'instrumentation appliqué à notre problématique

La démarche participative (Mackay et Fayard, 1997) consiste alors à intégrer les utilisateurs (analystes) au processus de conception et d'évaluation du système en devenir. Le principe méthodologique est de leur permettre d'exprimer leurs besoins et leurs points de vue, et d'explicitier leur pratique. Mackay et Fayard expliquent que les démarches de recherche prennent en compte l'introduction de l'artefact dans la situation d'usage. La conception d'un artefact se nourrit (Figure 1), d'une part d'une approche théorique vérifiée par expérimentation, et d'autre part d'une observation de terrain se structurant en modèle théorique. Le principe est de compléter l'analyse de la tâche a priori par l'analyse des besoins des acteurs qui effectuent cette tâche sur le terrain. Chaque cycle permet de faire évoluer le prototype, d'affiner le modèle tout en créant de nouveaux usages. Les trois champs d'étude concernés sont donc le modèle théorique, l'artefact informatique et l'usage réel. Cette démarche de conception itérative et participative permet d'identifier les besoins essentiels et de définir les outils correspondants, sans pour autant disposer

au départ d'un modèle complet de l'activité de l'analyste. Pour chaque cycle de conception, l'étape d'évaluation est essentielle.

Dans le cadre de la diffusion de corpus, l'accessibilité aux données peut s'évaluer à travers les critères de qualité proposés par le projet Mulce⁶ : pérennité, lisibilité, homogénéité et validité. Nous illustrons ces critères à travers les caractéristiques des corpus disponibles à l'origine au sein du projet Odil.

- Lisibilité : les données sont-elles lisibles à travers (au moins) un moyen de visualisation, dépendant ou non du dispositif d'origine ? Concernant le corpus Simuligne, l'accès aux données étant dépendant de la plateforme hôte (disparue), les acteurs de la formation eux-mêmes étaient initialement dans l'impossibilité de retrouver les traces de leur activité.

- Pérennité : les données sont-elles accessibles en dehors de la plateforme d'origine, pour une durée de vie indépendante du temps de la formation concernée ? Concernant les corpus du projet Odil, de nombreuses données sont dispersées ou incomplètes. Leur accès était limité aux participants à la formation, et interdit aux chercheurs extérieurs.

- Homogénéité : la structure des corpus permet-elle une manipulation homogène des données pour l'indexation ou la recherche ? Dans le cadre du projet Odil, les données "brutes" étaient toutes dépendantes du système d'information d'origine. Quand ils existent, les structures de bases de données et les formats d'échange sont hétérogènes.

- Validité (écologique, durée, volume) : le corpus est-il issu d'une formation réellement "écologique", et donc non expérimentale, en termes de conception de dispositif pédagogique, de durée de la formation et de nombre de participants ?

Afin de tenir compte des tâches et des démarches des analystes, l'élaboration de l'outil ViCoDiLi a suivi une méthodologie de conception itérative et participative (un informaticien et deux chercheurs en didactique). Les aspects méthodologiques de l'analyse des interactions en ligne sont présentés dans la partie suivante.

3. Analyse qualitative des interactions en ligne

Les recherches sur les "interactions pédagogiques en ligne" (Mangenot, 2006) se fondent sur des approches quantitatives, comme l'analyse de contenu par exemple, ou sur des approches qualitatives ayant recours à l'analyse du discours et à l'analyse des interactions verbales. L'analyse de contenu, telle qu'elle est définie par Charaudeau et Maingueneau (Charaudeau et Maingueneau, 2002) dans leur Dictionnaire d'analyse du discours, est "*une technique de recherche pour la description objective, systématique et quantitative du contenu manifeste de la communication*" (p. 39). Cette méthodologie implique "*deux opérations fondamentales : la précatégorisation thématique des données textuelles, et leur traitement quantitatif, généralement informatisé*" (*ibid.*). La linguiste américaine S. Herring, qui s'intéresse aux échanges en ligne, non pédagogiques, et qui propose une approche intitulée Computer Mediated Discourse Analysis, souligne les limites de l'analyse de contenu quantitative :

"[...] quantitative content analysis may not be the best approach for analyzing complex, interacting, ambiguous or scalar phenomena, which risk distortion by being forced into artificially discrete categories for purposes of counting. Such phenomena may be more richly revealed by qualitative, interpretive approaches [...]" (Herring, 2004) (p. 369).

L'analyse qualitative des interactions permet d'étudier l'organisation séquentielle des échanges et les procédures de co-construction mises en œuvre par les participants pour interagir et pour réaliser les tâches à accomplir en situation de formation. En outre, comme l'écrit F. Mangenot (Mangenot, 2006), ce type d'analyse permet d'observer plus finement les dimensions socio-affectives et socio-culturelles des échanges.

La conception de l'outil ViCoDiLi s'inscrit plutôt dans cette deuxième perspective d'une analyse qualitative des interactions. Celle-ci passe par des opérations de manipulation des échanges que nous allons maintenant présenter avant d'évoquer quelques logiciels d'aide à l'analyse des données.

Pratiques et besoins de l'analyse

Selon Tesch (Tesch, 1990), la tâche centrale du chercheur qui met en œuvre une analyse qualitative des données revient à construire une structure, à manipuler des extraits de corpus, à les comparer, à les associer dans le but de leur donner du sens. Dans le cadre de l'analyse des interactions en ligne, le corpus est principalement constitué des traces des échanges entre les participants, ce qui n'exclut pas que le chercheur puisse croiser ces traces avec d'autres types de données tels que des questionnaires ou des entretiens (Jeannot et al., 2006). Pour analyser les interactions en ligne, le chercheur est amené à effectuer différents types d'opérations impliquant la manipulation des corpus. Voici quelques unes de ces opérations identifiées à travers différents articles proposant des analyses de ce type :

- **Manipuler des échanges provenant de différents outils de communication.** Dans le cadre de formations en ligne, les apprenants ont à leur disposition des outils de communication leur permettant d'interagir de manière synchrone ou asynchrone. Les chercheurs souhaitant travailler sur ces différents types d'interactions ont besoin de pouvoir passer facilement d'un type de données à un autre. Ainsi Dejean-Thircuir (Dejean-Thircuir, 2008) utilise un corpus composé de messages provenant de forums et de clavardages dans le cadre d'une étude sur la constitution d'une communauté d'apprenants engagés dans la réalisation d'une tâche collaborative.

- **Isoler des interventions ou des extraits d'intervention.** Dans le cadre du projet de recherche Galanet portant sur la mise en place du contrat didactique dans une formation à l'intercompréhension en langues romanes (Degache, 2006), l'auteur commente des extraits d'interventions de différents locuteurs afin d'illustrer un point précis. Dans l'exemple ci-dessous, le lecteur ne sait pas ce qui a précédé ou ce qui a suivi cette intervention. Cette dernière est exploitée dans l'article en tant que trace d'un contrat codique établi entre les participants de la formation. Pour analyser des échanges en ligne, le chercheur repère des interventions pertinentes et les isole.

"Même lorsqu'il n'est pas négocié, on trouve des traces de ce contrat codique dans des opérations appréciatives positives ou négative : Je trouve ça génial que l'on soit capable d'animer un forum en quatre langues !!!..." (Degache, 2006) (p. 67)

- **Isoler des échanges.** Dans le cadre de leur étude consacrée aux rôles joués par les tuteurs lors d'échanges franco-japonais en ligne (Mangenot et Tanaka, 2007), les auteurs s'appuient sur de nombreux extraits de corpus présentant des échanges. Ainsi, à la suite de l'extrait présenté ci-dessous, les auteurs donnent l'exemple d'un échange à deux constituants (question-réponse), tiré d'un forum. Les deux interventions citées font peut être partie d'un échange plus important, la question ayant suscité de nombreuses réponses, mais cet échange minimal a suffi pour illustrer une des dimensions de l'analyse. De même qu'il identifie et isole des interventions, le chercheur repère et isole des échanges.

"Concernant l'inquiétude de certains étudiants français quand ils constatent que peu d'étudiants ont participé à leur tâche, l'enseignante japonaise tente d'y répondre en décrivant le style d'apprentissage des japonais comme dans l'échange suivant..." (Mangenot et Tanaka, 2007) (p. 9).

- **Référencer les extraits de corpus.** Les extraits de corpus utilisés par les chercheurs sont souvent accompagnés d'informations concernant le contexte spécifique de production des interventions (auteur, fil de discussion, thème, date et heure, rang). C'est le cas de l'exemple suivant :

"17/03/2004 C1, salon rouge, entre Vasco, Portugais, et Liliana Argentine en espagnol [E179C]" (Degache, 2006) (p. 63).

- **Regrouper des messages ou des extraits de messages.** Dans le cadre d'une étude sur les positionnements de tutrices en formation tutorant des apprenants de français en Australie (Dejean-Thircuir et Mangenot, 2006), les auteurs ont rassemblé des extraits de messages de différentes participantes pour rendre compte de certains traits stylistiques de leurs messages (en l'occurrence l'usage d'un registre familier).

"Waouh! C'est génial, tu nous as envoyé pleins de message et tu as carrément bien avancé dans les activités ...[Julie]

Tout baigne Judith ! [Cristina]" (Dejean-Thircuir et Mangenot, 2006) (p. 81)

- **Mettre en forme un texte.** Dans un article présentant une intervention provenant d'un forum de discussion (Fustenberg et English, 2006), les auteurs précisent : "c'est nous qui mettons en gras" (p. 188).

Le chercheur peut donc remettre en forme (sur le plan typographique) des extraits de corpus afin de faciliter le repérage et la lecture des éléments essentiels liés à l'analyse en cours.

Les différents exemples présentés ci-dessus montrent que dans le cadre d'analyses qualitatives des interactions en ligne, les chercheurs sont, entre autres, amenés à effectuer des opérations de repérage, d'extraction, de contextualisation, de regroupement, et de mise en forme des traces des échanges. Ces manipulations "directes" des données supposent de pouvoir accéder aux traces des échanges avec une structuration et une mise en forme proches de celles d'origine.

Logiciels d'assistance à l'analyse

De nombreux logiciels ont été élaborés afin de faciliter l'analyse de corpus écrits. Les plus connus sont The Ethnographer, Atlas/TI, NUD*IST, et Xsight. Ces logiciels permettent la manipulation des corpus en regroupant dans un même espace des documents textes variés et volumineux. Grâce à eux, le chercheur peut passer aisément d'un document à un autre afin d'effectuer des opérations d'extraction, de contextualisation, de regroupement, et de mise en forme. Ils permettent en outre de gérer des annotations et de garder un lien constant entre chaque extrait et son contexte de production. Ces logiciels sont destinés à favoriser les analyses qualitatives mais ils sont essentiellement conçus pour des corpus issus d'observations de terrain ou d'entrevues effectuées par le chercheur et ne sont pas adaptés à l'analyse de données textuelles ayant une structure complexe, comme celle des échanges dans certains types de forums par exemple.

D'autres logiciels ont été créés dans le but de représenter graphiquement les interactions en ligne ([Heer et Boyd, 2005](#)), ([Gibbs et al., 2006](#)). [Reffay et Lancieri \(2006\)](#) proposent de visualiser le corpus Simuligne sous forme de "réseaux sociaux" pour permettre une analyse structurale et quantitative des interactions.

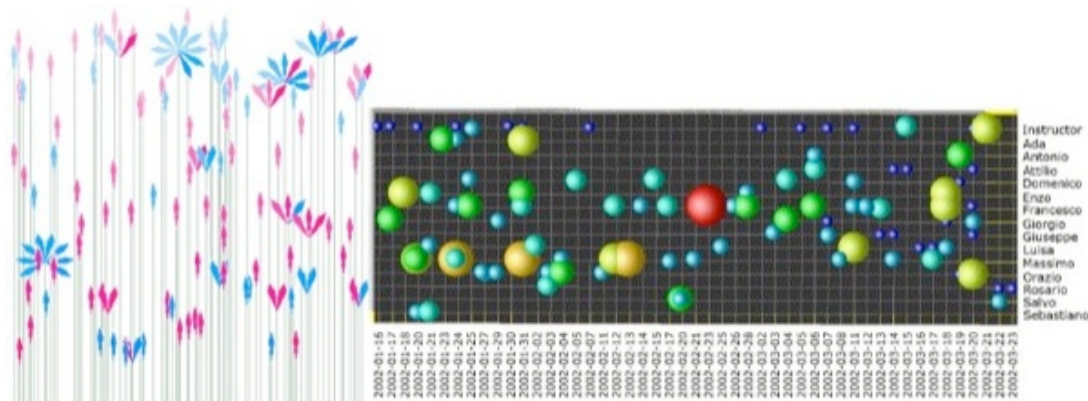


Figure 2, Visualisation de discussion avec a) PeopleGarden et b) CourseVis

Le logiciel PeopleGarden ([Xiong et Donath, 1999](#)) représente les échanges qui se construisent dans un forum grâce à la métaphore de la croissance d'une fleur en tenant compte de la durée de la participation et du nombre de réactions à un message initial (Figure 2a). CourseVis ([Mazza et Dimitrovab, 2007](#)) affiche quant à lui les fils de discussion sous forme de sphères qui représentent l'impact provoqué par chaque intervention initiale (Figure 2b).

Ces logiciels permettent de réaliser des analyses quantitatives des interactions. Ils apportent des informations sur les attributs des messages (origine, destinataire, date, impact quantitatif, etc.) et non pas sur leur contenu. En outre, ces logiciels sont généralement conçus dans le but d'analyser des interactions provenant d'un seul outil de communication (forum principalement), mais pas d'un ensemble d'outils de communication hétérogènes. L'outil ViCoDiLi, en revanche, permet d'effectuer des opérations de manipulation liées à la structure et au contenu des interactions synchrones et asynchrones en ligne provenant d'outils de communication variés.

4. Présentation de l'outil ViCoDiLi

ViCoDiLi est une application web de Visualisation de Contenus de Discussions en Ligne qui permet de

manipuler des échanges écrits, structurés par des outils de communication de type forum, clavardage, courriel ou blog. Il fonctionne sur le corpus Simuligne issu d'une plateforme WebCT ainsi que sur des corpus issus de la plateforme Moodle. L'outil a pour vocation d'être étendu à d'autres corpus issus de différentes plateformes de formation en ligne. Cette partie présente successivement l'environnement proposé à l'utilisateur, le format de description des échanges et l'architecture générale du système.

Interfaces de visualisation et de manipulation

La méthode de conception participative utilisée pour le projet et la prise en compte d'études portant sur l'analyse d'échanges en ligne ont abouti à la définition de spécifications fonctionnelles pour l'outil ViCoDiLi. Les interfaces support de ces fonctionnalités sont illustrées avec le corpus anonymisé de Simuligne (Reffay et Teutsch, 2007).

ViCoDiLi permet de naviguer rapidement du niveau le plus général au niveau le plus détaillé des corpus disponibles (Figure 3). Les différents niveaux de navigation correspondent aux choix de corpus, de lieu et de fil de discussion. L'utilisateur sélectionne successivement un des corpus qu'il a à sa disposition (liste déroulante en haut de la figure), un des lieux de discussion de ce corpus (notion de groupe de travail dans le cas de Simuligne, notion de salle pour d'autres dispositifs), et enfin le fil de discussion qui l'intéresse (colonne de gauche pour ces derniers niveaux).

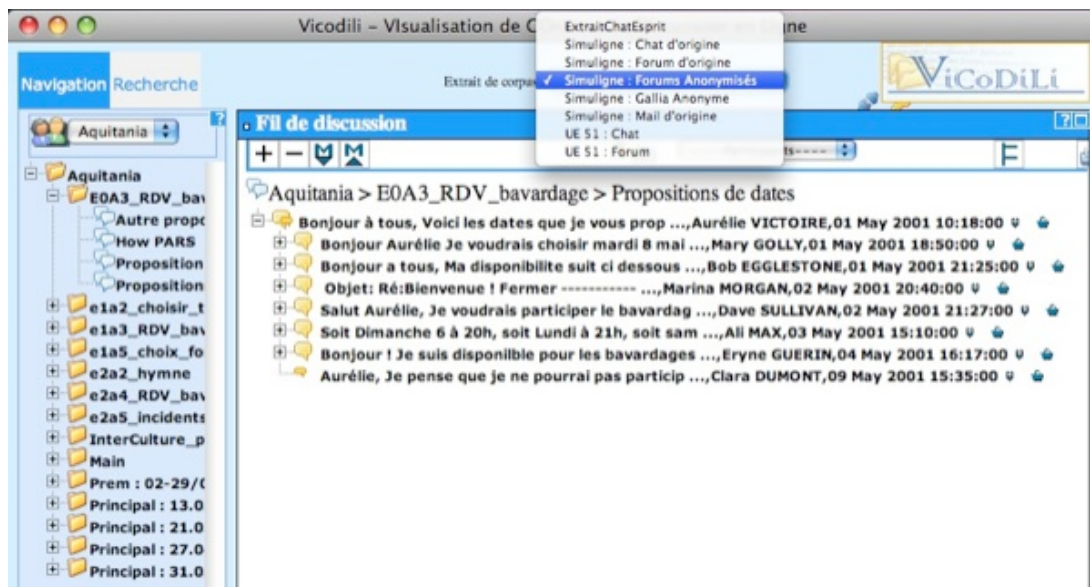


Figure 3, Choix de corpus et de discussion dans ViCoDiLi

Les possibilités d'affichage de chaque fil de discussion dans l'espace central (Figure 4) reprennent les conditions existant généralement dans les environnements de discussion en ligne. Chaque fil de discussion est ainsi présenté sous une forme arborescente que l'utilisateur peut déployer, ou réduire, pour l'ensemble de l'arbre ou pour chaque noeud. Les messages peuvent être présentés sous forme résumée (titre ou première ligne en complément de l'auteur et de la date) ou intégralement, là aussi pour l'ensemble de l'arbre et/ou pour chaque message. Une fonction de repérage visuel permet de souligner en couleur le nom de chacun des participants impliqués dans la discussion et de déployer toutes ses interventions (Mary Golly sur la figure).

Les fonctionnalités d'assistance spécifiques à l'analyse fournies par ViCoDiLi concernent la recherche et le regroupement des interventions. La recherche s'effectue par auteur ou par mot clé sur l'ensemble d'un corpus. La liste des occurrences trouvées permet de visualiser chacune d'entre-elles (via l'espace central) dans son contexte de production : fil de discussion et intervention concernés. L'utilisateur peut également sélectionner et regrouper des interventions ayant un intérêt particulier et ainsi recomposer son propre matériel d'étude. Chaque message de l'espace central peut être sélectionné pour rejoindre l'espace "panier" situé en bas de la page (Figure 5). Cet espace possède ses propres fonctionnalités permettant de trier, de regrouper et de supprimer les interventions sélectionnées. La liste d'interventions obtenue peut être exportée sous forme de fichier texte (pour effectuer un traitement complémentaire) ou sous forme de fichier .pdf (pour diffusion). L'avantage du panier est de pouvoir regrouper dans un même format des

interventions issues de corpus de différents types.

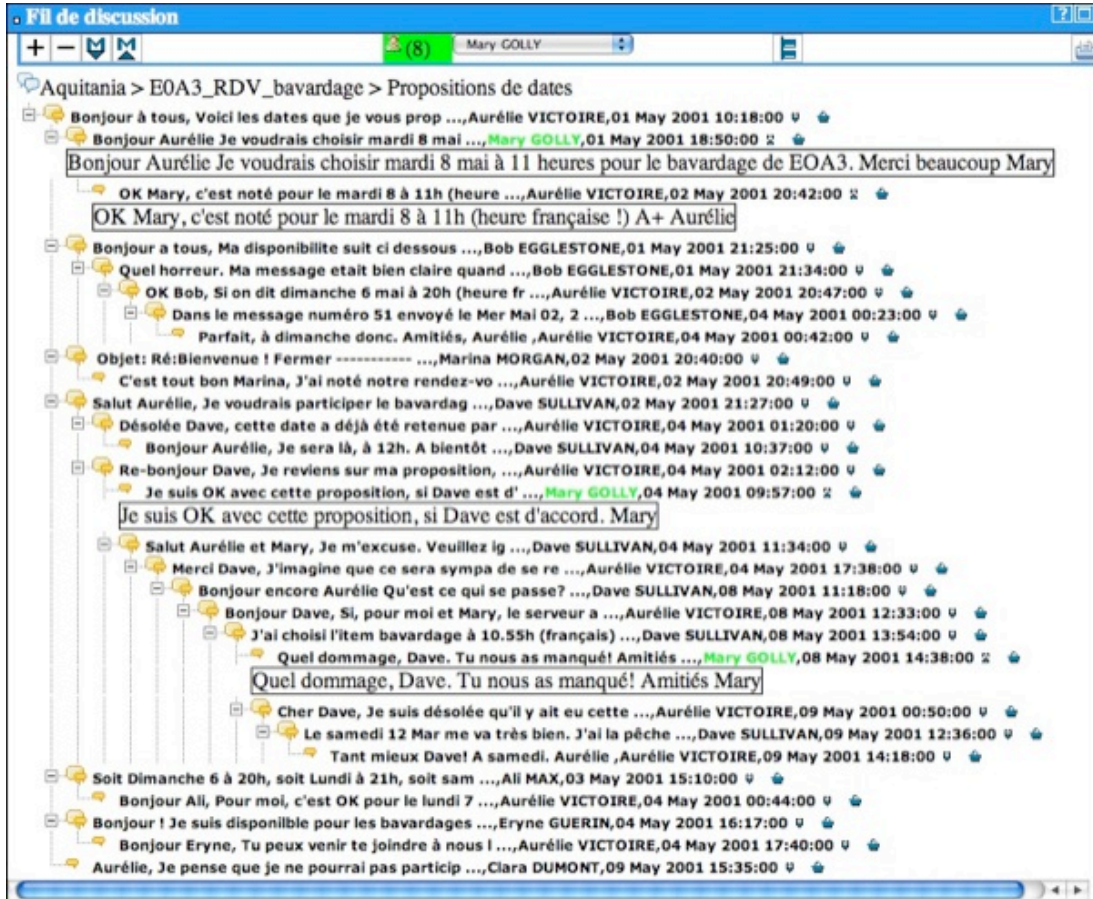


Figure 4, Affichage des interventions dans leur contexte

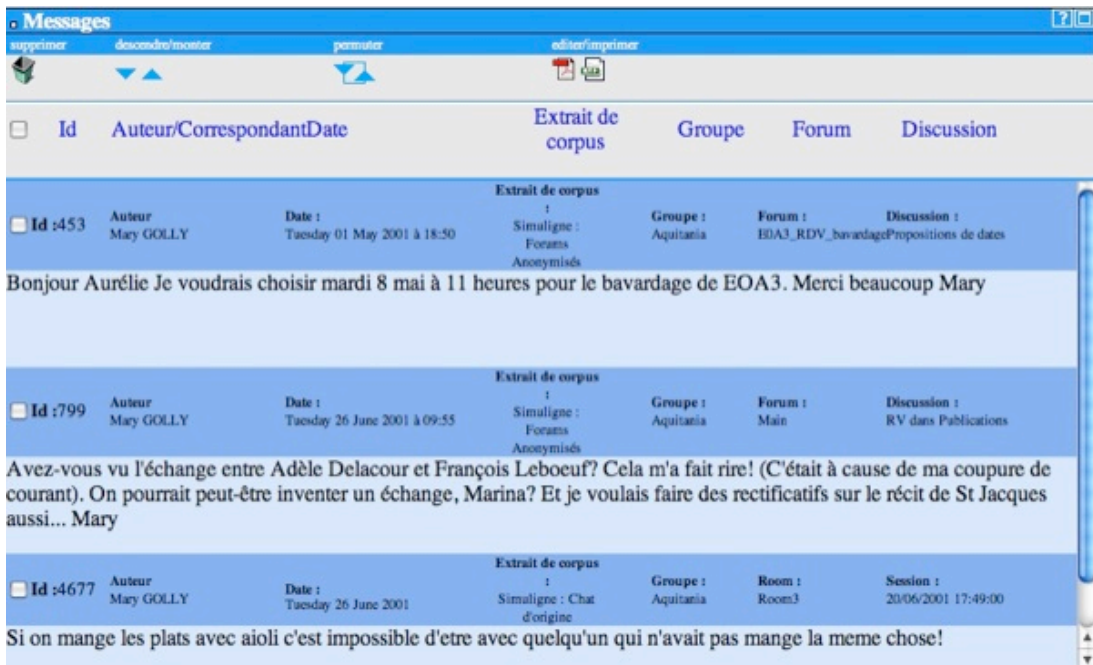


Figure 5, Regroupement d'énoncés d'origines diverses dans le "panier"

Format de description des échanges

Les fonctionnalités de manipulation homogène de corpus hétérogènes s'appuient sur une représentation commune et unique des différentes structures de discussion concernées. Chaque intervention (message) est définie par son auteur, sa date de dépôt, le corps du texte et éventuellement l'intervention "mère" dont

elle dépend dans le cas d'échanges structurés (forum par exemple).

Le format de description des corpus de discussion (Figure 6) s'appuie sur une structure arborescente qui supporte un nombre variable de niveaux intermédiaires entre le niveau supérieur de corpus et le niveau terminal de message. Le principe est de décrire récursivement l'ensemble du corpus à travers ce modèle et de disposer d'un fichier d'échange normalisé (format XML) interprétable par le système de visualisation. Par exemple, un corpus de forum extrait de Simuligne est progressivement décrit à travers les 3 niveaux intermédiaires suivants (Figure 7) : groupe de participants (groupe), type de discussion (forum), et fil de discussion (discussion). Par contre, 2 niveaux intermédiaires suffisent à décrire un corpus de clavardage extrait de Moodle : thématique de discussion et session d'échanges. Ce modèle permet d'unifier l'accès à des interventions de différents types produites par différents scénarios sur différentes plateformes de formation en ligne.

```

<ELEMENT activités (activité*)>
  <ELEMENT activité (forum*)>
    <!ATTLIST activité nom CDATA #REQUIRED>
  <ELEMENT forum (discussion*)>
    <!ATTLIST forum nom CDATA #REQUIRED>
  <ELEMENT discussion (message*)>
    <!ATTLIST discussion nom CDATA #REQUIRED>
  <ELEMENT message (texte, attachement*)>
    <!ATTLIST message idmess ID #REQUIRED
      mess_pere IDREF #IMPLIED
      jour CDATA #REQUIRED
      date CDATA #REQUIRED
      heure CDATA #REQUIRED
      nom_auteur CDATA #REQUIRED>
  <ELEMENT texte (#PCDATA)>
  <ELEMENT attachement (#PCDATA)>
  <!ATTLIST attachement reference CDATA #REQUIRED
    nom CDATA #REQUIRED
    type CDATA #REQUIRED>

```

Figure 6, DTD de corpus de discussion

Techniquement, il est nécessaire de transformer les corpus d'origine afin qu'ils respectent ce format de description générique. Si la plateforme d'origine permet d'exporter des fichiers respectant un format d'échange à balises de type XML, l'adaptation au format ViCoDiLi est simple et automatisable. Pour le corpus Simuligne, par contre, le fichier de description XML a été entièrement reconstruit à partir de la nomenclature des répertoires et fichiers WebCT fournie par les porteurs du projet ICOGAD (Reffay et al., 2002).

Dans tous les cas, seules les données "utiles" à l'affichage et à la manipulation des échanges sont reportées dans le document de description XML (Figure 7).


```

<?xml version="1.0" encoding="ISO-8859-1"?>
<groupes>
<groupe nom="Gallia">
<forum nom="Prem : 02-29/04/01">
<discussion nom="Salutations">
<message date="3 4 2001" heure="23h10" idmess="i3" jour="Mardi"
  mess_pere="inull" nom_auteur="Sabrina GLISTON">Salut tout le monde.
  J'espere que tout va bien marcher, on ne sait jamais avec les ordinateurs!! Je
  serais heureuse de vous tous connaitre et de commencer un dialogue avec
  vous. Amities, Sabrina
</message>
<message date="4 4 2001" heure="15h56" idmess="i5" jour="Mercredi"
  mess_pere="i3" nom_auteur="Sylvie TESSIER">Bonjour Sabrina ! Nous
  vous souhaitons la bienvenue sur WebCT ! Nous sommes heureux de voir
  que vous êtes entrée dans votre groupe sans trop de difficultés. Nous
  attendons à présent la venue de vos partenaires. Cordialement. Sylvie.
</message>
<message date="5 4 2001" heure="22h9" idmess="i7" jour="Jeudi"
  mess_pere="inull" nom_auteur="Tony JENKINS"> Salut tout le monde! je
  fais annonce de mon arrivee sur WebCT. J'espere que l'experience sera
  tres utile pour tous et toutes et j'attends avec impatience le plaisir de
  rencontrer les autres membres de la groupe. Amities Tony JENKINS
</message>
</discussion>
...

```

Figure 7, Extrait de fichier XML de forum dans Simuligne

Module d'anonymisation

Le projet ViCoDiLi a dès le départ intégré une dimension d'anonymisation permettant de respecter l'identité des participants aux formations. Le processus d'anonymisation s'appuie sur la définition des données d'identité à protéger, sur une table de correspondance attribuant un masque de remplacement à chaque donnée d'identité, et sur un traitement en plusieurs phases du corpus (Reffay et Teutsch, 2007). L'environnement ViCoDiLi propose un module d'anonymisation qui intervient sur le fichier XML de description des discussions et y modifie les identités.

Concernant l'identité de la personne dans le corpus, nous distinguons, d'une part les composants d'identification manipulés par la plateforme de formation, et d'autre part les éléments de désignation personnelle utilisés à l'intérieur des messages eux-mêmes. Les premières données font référence (directement ou indirectement) aux acteurs : nom, prénom, ou identifiant par exemple. Elles apparaissent sous la forme d'une chaîne de caractères constante facile à repérer et à remplacer de façon automatique. C'est le cas du nom de l'auteur d'un message posté dans un forum ou des pseudonymes qui précèdent un tour de parole dans un clavardage. Les secondes données se trouvent dans le corps des interventions produites par les acteurs eux-mêmes : signature, interpellation, réponse ou référence à un ou plusieurs des intervenants. Le processus d'anonymisation en lui-même consiste à appliquer les modifications dans le corpus d'origine (fichier XML) en deux phases : modification des identifiants des acteurs dans les en-têtes des interventions puis modification dans le corps des interventions. Ce processus transforme le contenu sans altérer la structure XML, ce qui permet à ViCoDiLi de visualiser également le nouveau corpus.

Bilan et perspectives

La [figure 8](#) résume le processus d'intégration d'un corpus à ViCoDiLi. Un premier corpus est construit à partir de la structure de données de la plateforme d'origine. Ce corpus est ensuite anonymisé avant d'être mis à disposition des chercheurs sur la plateforme ViCoDiLi : la base de données créée à partir du fichier XML est interrogée et visualisée dynamiquement à partir du navigateur.

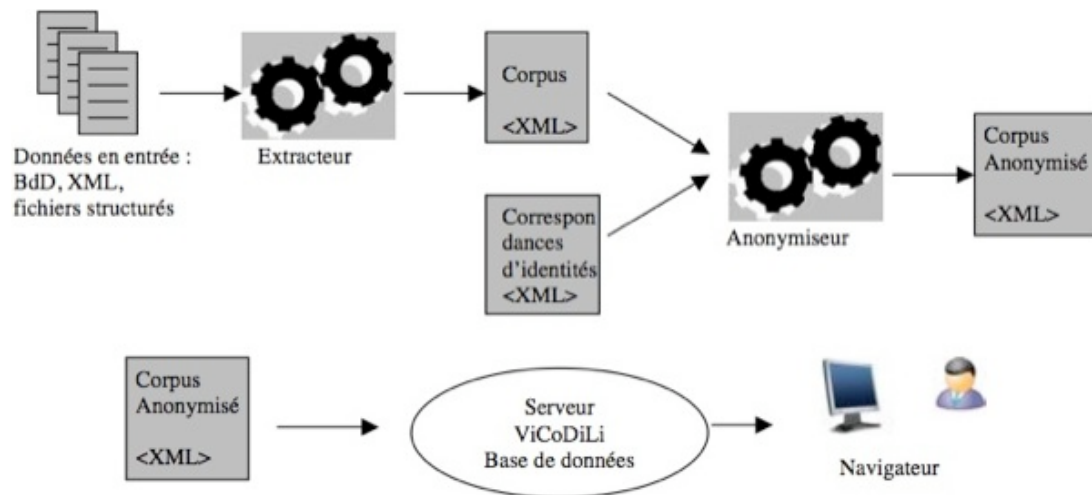


Figure 8, Processus d'intégration de corpus à ViCoDiLi

Grâce à ce processus, le projet ViCoDiLi a permis de restaurer entièrement le corpus Simuligne où une cinquantaine de personnes ont échangé plus de 12 000 interventions. Ce corpus est donc à nouveau visible, exploitable et analysable tout en étant indépendant de sa plateforme d'origine.

Le processus de structuration des interventions et d'intégration à l'outil ViCoDiLi a été appliqué à d'autres corpus issus de plateformes différentes, en particulier deux corpus d'échanges (quelques centaines de messages sur forums dédiés à des activités d'échanges de points de vue et de réflexion) produits lors d'une formation de master en didactique des langues et extraits de la plateforme Moodle via un fichier de format XML.

Ces différentes expérimentations d'intégration de corpus permettent de valider le principe de distinction entre contenus d'interactions (interventions, échanges) et supports d'interaction (plateforme d'origine). Une fois transformés dans le format de référence et intégrés à ViCoDiLi, les corpus vérifient trois des critères de qualité préconisés par la communauté de recherche sur les interactions en ligne : pérennité, lisibilité, homogénéité. La validité reste dépendante de la formation étudiée.

L'outil ViCoDiLi a néanmoins besoin d'être enrichi de nouvelles fonctionnalités avant d'être expérimenté en réelle situation d'analyse. Les retours sur l'utilisation de l'interface ont souligné la faiblesse des modalités d'expression des recherches et l'impossibilité de gestion des extraits regroupés dans le panier comme un corpus en soi, qui bénéficierait à son tour des modalités de décontextualisation. La méthode de conception entamée en partenariat avec des praticiens doit donc prolonger ses itérations pour faciliter l'évolution de l'artefact vers un instrument réellement modelé par ses utilisateurs avec leurs propres schèmes d'utilisation (Rabardel, 2005). Au-delà de la vérification de l'utilisabilité du logiciel, les questions de l'utilité et de l'acceptabilité (Tricot et al., 2003) pourront alors se poser.

5. Conclusion

Dans le cadre d'une démarche de diffusion de corpus de formations médiatisées, cet article a présenté les différentes facettes d'une recherche en ingénierie d'outils d'assistance à l'analyse des échanges en ligne. Celle-ci a abouti à la création d'un outil permettant d'accéder à des traces d'échanges indépendamment de leur plateforme d'origine. ViCoDiLi s'appuie sur un format de description des données qui facilite et harmonise l'accès à des traces d'échanges issues de contextes techniques et pédagogiques hétérogènes. Les fonctionnalités de l'outil permettent néanmoins de manipuler des échanges dans une structure proche de celle d'origine. Cet ensemble de propositions a été validé à l'occasion de la restauration du corpus Simuligne. Par la suite, afin de faciliter la compréhension du contexte de production des échanges, l'outil devra s'enrichir de moyens de mise en relation avec les autres composants d'un corpus d'apprentissage "complet" (Reffay et al., 2008) (dans ce numéro) tels que le scénario pédagogique, le contexte de la session de formation ou les retours d'enquête auprès des participants (Mangenot, 2006).

Les questionnements prolongeant cette recherche concernent la pratique de l'analyse des interactions

elle-même. Au-delà de la mise à disposition d'un outil logiciel facilitant l'accès aux échanges médiatisés, la démarche d'instrumentation de l'activité d'analyse nécessite de tenir compte de la mise en œuvre de l'outil et des pratiques réelles des usagers. Ainsi Savoie-Zajc (Savoie-Zajc, 2000) souligne que la structure même des logiciels d'aide à l'analyse "influence la logique du chercheur" (p. 116). Lors de l'élaboration de ViCoDiLi, il s'agissait de concevoir un outil générique d'assistance à l'analyse mais, de fait, nous avons été amenés à faire des choix ayant pour conséquence de particulariser les approches méthodologiques permises par l'utilisation de l'outil. Ainsi, en choisissant le message comme unité d'analyse minimale ou en permettant aux chercheurs d'avoir accès à certains types de vues sur les échanges en ligne, l'outil ViCoDiLi peut induire un certain type de navigation dans les données et ainsi influencer la pratique de l'analyse de ces données. La conception d'un tel outil ne représente donc qu'une étape dans l'accès aux échanges en ligne, dans la diffusion des corpus et dans l'appropriation de ces corpus par les chercheurs.

¹ Pour l'analyse des interactions en ligne, voir le récent colloque Épal "Échanger pour apprendre en ligne" : <http://w3.u-grenoble3.fr/epal/>

² ODIL : Outils et Didactique pour les Interactions en Ligne : projet (2004-2007) de l'ACI "Éducation et formation", piloté par Grenoble 3 (F. Mangenot) et regroupant 7 laboratoires. Projet centré sur l'analyse qualitative des interactions en ligne et sur les outils qui instrumentent ces interactions. Résumé du projet : <http://w3.u-grenoble3.fr/epal/image/resume-odil.pdf>

³ Mulce : Multimodal Learning Corpus Exchange <http://mulce.univ-fcomte.fr>

⁴ ViCoDiLi : Visualisation de Contenus de Discussions en Ligne <http://demolium.univ-lemans.fr>

⁵ En classe de langue étrangère, la simulation globale est une pratique pédagogique qui consiste à proposer à un groupe d'élèves un projet autour de la création d'un univers de référence (fonctionnant comme thème et comme lieu) et d'identités fictives.

⁶ Texte de T. Chanier et C. Reffay au symposium "Corpus d'apprentissage en ligne" du colloque Épal (juin 2007) : http://mulce.univ-fcomte.fr/epal_symposium/actes/Mulce.pdf

BIBLIOGRAPHIE

BOISSIÈRE P., SCHADLE I. (2006). Proposition d'un cadre méthodologique d'évaluation des systèmes d'assistance à la saisie de textes : Applications aux systèmes Sibylle et VITIPI. Actes de *Handicap 2006*, Paris, p. 149-154.

BRUILLARD É. (2006). Le forum de discussion : un cas d'école pour les recherches en EIAH. Éditorial du numéro spécial "Forum de discussion en éducation", *Revue STICEF*, Volume 13, <http://sticef.org> (consulté en juin 2008)

CHARAUDEAU P., MAINGUENEAU D. (2002). *Dictionnaire d'analyse de discours*, Paris. Seuil.

CHOQUET C., LUENGO V., YACEF K. (2007). Usage Analysis of Learning Systems *Journal of Interactive Learning Research special issue* - Vol 18, n° 2, AACE Publications, 174 pages.

DE WEVER B., SCHELLENS T., VALCKE M., VAN KEER H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *In Computers and Education*, 46, p. 6-28.

DEJEAN-THIRCUIR C. (2008, à paraître). Modalités de collaboration entre étudiants et constitution d'une communauté dans une activité à distance. In Nissen, E., Blin, F., *revue ALSIC, Apprentissage des Langues et Systèmes d'Information et de Communication*, vol.12. <http://alsic.org> (consulté en juin 2008)

DEJEAN-THIRCUIR C., MANGENOT F. (2006). Pairs ou tutrices ? Pluralité des positionnements d'étudiants de maîtrise FLE lors d'interactions en ligne avec des apprenants australiens, *Le Français dans le monde*, Recherches et applications n° 40, Paris, CLE International, p. 75-87.

DEGACHE C. (2006). Aspects du contrat didactique dans une formation plurilingue ouverte et à distance ». *Le Français dans le monde*, Recherches et applications n° 40, Paris, CLE International, p. 58-74.

DILLENBOURG P., BAKER M., BLAYE A. MALLEY C. (1996). The evolution of research on collaborative learning. In E. Spada & P. Reiman (Eds) *Learning in Humans and Machine: Towards an interdisciplinary learning science*, Oxford: Elsevier, p. 189-211.

- FUSTENBERG G., ENGLISH K. (2006). Communication interculturelle franco-américaine via internet : Le cas de Cultura. *Le Français dans le monde*, Recherches et applications n° 40, Paris, CLE International, p 178-191.
- GEORGE S., BOTHOREL C. (2006). Conception d'outils de communication spécifiques au contexte éducatif, *Revue STICEF*, Volume 13, <http://sticef.org> (consulté en juin 2008)
- GIBBS W. J., OLEXA V., BERNAS R.S. (2006). A visualization tool for managing and studying online communication. *Educational Technology & Society*, 9 (3), 232-243.
- HEER J, BOYD D. (2005). Vizter: Visualizing Online Social Networks. In: *InfoVis 2005*, IEEE Symposium on Information Visualization. <http://jheer.org/publications/2005-Vizster-InfoVis.pdf> (consulté en juillet 2008)
- HENRI F., CHARLIER B. (2005). L'analyse des forums de discussion pour sortir de l'impasse. In Sidir Mohamed, Bruillard Eric et Baron Georges-Louis (coord.), *Symposium Symphonic, formation et nouveaux instruments de communication*, Amiens, http://www.dep.u-picardie.fr/sidir/articles/henri_charlier.htm (consulté en juin 2008)
- HENRI F., PERAYA D., CHARLIER B. (2007). La recherche sur les forums de discussion en milieu éducatif : critères de qualité et qualité des pratiques, *Revue STICEF*, Volume 14, 2007, <http://sticef.org> (consulté en juin 2008)
- HERRING S. C. (2004). Computer-mediated discourse analysis: An approach to researching online behavior. In: S. A. Barab, R. Kling, and J. H. Gray (Eds.), *Designing for Virtual Communities in the Service of Learning*, New York: Cambridge University Press, p. 338-376.
- JEANNOT L., VETTER A., CHANIER T. (2006). Repérage des stratégies des apprenants et du tuteur dans un environnement audio-graphique synchrone, *Le Français dans le monde*, Recherches et applications n° 40, Paris, CLE International, p. 151-163.
- LAMY M-N. (2006). Conversations multimodales : l'enseignement-apprentissage de l'oral à l'heure des écrans partagés. *Le Français dans le monde*, Recherches et applications n° 40, Paris, CLE International, p. 129-138.
- MACKAY W., FAYARD A.L. (1997). HCI, Natural Science and Design: A Framework for Triangulation Across Disciplines, In *DIS'97: Designing Interactive Systems*. Amsterdam, Holland, p. 223-234.
- MANGENOT F. (2006). Analyser les interactions pédagogiques en ligne, pourquoi, comment ? *Intercompreensão*, 13, Comunicação electrónica em contextos de educação linguística. Teorias e praticas. Editions Cosmos / Escola Superior de Educação de Santarem, Portugal, p. 11-28.
- MANGENOT F. (2008). Pratiques pédagogiques instrumentées et propriétés des outils : le cas des forums, *Revue STICEF*, Volume 15, 2008, <http://sticef.org> (consulté en juin 2008)
- MANGENOT F., TANAKA S. (2007). Les enseignants de langue comme médiateurs entre deux cultures dans les interactions en ligne : le cas d'un échange franco-japonais, *colloque Épal "Échanger Pour Apprendre en Ligne"*, Grenoble, France, <http://w3.u-grenoble3.fr/epal/> (consulté en juin 2008)
- MAZZA R., DIMITROVAB V. (2007). CourseVis: A graphical student monitoring tool for supporting instructors in web-based distance courses. *International Journal of Human-Computer Studies*, Volume 65, Issue 2, 2007, p. 125-139
- NORMAN D.A. (1988). *The Design of Everyday Things*. New York, NY: Basic Books.
- OLLAGNIER-BELDAME M., MILLE A. (2008, à paraître). E-learning : Tracer l'activité des apprenants pour favoriser leur appropriation des systèmes ? *Revue du Management Technologique*, 15(3).
- PERAYA D. (2005). Axes de recherches sur les analyses de communication dans les forums Notes pour un texte, *Symposium Symphonic, formation et nouveaux instruments de communication*, Amiens, <http://www.dep.u-picardie.fr/sidir/articles/peraya.htm> (consulté en juin 2008)
- PREECE J., MALONEY-KRICHMAR D. (2005). Online communities: Design, Theory, and Practice. *Journal of Computer-Mediated Communication*, 10(4), <http://jcmc.indiana.edu/vol10/issue4/preece.html> (consulté en juin 2008)
- RABARDEL P. (1995). *Les hommes et les technologies, approche cognitive des instruments contemporains*, Colin, Paris.
- REFFAY C., CHANIER T., NICOLET J. (2002). Produire ensemble pour apprendre : expérience d'une simulation globale en ligne. *Actes du Colloque national Apprendre avec l'Ordinateur*, Bordeaux, France, p. 24.
- REFFAY C., LANCIERI L. (2006). Quand l'analyse quantitative fait parler les forums de discussion, *Revue STICEF*, Volume 13, p 255-288. <http://sticef.org> (consulté en juin 2008)
- REFFAY C., TEUTSCH P. (2007). Anonymisation de corpus réutilisables : Masquer l'identité sans altérer l'analyse des interactions. Poster In: *Actes de EIAH'2007*, Lausanne (Suisse). <http://edutice.archives-ouvertes.fr/edutice-00158877> (consulté en juin 2008)
- REFFAY C., CHANIER T., NORAS N., BETBEDER M.-L. (2008), Contribution à la structuration de corpus d'apprentissage pour un meilleur partage en recherche, *Revue STICEF*, Volume 15, 2008, <http://sticef.org> (consulté en juin 2008)

ROURKE L., ANDERSON T., GARRISON R., ARCHER W. (2001). Methodological Issues in the Content Analysis of Computer Conference Transcripts. *In International Journal of Artificial Intelligence in Education*, 12, p. 8-22.

SAVOIE-ZAJC L. (2000). L'analyse des données qualitatives : Pratiques traditionnelles et assistées par le logiciel NUD*IST. *Recherches Qualitatives*, 21, p. 99-123.

TEASLEY S., ROSCHELLE J. (1993). Constructing a Joint Problem Space: The Computer as a Tool for Sharing Knowledge, *in C. O'Malley (Ed.), Computer-supported collaborative learning*, New-York, Springer-Verlag, 31 p. <http://ctl.sri.com/publications/downloads/JointProblemSpace.pdf> (consulté en juillet 2008).

TESCH R. (1990). *Qualitative Research: Analysis Types and Software Tools*, New-York, The Falmer Press.

TSELIOS N., AVOURIS N., KOMIS V. (2008). The Effective Combination of Hybrid Usability Methods in Evaluating Educational Applications of ICT: Issues and challenges, *Education and Information Technologies Journal*, Springer-Verlag, vol. 13 (1), pp. 55-76.

TRICOT A, PLÉGAT-SOUTJIS F., CAMPS J.-F., AMIEL A., LUTZ G, MORCILLO A. (2003). Utilité, utilisabilité, acceptabilité : interpréter les relations entre trois dimensions de l'évaluation des EIAH, *Conférence EIAH 2003*, Strasbourg, France, p. 391-402.

XIONG R., DONATH J. (1999). PeopleGarden: creating data portraits for users, *Proceedings of the 12th annual ACM symposium on User interface software and technology*, Asheville, North Carolina, United States, p. 37-44.

■ À propos des auteurs

Philippe TEUTSCH est maître de conférences en informatique à l'Université du Maine, chercheur en EIAH au LIUM. Ses recherches portent sur l'instrumentation des activités de suivi et d'analyse de situations de formation médiatisées : interfaces de tutorat, outils de visualisations de traces issues de formations en ligne.

Adresse : LIUM, Université du Maine, Avenue Messiaen, 72085, Le Mans cedex 9, France

Courriel : Philippe.Teutsch@univ-lemans.fr

Toile : <http://www-lium.univ-lemans.fr/~teutsch/>

Charlotte DEJEAN-THIRCUIR est maître de conférences en sciences du langage à l'université du Maine et fait partie du LIUM. Elle étudie, dans une perspective linguistique et didactique, les interactions verbales en contexte d'apprentissage des langues et en contexte de formation de formateurs, en présentiel et en ligne. Elle s'intéresse tout particulièrement aux activités collaboratives et aux échanges entre pairs.

Adresse : UFR Lettres Langues et Sciences Humaines, Université du Maine, Avenue Olivier Messiaen, 72085 Le Mans cedex 9, France

Courriel : charlotte.dejean@univ-lemans.fr

Francis BANGOU est professeur adjoint à l'Université d'Ottawa en didactique des langues secondes. Ses travaux portent principalement sur l'étude des processus d'appropriation des technologies émergentes dans un contexte didactique. Il a eu l'opportunité de travailler pendant deux ans au sein du LIUM en tant que chercheur associé.

Adresse : Université d'Ottawa, Faculté d'éducation, 145 Jean-Jacques Lussier, Ottawa, ON K1N 6N5, Canada

Courriel : fbangou@uottawa.ca

Référence de l'article :

Philippe TEUTSCH, Francis BANGOU, Charlotte DEJEAN-THIRCUIR, Faciliter l'accès aux échanges en ligne et leur analyse, le cas de ViCoDiLi, *Revue STICEF*, Volume 15, 2008, ISSN : 1764-7223, mis en ligne le 17/10/2008, <http://sticf.org>

© Revue Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation, 2008