

ReaderBench : Un outil d'aide à l'analyse de discussions philosophiques

Philippe Dessus, Jean-Pascal Simon, Mihai Dascalu, Stefan Trausan-Matu

▶ **To cite this version:**

Philippe Dessus, Jean-Pascal Simon, Mihai Dascalu, Stefan Trausan-Matu. ReaderBench : Un outil d'aide à l'analyse de discussions philosophiques. Jean-Pascal Simon; Michel Tozzi. Paroles de philosophes en herbe : Regards croisés de chercheurs sur une discussion sur la justice en CM2, UGA Éditions, pp.151-171, 2017, 2377470084. hal-01587118

HAL Id: hal-01587118
https://hal.science/hal-01587118

Submitted on 13 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

READERBENCH : UN OUTIL D'AIDE A L'ANALYSE DE DISCUSSIONS PHILOSOPHIQUES

Philippe Dessus*, Jean-Pascal Simon**, Mihai Dascalu*** & Stefan Trausan-Matu***

* LSE (EA 602), Univ. Grenoble Alpes, France ** LIDILEM (EA 609), Univ. Grenoble Alpes, France *** RACAI, Université "Politehnica", Bucarest, Roumanie

Introduction

Les recherches conduites depuis plusieurs années autour des discussions à visée philosophique ont permis de collecter des corpus oraux multimodaux assez conséquents[1]. Si on peut se réjouir de cette abondance de données on sait combien leur traitement est lourd et il est à craindre que du recueil à l'analyse la route soit longue, car il faut passer par la transcription des propos, d'éléments non linguistiques (gestes, attitudes, etc.), pour obtenir *in fine* plusieurs centaines de pages qu'il faut ensuite analyser.

Cette même problématique se retrouve dans les travaux de recherche sur l'apprentissage collaboratif assisté par ordinateur (CSCL pour *Computer-Supported Collaborative Learning*). Les systèmes étudiés sont de plus en plus répandus et sophistiqués (Stahl, 2006) et sont utilisés dans de nombreux contextes, tels l'enseignement hybride (Dessus, Trausan-Matu, et al., 2011) ou les MOOC (Massive Open Online Courses, Chaturvedi, Goldwasser, & Daumé, 2014). Ils recueillent et traitent des conversations par *chat* ou forum, pour lesquelles Trausan-Matu (2010) mentionne que la durée d'une analyse dépasse largement la durée de la conversation elle-même. Il se trouve que les techniques d'analyse automatique de la langue rendent possible une analyse quasi-immédiate des productions des élèves, ce qui permet des rétroactions rapides pouvant leur être utiles pour améliorer leur apprentissage (Dessus, Lemaire, et al., 2011). Cette analyse partiellement automatisée peut également être d'un grand intérêt pour les enseignants et les chercheurs, en effet, le traitement de données orales est chronophage. Ainsi, Mu *et al*. (2012), par exemple, signalent que l'on peut utiliser jusqu'à 25 % des ressources totales d'un projet de recherche pour une analyse de conversations selon plusieurs dimensions.

Notre propos, dans ce chapitre, est d'explorer l'utilisation d'un système d'analyse automatique de productions à des fins d'apprentissage, *ReaderBench*. Le contexte n'est pas,

[1] Le corpus *Philosophèmes*, qui a été recueilli dans le cadre de plusieurs projets soutenus financièrement par les MSH de Lorraine (dirigé par Specogna et Halté en 2009), et de Clermont-Ferrand (dirigé par Auriac-Slusarczyk et Blasco-Dulbecco en 2010), et région Auvergne dans le cadre d'un appel à projet structurante en Sciences humaines et sociales. Un autre corpus en cours de constitution au LiDiLEM de l'université Grenoble Alpes collecté dans le cadre du projet Philéduc (soutenu par AGIR 2013) et d'une thèse financée par la Région Auvergne Rhône-Alpes.

comme dans ses précédentes utilisations, un apprentissage médiatisé par ordinateur, mais une une nouvelle configuration dans laquelle il n'a pas encore été testé : une interaction orale enseignant-élèves. Il ne s'agit pas, bien sûr, d'envisager cette étude comme le remplacement d'analyses humaines, mais comme leur complément. Dans cette contribution nous étudions une partie des fonctionnalités de ce logiciel : analyse de mots-clés et des thèmes abordés (dans la globalité de la DVDP et par locuteur), les proximités thématiques des contributions des différents locuteurs ; l'apport de chacun à la discussion et la dimension dialogique.

ReaderBench, un atelier multilingue d'analyse automatique de textes

ReaderBench (Mihai Dascalu, Dessus, Bianco, Trausan-Matu, & Nardy, 2014) est un outil d'analyse automatique de productions écrites liées à l'apprentissage qui fait l'objet, depuis 2010, d'une conception et d'un développement, initialement franco-roumain, puis étendu à des chercheurs de nombreux pays (États-Unis, Allemagne, Italie, Pays-Bas). Il est multilingue (anglais, français, italien, espagnol, roumain, néerlandais), et utilise des méthodes de traitement automatique de la langue éprouvées pour analyser la cohésion entre composants du texte (phrases, tours de parole, documents, etc.). Les deux principales méthodes utilisées sont LSA (Latent Semantic Analysis) et LDA (Latent Dirichlet Allocation).

LSA (Landauer & Dumais, 1997; Manning & Schütze, 1999) représente sous la forme d'un espace vectoriel les relations entre termes et les documents (paragraphes) qui les contiennent, en se fondant sur l'analyse de leurs co-occurrences. Cela permet d'évaluer la similarité sémantique inter-termes et entre termes et documents. LDA (Blei, Ng, & Jordan, 2003) permet de réaliser une identification de thèmes *via* un mécanisme d'inférence probabiliste de structures thématiques dans les documents. Ces deux méthodes sont de type « paquets-de-mots » (« *bag of words* »), et ne tiennent pas compte de l'ordre des mots dans les phrases et documents. Cela est bien entendu une grande approximation, mais qui n'est pas trop gênante puisqu'il s'agit de récupérer les mots-clés principaux et des indices de similarités entre ces mots-clés, apparaissant dans un grand corpus.

Les modèles utilisés dans *ReaderBench*

ReaderBench a été conçu par la mise en œuvre de deux modèles différents et complémentaires (voir M. Dascalu, Trausan-Matu, McNamara, & Dessus, 2015). Le modèle de la cohésion, opérationnalisée comme la mesure de similarité moyenne entre entités du texte (mots, phrases, tours de parole, conversation complète). Un graphe de cohésion entre tours de

paroles est construit, composé d'un nœud central (la discussion complète) décomposé en tours de paroles, puis phrases, les liens entre ces nœuds symbolisent des relations de cohésion fortes. Ce mécanisme permet d'évaluer la contribution de chaque participant en rapport avec la discussion complète, mais aussi avec les contributions des autres participants.

Un autre modèle, plus général, utilisé pour rendre compte des relations sémantiques entre les différentes contributions, considère, dans la lignée des travaux de Bakhtine (1981; voir aussi Linell, 2009), que toute interaction, communication, ou même pensée, est fondée sur le dialogue. Cela permet d'étudier la manière dont les "voix" s'organisent dans la discussion : une voix est un point de vue, un thème discuté, qui survient à différents moments de la discussion et résonne, interagit avec d'autres. Pour rendre compte de cette interaction, les notions de "multivocalité" et de "polyphonie" ont été formulées. La multivocalité rend compte de la multitude des voix et la manière dont elles sont reliées thématiquement les unes aux autres. La polyphonie ajoute à cette dernière la notion de cohérence : les voix peuvent résonner plus ou moins en cohérence les unes avec les autres.

Utilisation de *ReaderBench* pour analyser les forums de discussion

Comme déjà signalé, *ReaderBench* a été testé dans le cadre de l'évaluation automatique de forums ou de *chats* (M. Dascalu, Trausan-Matu, Dessus, & McNamara, 2015; Nistor, Dascalu, Stavarache, Tarnai, & Trausan-Matu, 2016). L'étude de Dascalu et ses collègues a consisté à faire évaluer 10 conversations de *chat* (messagerie instantanée), dans lesquelles des étudiants en informatique, par petits groupes de quatre, discutaient des avantages et inconvénients des outils de communication collaborative. Des étudiants de licence et master en informatique ($n = 110$) ont ensuite évalué chaque participant à la conversation selon une échelle individuelle de participation et globalement, de collaboration. Les corrélations intra-classe obtenues sont très élevées (supérieures à .79). Le travail de Nistor et ses collègues (2016) a consisté à examiner 22 communautés différentes de blogueurs (sur différentes plateformes, comme *WordPress* ou *Blogspot*) : intégrative ou non intégrative, ce critère étant déterminé après diffusion d'une requête à participer à une enquête et la communauté y répondant positivement a été catégorisée comme intégrative. Ensuite, les conversations de ces communautés, comportant au total plus de 8 000 participants, ont été analysées avec *ReaderBench*, selon certains indicateurs comme la participation. Une analyse en grappes (*clusters*) révèle que les communautés qualifiées d'intégratives ont un indice de réponse communautaire (évaluée par la longueur moyenne des fils de discussion) plus important que celles qualifiées de non intégratives. De plus, la proximité moyenne des participants, évaluée

automatiquement *via* leurs contributions respectives est plus importante dans les communautés intégratives.

ReaderBench réalise quatre types différents de traitements automatiques des tours de parole, pouvant répondre à des questions et des problèmes rencontrés fréquemment par les chercheurs ou enseignants :

- la contribution des participants à la progression de la réflexion ;
- l'identification des principaux mots-clés de la discussion ;
- l'analyse des proximités thématiques des tours de parole des participants ;
- la mise au jour de "voix" tout au long de la discussion.

Pré-traitement du corpus

Le corpus est constitué de productions orales de participants à la discussion. Pour que *ReaderBench* réalise une analyse valide, il a fallu réaliser un certain nombre d'adaptations de la transcription initiale des échanges, qui a été faite dans un premier temps à l'aide du logiciel d'annotations de vidéos *ELAN* (https://tla.mpi.nl/tools/tla-tools/elan/) (Sloetjes & Wittenburg, 2008), en étiquetant les fichiers audio et vidéo de manière à obtenir une transcription alignée sur le son et l'image. On ne parle pas comme on écrit et depuis longtemps on sait que les unités langagières de l'oral et de l'écrit ne sont pas les mêmes. Ainsi, la notion de phrase est une réalité typographique et l'oral s'analyse plutôt en intervention, groupe de souffle[2]. Les analyses en grilles syntaxiques du corpus présentées dans ce même ouvrage par Blasco et Lebas-Fraczak (p. référence à mettre à jour) illustrent cette réalité.

Tout d'abord, il nous a fallu adapter la transcription initiale pour que le logiciel puisse la traiter dans les meilleures conditions. Pour cela nous avons dû supprimer toutes les indications de posture ou d'attitude comme : {volte-face}, les épiphénomènes (Tousseul, 2005) : {coupure son}, {rires}, etc. ; les incises quand il s'agissait simplement d'un retour positif ou encourageant et pas réellement d'un tour de parole ; les indications de prosodie, d'allongement, etc. Enfin l'orthographe a été renormalisée : les phénomènes d'élision des « e » supprimés, la prononciation [ja] transcrite « il y a », *chais : « je sais ». Ainsi la transcription initiale :

> *85.* ***Candice*** *: moi je pense qu'(il) y a plusieurs (en)fin heu // plusieurs sens de c'est pas juste parce que y a / heu (il) y a // c'est pas juste qui peut créer un conflit par exemple heu qui est vraiment dans l'inégalité heu // par exemple heu : // je sais pas trop {rires} // (il) y a aussi heu le c'est pas juste bah un peu heu // capricieux on va dire // (en)fin qu'on*

[2] Pour plus de détails on se reportera aux travaux du Groupe Aixois de Recherche sur la Syntaxe (GARS) et aux nombreux écrits de Claire Blanche-Benveniste, et à Morel et Danon-Boileau (2001)

pense pas trop // qui est plutôt heu // heu c'est pas juste heu, mais en fait c'est juste (en)fin heu voilà et ben après heu ça dépend des situations, mais :

devient :

85. ***Candice*** *: moi je pense qu'il y a plusieurs enfin heu // plusieurs sens de c'est pas juste parce que y a heu il y a // c'est pas juste qui peut créer un conflit par exemple // heu qui est vraiment dans l'inégalité heu // par exemple heu // je sais pas trop // il y a aussi heu le c'est pas juste bah un peu heu // capricieux on va dire enfin // qu'on pense pas trop // qui est plutôt heu // heu c'est pas juste heu, mais en fait c'est juste enfin heu // voilà // et ben après heu // ça dépend des situations //, mais*

La seconde étape a été d'effectuer un balisage afin d'obtenir une structure XML, compatible avec *ReaderBench* ce qui donne (les limites de groupes de souffle, codées // ci-dessus ont été remplacées par des points) :

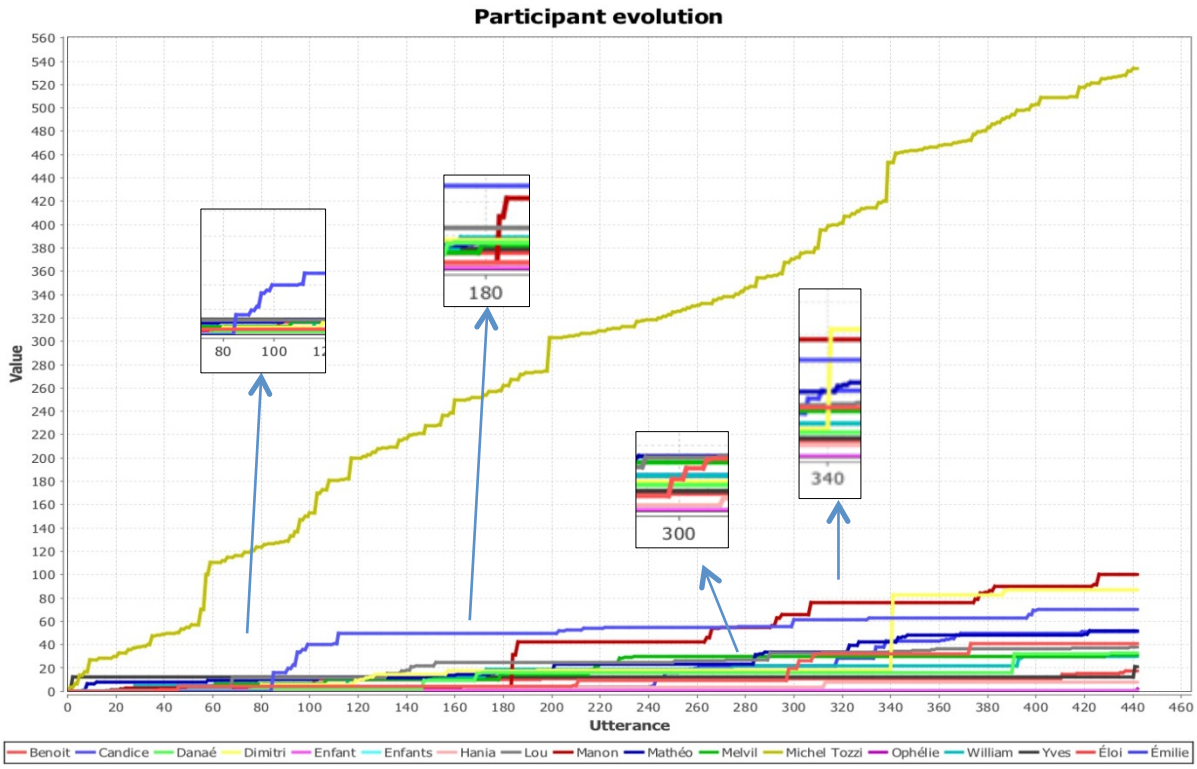
```
<Turn nickname="Candice">
   <Utterance genid="85">
moi je pense qu'il y a plusieurs enfin heu plusieurs sens de c'est pas juste parce que y a heu il y a. c'est pas juste qui peut créer un conflit par exemple. heu qui est vraiment dans l'inégalité heu. par exemple heu :. je sais pas trop. il y a aussi heu le c'est pas juste bah un peu heu. capricieux on va dire enfin. qu'on pense pas trop. qui est plutôt heu. heu c'est pas juste heu, mais en fait c'est juste enfin heu. voilà. et ben après heu. ça dépend des situations. mais
   </Utterance>
</Turn>
```

Quatre types d'analyse de la discussion

Contribution des intervenants à la progression de la réflexion

Une première fonctionnalité de *ReaderBench* est de pouvoir représenter graphiquement le poids cumulé au cours du temps des interventions de chaque participant en relation avec la discussion complète. Cela permet, d'une part, de repérer visuellement à quels endroits de la discussion ont pu avoir lieu des interventions notables (et analysables plus précisément par l'enseignant ou les participants). D'autre part, on peut identifier la ou les intervenants ayant le plus contribué à la discussion.

Figure 1 : Apport cumulé et au cours de la discussion des participants à la discussion.



La Figure 1 montre qu'à certains moments de la discussion des élèves interviennent de manière significative, nous allons en étudier quelques-uns :

- Entre 85 et 100 : Candice (courbe bleue)
- En 184 : Manon (courbe rouge)
- De 298 à 310 : Benoit (courbe orange)
- En 341 : Dimitri (courbe jaune)

Le premier passage correspond à ce que M. Tozzi (*cf.* page **Erreur ! Signet non défini.**) nomme un mini-entretien entre lui et Candice, les interventions de l'animateur permettant à celle-ci d'approfondir et préciser sa pensée[3]. En 85, Candice postule deux formes de « c'est pas juste » : un « c'est pas juste » fondé sur une inégalité et un « c'est pas juste » qui est un caprice.

*85.Candice : moi je pense qu'(il) y a plusieurs (en)fin heu // **plusieurs sens de c'est pas juste** parce que y a / heu (il) y a // c'est pas juste qui peut créer un conflit par exemple heu qui est **vraiment dans l'inégalité** heu // par exemple heu : // je sais pas trop {rires} // (il) y a aussi heu le **c'est pas juste bah un peu heu // capricieux** on va dire // (en)fin*

[3] Il est intéressant de constater que, sans concertation, d'autres analyses de cette même DVDP se soient aussi arrêtée sur ce passage, comme celles de V. Delille, N. Frieden, A. Herla et G. Jeanmart (cf. p. **mettre à jour**), A. Fournel cf. p. **mettre à jour**) et M. Blasco et L. Lebas-Fraczak (cf. p. **mettre à jour**).

qu'on pense pas trop // qui est plutôt heu heu c'est pas juste heu, mais en fait c'est juste (en)fin heu voilà et ben après heu ça dépend des situations, mais:

Pour faire avancer la réflexion, M. Tozzi lui demande d'illustrer son point de vue à l'aide d'exemples : (86) « *est-ce que tu peux nous clarifier ce que tu viens de dire avec des situations précises* », ce qu'elle n'arrive pas à faire dans un premier temps :

93. Candice : heu ben ça dépend des choses en fait parce par exemple quand on est petit des fois on est un peu bah heu j(e) sais pas capricieux on a l'impression que heu ben qu'on peut avoir plus de choses, mais (en)fin heu

(...)

*97. Candice : ben heu // **je sais pas** // (en)fin ça dépend heu de quoi pa(r)ce que c'est juste et pas juste (en)fin {rires}*

Ce n'est en qu'en 112 (dernier palier de la courbe) qu'elle arrive à formuler un exemple qui illustre son point de vue : « *ouais **j'ai trouvé un exemple** par exemple heu ma nièce heu // elle a heu un an et huit mois i(l) m(e) semble heu et heu par exemple j(e) lui ai dit non à:: à:: prendre des objets qui sont petits pour pas qu'elle heu les avale (en)fin et tout // et heu et après(en)fin j(e) (rire) elle sait pas dire c'est pas juste, mais j(e) pense qu'elle le pense //, mais au fond c'est pour la protéger (en)fin c'est pas:: méchamment (en)fin c'est pas:: de l'inégalité parce que après je vais pas faire devant elle heu ouais je joue avec cque t'as pas le droit et tout, mais heu pour la protéger (en)fin*.

Cette séquence est effectivement d'un apport intéressant. M. Tozzi (p. **Erreur ! Signet non défini.** de cet ouvrage) dans son auto-analyse note qu'à ce moment de la discussion la « pensée collective se complexifie », c'est d'ailleurs ce qu'il retourne à Candice et aux autres élèves quand il dit en 98 « *(...) c'est intéressant parce qu'on rentre dans la/ dans la complexité de la chose là hein c'est vrai que // ce que tu disais tout à l'heure c'est que des fois on dit c'est pas juste, mais en fait c'est un/ caprice c'est intéressant hein* ».

Dans le deuxième passage, en 184, Manon intervient pour faire un résumé de ce qui a été dit jusqu'alors : *« bah déjà le sujet c'était pourquoi on dit c'est pas juste quand on est enfant ? // donc c'est Melvil qui a commencé en parlant de sa sœur et // heu elle a des droits et:: qu'il n'a pas heum {rire} // Lou heum c'est plus de l'injustice quand on dit heu:: // c'est pas juste heu:: parce que quand on on l(e) pense vraiment (en)fin voilà // heum Danaé elle a dit quand on dit c'est pas juste c'est qu'on n'est pas d'accord // heu après y a Candice qui a dit il y a plu/:: sieurs c'est pas juste heu y en a un que j'ai pas pu noter et quand on pense heu y a l'injustice qu'on pense pas trop et donc l'autre je j'ai pas pu l'écrire et // il et ça dépend des situations quand on y pense // c'est pas juste heu {rire} // après y a Éloi quand on est petit*

les grands peuvent venir heu // les grands heu peuvent venir quand les petits i(ls) jouent, mais heu il ne pourra pas jouer, car ils ont peur qu'il se fasse mal heu après y a Candice qui a dit heum je crois qu'elle a parlé de sa petite cousine c'est ? ». La significativité de cette intervention consiste dans le fait qu'elle reprend et résume les différents apports des élèves précédents, il n'y a rien de fondamentalement nouveau dans les dires de Manon, d'ailleurs vu son rôle, c'est la secrétaire, elle n'a pas le droit de prendre la parole. De même en 340 (identifié comme le quatrième passage), Dimitri intervient pour faire un résumé de la seconde partie de la discussion.

Dans le troisième passage, de 298 à 310, *ReaderBench* identifie un apport significatif de la part de Benoit. Si en 298 la première formulation de son point de vue est assez approximative « *moi je pensais que peut-être heumm // heu:: heu y a pas heum heu le le le bon:: y a pas assez de:: de parts pour tout le monde en fait y en a trop // c'est pas pareil heumm alors en fait je pensais que:: quelqu'un il prendra la plus petite part et il l/ là où:: i(ls) ont fait le gâteau alors les autres i(ls) vont penser ha ha:: lui il en a moins autrement dit on en a eu plus qu'on que qu'on aurait dû en avoir //, mais après une fois qu'ils sont partis ben ça pourrait être un peu plus // heum juste pour heu pour celui qui en a mangé un peu moins parc(e) qu'ils vont donner heum heu le reste du gâteau à lui // s'il en reste* », ce qui laisse M. Tozzi perplexe : « *est-ce que quelqu'un pourrait aider à comprendre ce que Benoît vient de dire ?* », la pensée de Benoit va se préciser au cours de cette mini-conversation pour aboutir à une définition de ce qui est juste en 310 : « *c'est c'est juste pour tout le monde parc(e) que tout le monde va penser que c'était juste* ».

Ces quelques exemples illustrent bien l'aide que peut apporter *ReaderBench* dans l'analyse des discussions. Le graphe de la contribution des participants à la discussion permet d'identifier assez vite certains moments de la discussion où manifestement « il se passe quelque chose ». Le travail de l'analyste consiste alors à se pencher sur ces moments et de voir en quoi l'apport identifié peut être significatif. Ces exemples montrent que si les interventions de ces quatre élèves sont significatives au niveau sémantique, les apports ne sont pas de même nature et ne témoignent pas des mêmes compétences langagières et cognitives à l'œuvre. Candice et Benoit font avancer la réflexion collective par leur apport dans la conceptualisation collective : Candice propose deux types d'injustice, Benoit une définition de la justice reposant sur le sentiment de justice (*vs*. l'égalité dans la dimension des parts de gâteau). L'apport de Manon et Dimitri est ici un effet du rôle qui leur a été donné dans la discussion : ils ne participent pas à la discussion, ce sont les secrétaires, leur mission est d'écouter, noter et restituer les propos. Leurs interventions ont plutôt une fonction d'ancrage

de la discussion dans ce qui a déjà été dit. Si les apports sont significatifs de part et d'autre, ils n'ont pas la même fonctionnalité dans la construction de la pensée, et la co-construction du discours.

Identification des principaux mots-clés de la discussion

Deuxièmement, *ReaderBench* permet d'identifier les différents mots-clés traités dans une discussion. L'analyse de ces mots-clés ne se fait pas par le simple comptage des occurrences des termes. Les différents mots composant les tours de parole sont tout d'abord analysés d'un point de vue morpho-syntaxique (de manière à déterminer notamment leur catégorie grammaticale et leur lemme). Ensuite, ce que nous nommerons leur pertinence, c'est-à-dire leur relation sémantique avec le contenu de la discussion, est calculée comme moyenne de moyennes de leur proximité avec la phrase, le tour de parole, et la discussion dans laquelle ils sont insérés. Cette pertinence peut être affichée, soit pour la conversation globale, soit par participant. Dans ce cas, la comparaison est faite avec la phrase, le tour de parole et l'ensemble des tours de parole du participant.

Nous avons choisi d'examiner les mots-clés qui apparaissent dans les propos de l'animateur et des quatre élèves repérés plus haut, notre critère de choix étant que ces élèves ont fait à un moment donné un apport significatif à la discussion et qu'il s'agit d'élèves ayant eu des rôles différents : Candice et Benoît étant « philosophes » tout au long de la DVDP, Manon et Dimitri ont été « secrétaires » pendant la première partie de la discussion pour Manon et lors de la seconde pour Dimitri. Le Tableau 1 ci-dessous reprend les mots-clés les plus représentatifs de l'ensemble de la discussion pour l'ensemble des cinq locuteurs (classés par ordre décroissant selon leur indice global), et leur indice pour les locuteurs sélectionnés. Nous avons mis en évidence les scores supérieurs à 2, valeur arbitraire qui signale une importance suffisante par rapport à la discussion, et complété le tableau pour les valeurs inférieures à 2, lorsque nécessaire.

Tableau 1 : Pertinence des termes employés

Mots-clés	**Indices par locuteur**				
	Tozzi	**Candice**	**Benoit**	**Manon**	**Dimitri**
penser (vb)	7,27	5,90	7,89	7,38	5,50
juste (adj)	6,18	5,37	4,74	6,33	1,99
part (n)	6,12	3,12	3,14	4,55	4,48
dire (vb)	7,14	4,14	N/A	4,00	2,89

Mots-clés	Indices par locuteur				
	Tozzi	**Candice**	**Benoit**	**Manon**	**Dimitri**
autre (adj)	3,22	3,30	5,99	2,80	2,50
faire (vb)	4,86	2,34	3,66	2,26	2,28
chose (n)	4,92	3,15	N/A	3,51	1,99
petit (adj)	2,98	1,58	3,53	2,03	3,01
savoir (vb)	2,06	3,54	1,35	3,20	1,61
raison (n)	6,86	N/A	1,33	1,60	1,69
prendre (vb)	2,83	2,48	1,45	2,00	2,35
bon (adj)	7,34	N/A	3,11	N/A	N/A
exemple (n)	2,21	2,80	1,88	1,32	1,66
sort (n)	4,74	1,08	1,11	1,35	1,42
pouvoir (n)	3,37	1,32	1,37	1,63	1,47
accord (n)	3,73	1,27	N/A	1,48	2,04
vrai (adj)	2,98	2,60	N/A	2,26	N/A
croire (vb)	2,42	N/A	N/A	2,43	1,97
falloir (vb)	2,50	1,03	N/A	1,72	1,26
important (adj)	3,52	N/A	N/A	1,96	N/A
essayer (vb)	2,13	1,41	N/A	1,44	N/A
travailler (vb)	2,01	0,93	N/A	N/A	1,36
justice (n)	3,07	N/A	N/A	N/A	0,92
venir (vb)	2,27	N/A	N/A	1,67	N/A
finalement (adv)	3,04	N/A	N/A	N/A	N/A
ensuite (adv)	2,47	N/A	N/A	N/A	N/A
exactement (adv)	2,46	N/A	N/A	N/A	N/A
réfléchir (vb)	2,44	N/A	N/A	N/A	N/A

Légende : seuls ont été pris en compte Michel Tozzi et quelques participants (Candice et Benoît, « philosophes » et Manon et Dimitri, « secrétaires »), les mots sont classés par somme totale décroissante. Les scores en gras sont supérieurs à 2. N/A pour valeur non disponible (quand le mot n'a pas été employé par le locuteur).

Nous pouvons classer les mots de ce tableau en plusieurs catégories. Tout d'abord, les sept mots qui ont une pertinence élevée (supérieure à 2) pour tous les participants : « penser », « juste », « part », « dire », « autre », « faire », « chose », ce qui signale d'une part l'existence d'une communauté de discours et d'autre part, on note que les trois premiers termes sont en lien direct avec la discussion : l'adjectif « juste » renvoie directement à la question traitée ; le

verbe « penser » à l'activité cognitive et langagière en jeu : réfléchir autour de la problématique, donner une opinion, la justifier ; le nom « part » renvoie au problème posé par M. Tozzi au début de la seconde partie de la discussion : *comment partager un gâteau de manière juste ?*

Ensuite, les mots qui ont une pertinence élevée seulement pour l'animateur, M. Tozzi, soit tout le bas de la liste : « falloir », « important », « essayer », « travailler », « justice », « venir », « finalement », « ensuite », « exactement », « réfléchir ». Ces mots sont tous liés, hormis « justice », à des champs de la validation (« exactement »), ou de la gestion pédagogique de la discussion (« travailler », « important », « réfléchir », « finalement »), ou de la prescription (« falloir »), ce qui est attendu compte tenu du rôle de M. Tozzi. Il est intéressant de comparer le degré de pertinence des mots « juste » (adj.) et « justice » (nom), chez les 4 élèves et chez l'animateur. En effet, on constate un écart : le terme « justice » n'est pas présent dans les propos de 3 des 4 élèves alors qu'il est repéré avec un taux de pertinence de 3,07 chez l'animateur. On peut voir là un décalage conceptuel entre les discours de ces quatre enfants et celui de l'animateur : avec un indice de pertinence fort sur l'adjectif seulement les propos des enfants sont plutôt du côté de la « qualité » des situations qu'ils proposent alors que l'animateur mobilise aussi le concept (Dimitri aussi, mais faiblement). Par ailleurs, les indices globaux de pertinence pour l'ensemble des locuteurs témoignent de cette tendance puisque le nom « justice » obtient un score de 9,059 et l'adjectif « juste » 2,915. On a là un indicateur intéressant pour mesurer le degré de conceptualisation de la discussion sur cette dimension de la réflexion.

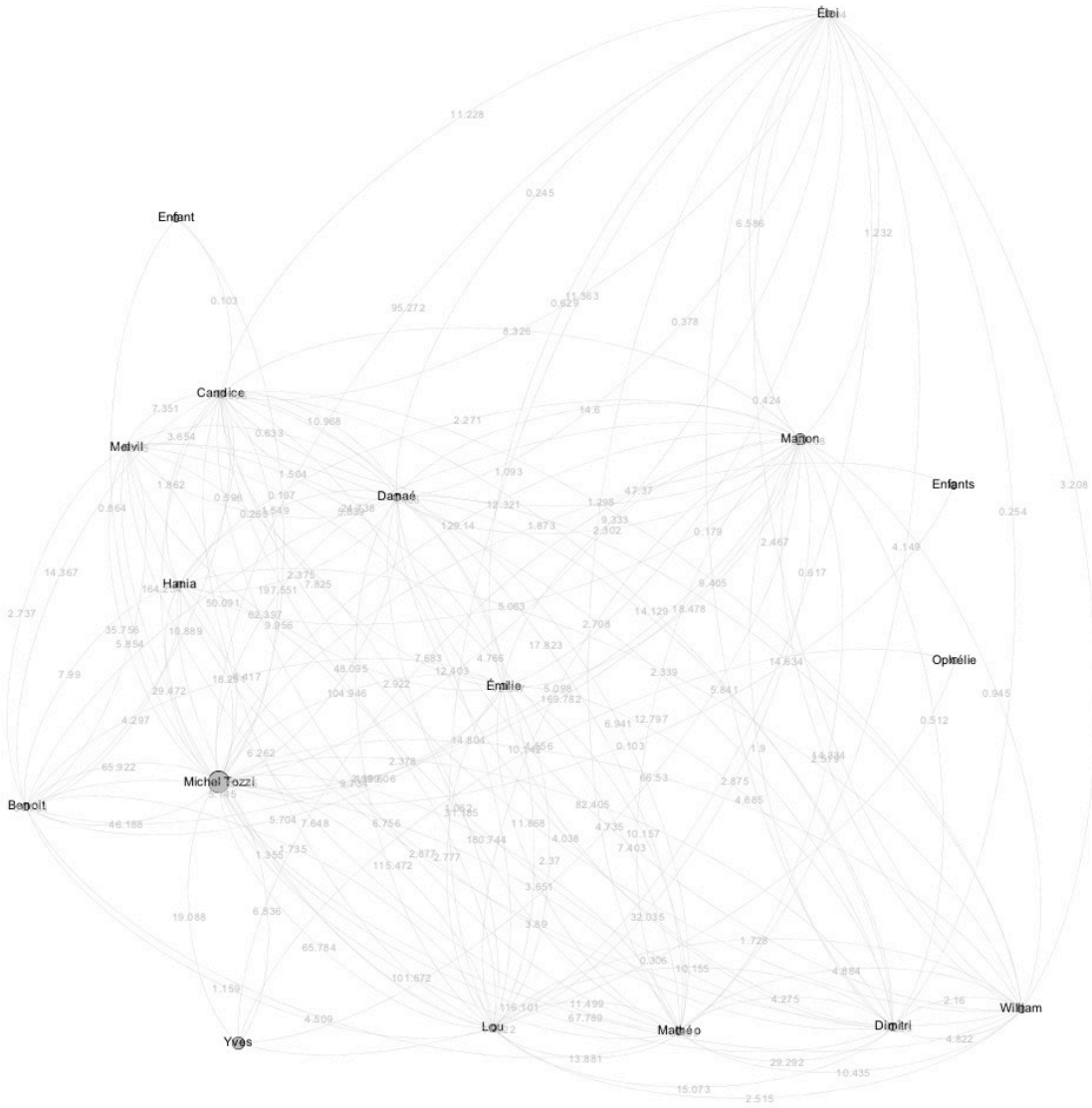
Enfin, quelques mots qui ont une pertinence différente selon le statut des élèves (secrétaire ou philosophe) : « exemple » est plus pertinent et employé chez les philosophes que chez les secrétaires, « accord », « croire » sont plus pertinents et employés chez les secrétaires que chez les philosophes. Il est à noter que les philosophes ne sont bien sûr pas tenus d'intervenir sur tous les thèmes traités, ce qui apparait par leur utilisation plus inégale que chez les secrétaires de certains mots-clés (*p. ex.*, Benoît n'utilise pas « accord » et « croire », alors que Candice n'utilise pas ou peu « raison » et « croire »).

Ainsi, *ReaderBench* pourrait être utilisé pour déterminer l'efficacité d'un-e secrétaire à sa manière d'utiliser, dans ses interventions, les principaux mots-clés d'une discussion entière ainsi que le niveau de conceptualisation en comparant les degrés de pertinence des adjectifs et des noms ayant la même racine comme nous l'avons fait pour « juste » et « justice »

Analyse des proximités thématiques des tours de paroles des participants

Une troisième fonctionnalité de *ReaderBench* est de calculer, avec des techniques d'analyse de réseaux sociaux (Reffay & Martinez-Mones, 2013), la proximité thématique des propos tenus par chacun des participants. L'ensemble des tours de parole de chaque participant est comparé sémantiquement à ceux des autres.

Figure 2 : Analyse en réseau social des propos de chaque participant de la discussion philosophique



Sur ce réseau (Figure 2), qui prend en compte l'ensemble de la discussion, on visualise bien un premier cercle de participants autour de l'animateur, qui contribuent donc de manière similaire à la discussion : Mathéo, Dimitri, Benoit et Danaé. Toutefois, s'en tenir à cette analyse n'est pas satisfaisant, car les valeurs calculées peuvent être influencées par le rôle assigné aux élèves, rôle qui change selon les moments de la DVDP.

Perspective dialogique : la mise au jour de thèmes

La quatrième fonctionnalité de *ReaderBench* concerne la sélection et visualisation de « voix », ou thèmes, tout au long de la discussion. Il revient alors au chercheur de sélectionner les thèmes en lien avec la problématique de la discussion (Figure 3 : Copie d'écran de la procédure de sélection des voix.) qu'il pourra alors afficher longitudinalement afin de visualiser la dynamique de la discussion (Figure 4). Chaque barre de cette dernière figure représente une occurrence de thème, localisée par tour de parole (en abscisse). Cliquer sur une barre permet d'afficher son contenu. Les couleurs ne sont pas signifiantes.

Figure 3 : Copie d'écran de la procédure de sélection des voix.

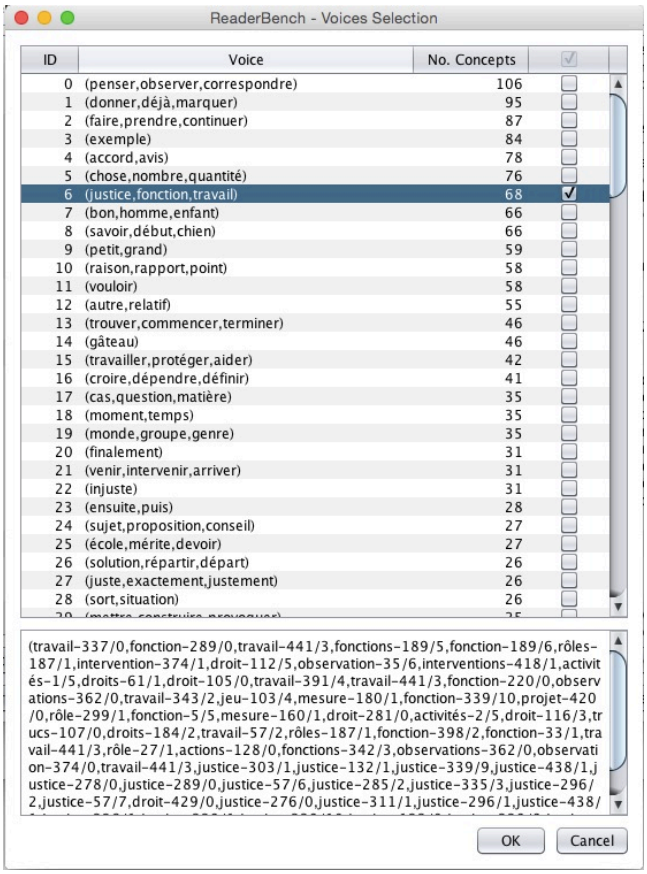
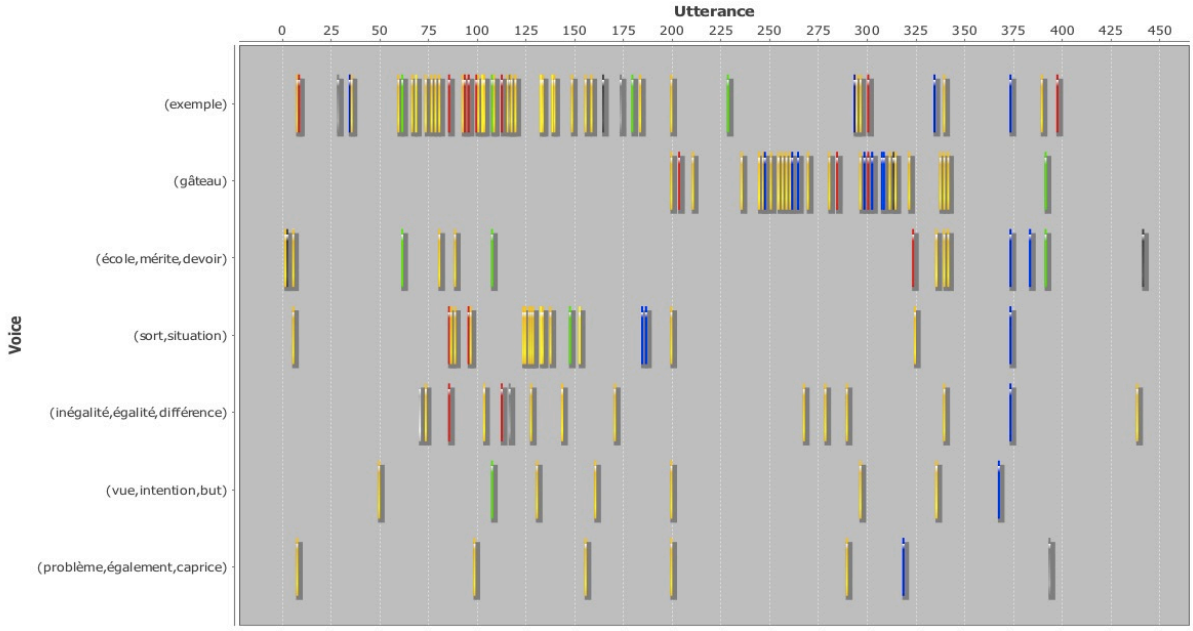


Figure 4 : Vue longitudinale de quelques thèmes sélectionnés.

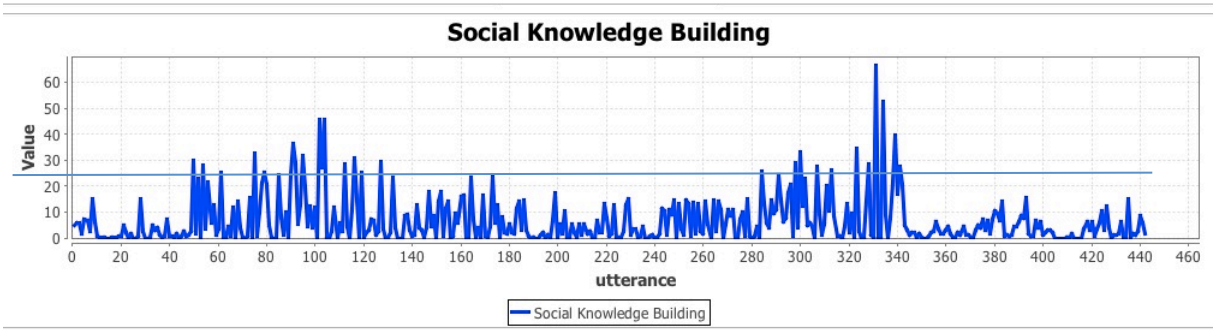


L'exemple affiché ci-dessus, permet de voir que la discussion s'organise autour de deux temps : une phase de *brainstorming* à partir d'exemples proposés par les enfants (du tour de parole 60 à 180) puis autour d'un exemple particulier : le partage du gâteau (de 200 à 330 environ). Cette visualisation permet d'affiner l'analyse sémantique : on peut ainsi explorer les différentes dimensions sémantiques de la discussion, et donc ses phases. Le logiciel permettra dans certains cas de mettre au jour des univers sémantiques qu'une lecture humaine n'aurait pas perçus. Cette fonctionnalité permet aussi de voir qui apporte quelle thématique : l'animateur (ici la question du partage du gâteau), les élèves.

Une fois que ce positionnement de thèmes tout au long de la discussion a été réalisé, *ReaderBench* affiche également (voir

Figure 5) la manière dont les tours de parole des différents participants sont sémantiquement reliés à la discussion complète. Les pics du graphique de la figure indiquent qu'aux tours de parole concernés différents participants ont des tours de parole à la fois sémantiquement proches les uns des autres et proches de la discussion complète, ce qui peut rendre compte d'une construction de connaissance collaborative (Scardamalia & Bereiter, 1994).

Figure 5 : Représentation de la construction collaborative sur l'ensemble de la DVDP



Nous allons maintenant explorer quelques pics de la Figure 5. Il faut préciser que logiciel permet de zoomer sur un ensemble de tours de parole afin d'avoir une vision plus précise des échanges. La vision globale de la Figure 5 met en évidence quelques moments où la construction collaborative est plus forte : on observe des pics entre 49 et 134 entre 280-320 et 320-345. Nous allons maintenant explorer quelques-uns de ces pics.

À la question de l'animateur portant sur le rôle de philosophe : « *et alors les: philosophes vous qu'est-ce que qu'est-ce que vous allez faire* » (47) ce sont Melvil, Lou et Éloi qui interviennent avec les apports suivants :

- Melvil (48) : *(…) on va débattre sur […] le sujet (…)*
- Lou (50) : *(…) on va s'exprimer on va on va dire tout c'qu'on pense (…)*
- Éloi : (52) *ben il faut dire si on est d'accord ou pas d'accord avec euh c'que une personne a dit* et (54) *et si on n'est pas d'accord bah on essaie de dire pourquoi (…)*

Ces interventions se caractérisent par leur complémentarité, ce qui peut se résumer ainsi : « *être philosophe c'est débattre, s'exprimer, dire si on est d'accord ou pas, justifier son point de vue* ». Il y a co-construction de la réponse, l'animateur intervient essentiellement pour valider les propositions, les reformuler (ce qui est une forme de validation), et questionner. Cela est cohérent avec le score attribué à ces tours de parole par *ReaderBench*[5].

Nous avons présenté plus haut plusieurs interventions identifiées comme des apports significatifs, l'analyse des liens sémantiques entre une intervention et la discussion complète (sous la forme d'un score de *construction sociale de connaissances* attribué à chaque tour de parole) apporte un éclairage complémentaire à l'analyste.

En 85, un apport significatif de Candice qui propose « *plusieurs sens* » de *c'est pas juste*, même si les interventions qui suivent (91 et 85) ne constituent pas un apport significatif, elles illustrent cette pluralité de sens à partir du vécu (95) « *quand j'étais petite* » puis avec un exemple fictif « *(…) il y a un exemple, mais je l'ai pas je l'ai pas passé quoi c'est pas*

[5] TDP n°50 : 29,91 ; TDP n°52 : 23,20 et TDP n°54 : 28,26.

vrai (...) » ces interventions permettent une conceptualisation collective dont témoigne la reprise explicite d'Éloi (en 102) « *ben il y a comme elle a comme elle dit Candice il y a des exemples où* (…) », on peut aussi considérer qu'en 107 Mathéo va dans le même sens.

Idem pour l'intervention de Benoit, en 298, identifiée par *ReaderBench* comme apport significatif à la discussion, avec le score de 29,22 et aux tours de parole adjacents (n° 300 : Candice, 33,10), qui s'inscrit dans le cadre d'une construction collective. Si en 292, Benoit essaye d'introduire l'idée de justice apparente (*vs*. réelle), il formule un exemple qui est censé illustrer cela, mais qui n'est pas très clair, les hésitations témoignent ici d'une pensée en construction. La reformulation, en 300, de Candice par une formulation plus fluide qui explicite avec l'idée de sacrifice (réel et/ou virtuel) permet à Benoit en 302 de formuler plus clairement cette idée de justice apparente.

En revanche, si l'intervention de Manon en 184 est significative sur le plan sémantique, elle ne l'est pas sur le plan de la construction sociale de la connaissance, cela est cohérent avec ce que nous avions noté plus haut, ici Manon est dans son rôle de secrétaire, elle résume les différents apports faits jusqu'alors. Il en va de même de l'intervention de Dimitri en 341.

Analyser les discussions philosophiques

Il s'agit là d'une première analyse automatique d'une discussion philosophique, à partir d'une transcription orale. Tout d'abord, nous pouvons noter que les différentes étapes de transcription puis de formatage de la discussion pour qu'elle soit analysable par *ReaderBench* prennent un temps raisonnable. Ensuite, nous avons mis en évidence quelques fonctionnalités du logiciel qui peuvent signaler à l'enseignant ou aux participants des éléments à analyser de plus près. *ReaderBench* produit deux types de visualisations complémentaires. Des visualisations *centrées sur les tours de paroles* (qui figurent souvent en abscisses) permettent d'une part de visualiser les apports des participants à la discussion complète (Figure 1), d'autre part l'interanimation des thèmes tout au long de la discussion (Figure 4). Des visualisations *centrées sur les participants* (Figure 5), permettant mettre en évidence des rapprochements sémantiques entre les contributions des intervenants.

Nous tenons à rappeler que ce système doit être vu comme une aide à l'analyse humaine des discussions, pouvant amener l'enseignant ou les participants à analyser plus rapidement les contributions des participants et leur contenu thématique. Il ne peut être utilisé comme un outil d'évaluation sommative automatique des contributions, pouvant mener à des notes ou des certifications. Il n'est pas exempt de limites. Ainsi, nous avons parfois noté des analyses qui nous interrogent comme au tour de parole 104, où Mathéo demande : « *est-ce que Hania*

tu as quelque chose à dire ? » est affecté d'un score élevé de *construction sociale de connaissances* (45,8, sachant que le score maximal est de 66,59), alors que le même énoncé une quarantaine de tours de parole plus loin (TDP n° 149, toujours Mathéo), obtient un score de 8,63. Cela juste pour insister sur le nécessaire dialogue entre l'analyste (s'il n'est pas l'auteur du logiciel) et les informaticiens qui conçoivent les programmes d'analyse.

Pour ce qui est des corpus à analyser, il faut prendre en compte quelques contraintes : non seulement les discussions doivent être d'un volume (en nombre de mots) suffisant, mais cela doit être également le cas des tours de parole d'un participant donné. Les tours de parole très courts comme « *voilà* », « *c'est ça* », « *bah non, mais* » … ne sont pas analysés alors que sur le plan pragmatique ils ont une valeur certaine. En outre, un lexique de spécialité ne pourra être pris en compte si on ne fournit pas en amont de l'analyse un corpus de référence appartenant au domaine de spécialité.

Nous projetons de comparer ces différentes analyses automatiques avec des analyses réalisées par des chercheurs ou enseignants. Une autre possibilité sera d'utiliser d'autres fonctionnalités existantes, comme celle d'analyse des stratégies de compréhension de textes pour s'intéresser à la manière dont des élèves comprennent un texte philosophique donné, qui sera lui-même analysé (M. Dascalu, Dessus, Bianco, & Trausan-Matu, 2014). La fonctionnalité d'analyse de la complexité textuelle peut également être utile pour calibrer la difficulté d'étude de textes philosophiques (M. Dascalu, Stavarache, Trausan-Matu, Dessus, & Bianco, 2014). De plus, l'ajout d'un module d'analyse de sentiments à *ReaderBench* va pouvoir permettre l'analyse de la valence émotionnelle des tours de parole.

Remerciements

Nous remercions tous les chercheurs qui ont pris part dans la conception et le développement de *ReaderBench*. Son développement a été partiellement réalisé par les projets FP7 2008-212578 LTfLL et par le Projet européen H2020 RAGE No 644187 *(Realising and Applied Gaming Eco-System*) http://www.rageproject.eu/, ainsi que par le projet DEVCOMP (ANR-10-BLAN-1907-01) financé par l'Agence nationale de la recherche (ANR).

Références

Bakhtin, M. M. (1981). *The dialogic imagination: Four essays*. Austin: The University of Texas Press.

Blei, D., Ng, A., & Jordan, M. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research, 3*(4-5), 993–1022.

Chaturvedi, S., Goldwasser, D., & Daumé, H. (2014). *Predicting instructor's intervention in MOOC forums.* Paper presented at the Proc. 52nd Ann. Meeting of the Association for Computational Linguistics, Baltimore.

Dascalu, M., Dessus, P., Bianco, M., & Trausan-Matu, S. (2014). Are automatically identified reading strategies reliable predictors of comprehension? In S. Trausan-Matu, E. Boyer, M. Crosby & K. Panourgia (Eds.), *12th Int. Conf. on Intelligent Tutoring Systems (ITS 2014)* (Vol. LNCS 8474, pp. 456–465). New York: Springer.

Dascalu, M., Dessus, P., Bianco, M., Trausan-Matu, S., & Nardy, A. (2014). Mining texts, learners productions and strategies with ReaderBench. In A. Pena-Ayala (Ed.), *Educational Data Mining: Applications and Trends* (pp. 345–377). New York: Springer.

Dascalu, M., Stavarache, L. L., Trausan-Matu, S., Dessus, P., & Bianco, M. (2014). Reflecting comprehension through French textual complexity factors *26th IEEE Int. Conf. on Tools with Artificial Intelligence (ICTAI 2014)* (pp. 615–619). Limassol, Cyprus: IEEE.

Dascalu, M., Trausan-Matu, S., Dessus, P., & McNamara, D. S. (2015). Dialogism: A framework for CSCL and a signature of collaboration. In P. H. k. O. Lindwall, T. Koschman, P. Tchounikine, & S. Ludvigsen (Ed.), *Exploring the material conditions of learning: The Computer Supported Collaborative Learning (CSCL) Conference* (Vol. 1, pp. 86–93). Gothenburg, Sweden: The International Society of the Learning Sciences.

Dascalu, M., Trausan-Matu, S., McNamara, D. S., & Dessus, P. (2015). ReaderBench - Automated evaluation of collaboration based on cohesion and dialogism. *International Journal of Computer-Supported Collaborative Learning, 10*(4), 395–423. doi: 10.1007/s11412-015-9226-y

Dessus, P., Lemaire, B., Loiseau, M., Mandin, S., Villiot-Leclercq, E., & Zampa, V. (2011). Automated free-text assessment: Some lessons learned. *International Journal of Continuing Engineering Education and Life-Long Learning, 21*(2/3), 140–154. doi: 10.1504/IJCEELL.2011.040195

Dessus, P., Trausan-Matu, S., Wild, F., Dupré, D., Loiseau, M., Rebedea, T., & Zampa, V. (2011). Un environnement personnel d'apprentissage évaluant des distances épistémiques et dialogiques. *Distances & Savoirs, 9*(4), 473–492.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review, 104*(2), 211–240.

Linell, P. (2009). *Rethinking language, mind, and world dialogically*. Greenwich: Information Age Publications.

Manning, C., & Schütze, H. (1999). *Foundations of statistical Natural Language Processing*. Cambridge (Mass.): MIT Press.

Morel, M.-A., & Danon-Boileau, L. (2001). *Grammaire de l'intonation, l'exemple du français*. Paris: Ophrys.

Mu, J., Stegmann, K., Mayfield, E., Rosé, C., & Fischer, F. (2012). The ACDOEA framework: Developing segmentation and classification schemes for fully automatic analysis of online discussions. *Computer-Supported Collaborative Learning, 7*, 285–305. doi: 10.1007/s11412-012-9147-y

Nistor, N., Dascalu, M., Stavarache, L. L., Tarnai, C., & Trausan-Matu, S. (2016). Predicting newcomer integration in online knowledge communities by automated dialog analysis. In Y. Li, M. Chang, M. Kravcik, E. Popescu, R. Huang, Kinshuk & N.-S. Chen (Eds.), *State-of-the-art and future directions of smart learning* (pp. 13–17). Berlin: Springer.

Reffay, C., & Martinez-Mones, A. (2013). Basic concepts and techniques in social network analysis. In R. Luckin, S. Puntambekar, P. Goodyear, B. Grabowski, J. Underwood & N. Winters (Eds.), *Handbook of design in educational technology* (pp. 448–456). New York: Routledge.

Scardamalia, M., & Bereiter, C. (1994). Computer support for knowledge-building communities. *The Journal of the Learning Sciences, 3*(3), 265–283.

Sloetjes, H., & Wittenburg, P. (2008). *Annotation by category – ELAN and ISO DCR.* Paper presented at the Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008), Marrakech.

Stahl, G. (2006). *Group cognition. Computer support for building collaborative knowledge*. Cambridge: MIT Press.

Tousseul, S. (2005). L'origine du sens : l'Affect Inconditionnel. *Revue d'Intelligence Artificielle, 19*(1), 265–280.

Trausan-Matu, S. (2010). Automatic support for the analysis of online collaborative learning chat conversations. In P. M. Tsang, S. K. S. Cheung, V. S. K. Lee & R. Huang (Eds.), *Hybrid Learning, Third International Conference, ICHL 2010* (pp. 383–394). New York: Springer.