



HAL
open science

Emplois sous-spécifiés des marqueurs discursifs et / and à l'oral : stratégie (inter)subjective et variation en genre

Ludivine Crible

► To cite this version:

Ludivine Crible. Emplois sous-spécifiés des marqueurs discursifs et / and à l'oral : stratégie (inter)subjective et variation en genre. Les Cahiers FoReLLIS – Formes et Représentations en Linguistique, Littérature et dans les arts de l'Image et de la Scène, 2019, Traces de subjectivité et corpus multilingues, III. Les données des corpus comparables : comparaison des langues dans des genres ciblés. hal-02987097

HAL Id: hal-02987097

<https://hal.science/hal-02987097v1>

Submitted on 14 Apr 2021

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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License



Emplois sous-spécifiés des marqueurs discursifs et / and à l'oral : stratégie (inter)subjective et variation en genre

Par Ludivine Crible

Publication en ligne le 06 septembre 2019

Table des matières

1. Marqueurs discursifs et (inter)subjectivité
2. L'égo-centrisme des marqueurs discursifs sous-spécifiés
3. *DisFrEn* : un corpus oral comparable anglais-français
4. Méthode d'annotation
 - 4.1. Identification des marqueurs discursifs
 - 4.2. Annotation fonctionnelle des marqueurs discursifs
 - 4.3 Annotation des disfluences
5. Résultats quantitatifs sur corpus
6. Analyse qualitative des stratégies compensatoires
- Conclusion

Texte intégral

La communication orale dispose d'un ensemble riche et varié de formes linguistiques et paralinguistiques multimodales qui permettent aux locuteurs d'exprimer leur attitude vis-

à-vis de l'énoncé en cours de production et de se positionner par rapport à l'interlocuteur (Benveniste, 1966). Ces marques de subjectivité et d'intersubjectivité contribuent au bon déroulement d'une interaction et peuvent agir aux niveaux syntaxique, prosodique ou pragmatique séparément ou simultanément. En fonction du système linguistique et, à l'intérieur de celui-ci, du genre discursif activé, l'(inter)subjectivité peut endosser des formes et fonctions diverses qui révèlent différents degrés d'implication du locuteur dans son énoncé ainsi que des effets contextuels liés à la spontanéité et l'interactivité d'un dialogue. Dans cette perspective, les approches sur corpus multilingues permettent de distinguer les « universaux » des particularismes dans l'expression de la relation locuteur-énoncé et locuteur-interlocuteur.

Dans cet article, nous proposons une étude contrastive et variationniste d'un cas particulier de ces marques (inter)subjectives, à savoir la conjonction de coordination *et* et son équivalent anglais *and* dans leurs emplois de « marqueurs discursifs » (*discourse markers*, Schiffrin, 1987), c'est-à-dire, fonctionnant comme opérateurs pragmatiques portant sur deux ou plusieurs énoncés dont ils facilitent l'interprétation en explicitant leur relation discursive comme une cause ou un contraste, par exemple. Les marqueurs discursifs tels que *et*, *mais*, *donc*, *enfin*, *bon*, *alors*, *quoi*, etc., peuvent exprimer des fonctions très variées concernant à la fois le contenu et la structure des unités du discours. Leur multifonctionnalité est particulièrement grande à l'oral puisque les locuteurs sont engagés dans plusieurs activités à la fois : conceptualisation et articulation plus ou moins spontanée des énoncés, interprétation en direct, structuration des actes de parole et autres unités hiérarchiques, organisation des tours de parole, gestion de la relation interpersonnelle, etc. (Bunt, 2011). La plupart de ces fonctions ont un rapport direct avec les attitudes du locuteur à l'égard de son énoncé et/ou de son interlocuteur, ce qui confère aux marqueurs discursifs un rôle majeur dans l'expression de l'(inter)subjectivité, comme l'ont déjà étudié de nombreux auteurs (Haselow, 2012 ; House, 2013 ; Paillard, 2011 ; Traugott, 2007).

Parmi toutes les caractéristiques des marqueurs discursifs qui sont pertinentes à l'analyse de la subjectivité, il en est une qui reste encore largement à explorer : il s'agit de l'imprécision ou sous-spécification (*underspecification*) des marqueurs, c'est-à-dire le fait qu'un marqueur exprime une fonction en contexte qui n'est pas entièrement encodée dans son sémantisme mais qui s'appuie plutôt sur les savoirs partagés et d'autres indices linguistiques pour être interprétée (Frisson, 2009 ; Spooren, 1997). La sous-spécification concerne, par exemple, l'emploi du marqueur *et* pour signaler non pas son sens additif de base mais plutôt une relation plus spécifique comme un contraste ou une conséquence. Notre étude se concentre sur ces emplois sous-spécifiés de *et* et *and* afin d'en distinguer des usages subjectifs et d'autres plutôt intersubjectifs, sur base de leur distribution dans

différents genres de parole et de leur co-occurrence avec d'autres éléments du discours. L'hypothèse centrale de cette analyse suggère que la sous-spécification est fondamentalement égocentrique (*speaker-oriented*) ou subjective en ce qu'elle constitue une stratégie de production au bénéfice du locuteur et non de l'interlocuteur. Cet argument sera développé par une réflexion théorique ainsi que sur la base d'analyses quantitatives et qualitatives sur un corpus oral comparable.

1. Marqueurs discursifs et (inter)subjectivité

Dans l'introduction du volume qu'elle a dirigé, Hancil (2011 : 7) définit la subjectivité comme l'« expression, explicite ou implicite, des émotions et attitudes de l'énonciateur dans la construction et la cohésion discursives ». En ce sens, les marqueurs discursifs occupent une place privilégiée dans l'expression de la subjectivité. En effet, ils constituent des indices de contextualisation et rendent manifestes les intentions de sens du locuteur, que ce soit par le signalement de leur perspective sur la relation entre des énoncés ou par le marquage d'une attitude ou d'un commentaire sur une situation ou un énoncé donné (Traugott, 2007). Cette fonction subjective des marqueurs discursifs explique leur rôle dans l'argumentation (Anscombe & Ducrot, 1983), l'émotion (Romano & Cuenca, 2013) ou encore l'identité sociolinguistique (Andersen, 1997 ; Beeching, 2007).

La plupart des travaux portant sur la fonction subjective des marqueurs discursifs ne manquent pas de mentionner la notion voisine d'intersubjectivité, comme House (2013 : 58) qui affirme l'existence d'une « *close relationship between subjectivity and addressee orientation, and thus to what one may call 'intersubjectivity'* ». Au-delà des adresses directes à l'interlocuteur, typiquement exprimées par des marqueurs discursifs tels que *tu vois* ou *écoute*, la fonction intersubjective centrale des marqueurs consiste en l'inscription de leur énoncé-hôte dans l'ensemble des savoirs conversationnels partagés. Cette fonction procédurale est présente dans la définition même de la catégorie des marqueurs discursifs proposée par Hansen (2006 : 25) : les marqueurs discursifs « *provide instructions to the hearer on how to integrate their host utterance into a developing mental model of the discourse in such a way as to make that utterance appear optimally coherent* ». En d'autres termes, ils contribuent à enrichir les savoirs partagés ou *common ground* entre les participants. D'après Fetzer & Fischer (2007 : 8), « *a lexical marker contributes to organizing common ground by indicating the nature of the connectedness between particular information communicated by a particular piece of discourse with the discourse common ground and with other types of common ground* ». L'utilisation d'un marqueur discursif pour expliciter la relation entre deux énoncés relève donc d'une marque d'implication

subjective et intersubjective du locuteur vis-à-vis de son discours, qu'il tient à rendre cohérent et interprétable pour son interlocuteur, et ce d'autant plus que les marqueurs sont grammaticalement optionnels par définition, ce qui rend leur présence encore plus signifiante (bien que souvent non-intentionnelle).

Il apparaît que la frontière entre subjectivité et intersubjectivité est perméable et relative, surtout à l'oral où un interlocuteur est toujours présent (même s'il est virtuel). C'est pour cette raison que Degand & Fagard (2012) suggèrent de réserver la subjectivité au domaine de l'expression du locuteur et d'étendre l'intersubjectivité à tout ce qui concerne l'interaction. En ce sens, les fonctions structurantes et métadiscursives telles que la digression ou la reformulation seraient intersubjectives puisqu'elles « *materialize the strategic interaction between speaker and hearer and reflect the active role of the speaker to orient and to guide the hearer in his interpretational tasks* » (Carlier & De Mulder, 2010 : 269).

Dans la même veine, on peut considérer la subjectivité comme un égo-centrisme langagier (*speaker-orientation*), par opposition à l'intersubjectivité qui serait alors envisagée comme allo-centrisme (*addressee-orientation*). Dans cette perspective, qui s'apparente à la Théorie de la Pertinence (Sperber & Wilson, 1989), la subjectivité permet au locuteur de se concentrer sur ses propres besoins de production (avec un moindre effort cognitif), tandis que l'intersubjectivité consiste en une prise en charge des besoins d'interprétation de l'interlocuteur, au risque d'un surcoût cognitif. Lenk (1998) illustre cette approche dans son analyse de l'expression anglaise *what else* 'quoi d'autre', qu'elle considère comme un marqueur d'auto-élicitation utilisé par les locuteurs « *to aid their own orientation within the stretch of discourse being produced, or as an expression of their mental processes* », en particulier lorsque le locuteur éprouve des difficultés (1998 : 192). Comme le résume Haselow (2012 : 190), « *intersubjectivity is thus the coded expression of the speaker's attention to the self of the addressee and his/her needs concerning the processing of a verbal message* », par opposition à l'égo-centrisme des marqueurs subjectifs. Dans la suite de cette étude, c'est cette dernière définition de la notion d'(inter)subjectivité qui sera utilisée et éclairée par l'analyse des emplois sous-spécifiés de *et / and*, afin d'en montrer les potentiels corrélats cognitifs.

2. L'égo-centrisme des marqueurs discursifs sous-spécifiés

La sous-spécification (ou usage imprécis) des marqueurs discursifs est un phénomène qui a été introduit par Spooren (1997) dans son analyse pragmatique d'énoncés oraux élicités chez des enfants néerlandophones. Il propose d'interpréter la sous-spécification dans les termes de Horn (1984), en particulier les principes R et Q : le premier stipule qu'un locuteur doit fournir les informations nécessaires à la compréhension, ni plus ni moins (cf. les maximes de pertinence, manière et quantité chez Grice, 1981) ; le deuxième établit qu'un locuteur dit tout ce qu'il peut. Le principe R est un principe d'économie pour le locuteur, qui peut exploiter un nombre minimal de formes pour exprimer plusieurs sens. À l'inverse, le principe Q est un principe d'économie pour l'interlocuteur, qui a toutes les informations à sa disposition. L'analyse de Spooren (1997) a montré que la tendance à la sous-spécification diminue à mesure que la maîtrise de la langue augmente chez les enfants, ce qui suggère qu'à travers l'acquisition d'expérience, le locuteur se préoccupe de plus en plus des besoins de l'interlocuteur, allant ainsi à l'encontre du principe d'économie pour sa propre production (principe R).

Un parallèle apparaît clairement entre, d'une part, cette interprétation de la sous-spécification comme principe d'économie et, d'autre part, la définition de la subjectivité comme pratique égocentrique reflétant les processus de production du locuteur. De même, l'acquisition du principe Q de Horn, qui va à l'encontre de ce principe d'économie, fait écho à l'intersubjectivité définie comme prise en charge des besoins interprétatifs de l'interlocuteur. Concrètement, l'utilisation sous-spécifiée d'un marqueur discursif comme *et / and* pour exprimer une relation de contraste ou de conséquence permet au locuteur de « recycler » un marqueur hautement multifonctionnel et très fréquent, ce qui allège sa tâche de production en limitant la recherche lexicale tout en laissant plusieurs possibilités ouvertes quant à la suite de l'énoncé : une proposition introduite par *et* n'appelle pas nécessairement un contraste ou une conséquence, contrairement à *en revanche* ou *par conséquent*, respectivement. Les marqueurs discursifs sous-spécifiés n'en restent pas moins subjectifs dans le sens de la première définition (i.e. attitude et positionnement du locuteur à l'égard de l'énoncé) puisqu'ils reflètent une relation – certes imprécise – que le locuteur envisage entre deux énoncés et rendent donc compte de la cohérence que le locuteur entend conférer à son discours.

Cette association entre principe d'économie et subjectivité suggère que les marqueurs discursifs sous-spécifiés devraient être privilégiés dans des types de contextes hautement interactifs et spontanés, où le locuteur ne dispose pas du temps ni des ressources cognitives nécessaires pour sélectionner avec précision le marqueur discursif le plus spécifique possible pour exprimer la relation envisagée entre une paire d'énoncés en cours de production. À l'inverse, dans les situations de parole plus formelles où le locuteur bénéficie d'un temps de préparation préalable (voire d'un script complet), on peut

s'attendre à ce que les besoins interprétatifs de l'interlocuteur soient mieux pris en charge par le locuteur, afin de favoriser une compréhension optimale, par exemple dans des contextes publics et professionnels comme un discours politique. Par conséquent, la proportion de marqueurs discursifs sous-spécifiés devrait diminuer à mesure que la préparation du discours augmente, conformément au principe Q. Cette hypothèse peut aisément se vérifier sur des échantillons de corpus représentatifs d'une diversité de genres discursifs.

Toutefois, des travaux antérieurs suggèrent que la situation n'est peut-être pas aussi simple que cette opposition binaire entre subjectif/informel et intersubjectif/formel. En effet, Asr & Demberg (2012) et Webber (2013) ont montré que la production et l'interprétation de relations de cohérence telles que cause ou contraste ne reposent pas uniquement sur la présence d'un marqueur discursif plus ou moins spécifique, mais qu'un ensemble d'autres éléments linguistiques peuvent nourrir les inférences pragmatiques, tels que les marques de négation, les modalisateurs ou les temps verbaux. D'après Asr & Demberg (2012 : 7), les relations causales marquées par un *et* (par exemple *Jean est tombé et s'est cassé la jambe*) « *are not really explicit causal relations, rather the causality is left implicit in the content of the arguments* ». De même, dans son étude sur les relations d'alternative, Webber (2013) conclut que la présence ou l'absence d'un marqueur va de pair avec la présence d'indices plus ou moins subtils dans le contexte. En somme, l'usage sous-spécifié d'un marqueur discursif n'est peut-être pas tant l'effet d'une stratégie subjective d'économie de production (principe R) que le résultat d'autres choix linguistiques visant à limiter la redondance d'information, dans un souci d'économie de la tâche interprétative (principe Q). Dans une perspective multilingue, il serait alors intéressant d'observer comment différentes langues « compensent » la sous-spécification des marqueurs discursifs.

Parmi ces stratégies de compensation, il en est une qui est propre à l'oral, et en particulier aux genres informels : la présence de « disfluences » dans le contexte immédiat de marqueurs discursifs. Par « disfluences », nous entendons des éléments qui témoignent du travail de production en cours comme des pauses silencieuses ou remplies (*euh*) ou encore des interruptions de mots ou de structures syntaxiques (Candea, 2000 ; Pallaud *et al.*, 2013 ; Shriberg, 1994). Ces éléments (voir la liste complète à la Section 4.3) sont souvent envisagés comme des symptômes d'une hésitation passagère, bien que de nombreuses études expérimentales aient également démontré leurs effets positifs, notamment au niveau de l'accès référentiel ou de la rétention d'information (Arnold *et al.*, 2003 ; Bosker *et al.*, 2014). Dans la plupart des travaux, comme ici, les marqueurs discursifs eux-mêmes sont considérés comme des disfluences, bien que leur inclusion dans cette catégorie reste encore largement partielle et soumise à condition dans beaucoup d'études (Götz, 2013 ; Meteer *et al.*, 1995). Une approche exhaustive et inclusive des marqueurs discursifs et des

disfluences nous a permis de révéler que ces éléments sont très souvent produits ensemble, dans des séquences de co-occurrence pouvant contenir de deux à plus d'une trentaine de disfluences adjacentes (Crible, 2017a).

Si les disfluences n'ont *a priori* pas de rapport direct avec la subjectivité, dans son sens premier, il n'est pas exclu qu'elles participent aux stratégies de compensation par co-occurrence définies plus haut. En effet, on peut aisément concevoir que l'adjonction d'une pause ou d'un *eu*h à un marqueur discursif peut générer des inférences pragmatiques différentes que dans un cas où la co-occurrence n'est pas réalisée : une séquence de type [pause + *et*] pourrait renforcer la segmentation entre les deux énoncés ainsi reliés et privilégier une interprétation contrastive plutôt qu'additive. Ces spéculations sur le rôle compensatoire des disfluences dans le contexte des marqueurs discursifs sous-spécifiés seront testées par plusieurs analyses d'exemples en Section 6. Les résultats de l'étude mentionnée ci-dessus (Crible, 2017a) semblent déjà confirmer cette hypothèse : il a été montré que certains types de fonctions des marqueurs discursifs sont en effet statistiquement associés à certains types de disfluences. Par exemple, les fonctions structurantes (ouverture d'un tour de parole, changement de sujet) entrent souvent en co-occurrence avec des pauses silencieuses ou pleines. Reste à savoir si ce type d'associations est également valide pour le cas particulier de *et* / *and* à travers différents genres discursifs en anglais et en français.

3. *DisFrEn* : un corpus oral comparable anglais-français

La présente étude a été réalisée à l'aide du corpus comparable *DisFrEn* qui rassemble des transcriptions (alignées au son) en anglais et en français, échantillonnées à partir de plusieurs corpus d'origine : pour l'anglais, la composante britannique du *International Corpus of English* (Nelson *et al.*, 2002) et les données du projet *BACKBONE* (Kohn, 2012) ; pour le français, les corpus VALIBEL (Dister *et al.*, 2009), CLAPI (Palisse, 1997), LOCAS-F (Degand *et al.*, 2014), C-Phonogenre (Goldman *et al.*, 2014), Rhapsodie (Lacheret *et al.*, 2014) et C-Humour (Grosman, 2016). À partir de ces différentes sources, le corpus *DisFrEn* a été construit dans une perspective variationniste afin d'étudier la distribution et diverses caractéristiques des marqueurs discursifs à travers plusieurs genres de parole. Les textes ont donc été sélectionnés sur un critère de comparabilité inter-linguistique des genres, avec la condition supplémentaire que le fichier audio devait être fourni avec la transcription. La taille du corpus par langue et par genre est rapportée dans le Tableau 1.

Comme on peut le constater, *DisFrEn* comprend aussi bien des dialogues que des monologues, couvrant différents degrés de préparation (conversation spontanée, cours en classe semi-préparé, journal parlé scripté) et d'interactivité (échange téléphonique à bâtons rompus, interview semi-interactive, discours politique monologal), ainsi qu'une variable de médiatisation qui concerne la moitié des sous-corpus. Cet échantillon se veut représentatif de différents types de production linguistique, finement déclinés en huit sous-corpus afin de permettre des comparaisons et recoupements multiples (par ex. dialogue vs monologue ; préparé vs spontané), complétant ainsi les approches sur seulement deux ou trois registres différents.

Sous-corpus	Anglais	Français	Total
conversation	17479	17432	34911
interview face-à-face	17055	18043	35098
interview radio	8773	8416	17189
appel téléphonique	9747	6783	16530
cours en classe	9425	3723	13148
discours politique	8650	7824	16474
journal parlé	7046	6788	13834
commentaires sportifs	8237	6279	14516
Total	86412	75288	161700

Tableau 1 – Nombre de mots par langue et par genre dans *DisFrEn*

En dépit de la variété de ses sources, le format final de *DisFrEn* a été homogénéisé à la suite d'un traitement technique permettant de consulter la bande sonore alignée à la transcription et d'analyser différents phénomènes linguistiques sur plusieurs couches d'annotation à l'aide du logiciel EXMARaLDA (Schmidt & Wörner, 2012).

4. Méthode d'annotation

Notre analyse repose sur les annotations manuelles réalisées sur le corpus *DisFrEn* en suivant une procédure en trois étapes principales : 1) identification en contexte des marqueurs discursifs ; 2) annotation de la fonction pragmatique des marqueurs identifiés ; 3) annotation des disfluences adjacentes à un marqueur discursif. Chaque étape sera brièvement présentée ci-dessous.

4.1. Identification des marqueurs discursifs

La sélection des marqueurs discursifs est intégralement manuelle et onomasiologique (c.-à-d. sans liste fermée) et se base sur la définition formelle et fonctionnelle proposée dans Crible (2017b) qui suggère les critères prescriptifs suivants : stricte optionalité syntaxique ; portée discursive (présence d'un prédicat conjugué) ; haut degré de grammaticalisation ; sens procédural exprimant une relation discursive (par exemple cause, contraste, conséquence, etc.), une attitude subjective, une fonction structurante (gestion des thèmes et tours de parole) ou une adresse à l'interlocuteur. Cette définition inclut *de facto* des expressions variées telles que des conjonctions de coordination et de subordination, des adverbes, des locutions verbales, etc., à condition qu'elles remplissent les critères mentionnés ci-dessus. En revanche, elle exclut certains cas-limites comme les pauses pleines (*euuh*), les interjections (*ah, oups*) ou les marques de réponse (*oui, non*), qui sont considérées comme des catégories pragmatiques distinctes.

Pour ce qui concerne la présente étude et l'analyse des conjonctions *et / and*, il convient de préciser que seules les occurrences reliant deux énoncés au statut de proposition principale ou subordonnée ont été incluses. Ce critère inclut les cas d'ellipse du sujet (*il arrive et repart*) et exclut les occurrences reliant des groupes nominaux, des propositions relatives ou toute autre unité syntaxiquement et fonctionnellement incomplète. Ces décisions ont sans doute des conséquences sur la quantité des marqueurs discursifs identifiés, surtout si l'on considère la syntaxe complexe et souvent pluri-nucléaire de l'oral.

4.2. Annotation fonctionnelle des marqueurs discursifs

Une fois identifiés, les marqueurs discursifs sont désambiguïsés au niveau de leur fonction pragmatique sur la base d'une typologie élaborée par Crible (2017b), qui distingue trente fonctions discursives réparties dans quatre grands domaines, représentés dans le Tableau 2.

Idéationnel	Rhétorique	Séquentiel	Interpersonnel
cause	motivation	ponctuation	suivi
conséquence	conclusion	ouverture	politesse (face)
concession	opposition	fermeture	désaccord
contraste	spécification	retour au thème	accord
alternative	reformulation	nouveau thème	ellipse
condition	pertinence	Citation	
temporel	emphase	addition	
exception	commentaire	énumération	
	approximation		

Tableau 2 – Typologie fonctionnelle des marqueurs discursifs

Le domaine idéationnel comprend les relations de cohérence portant sur des événements objectifs et externes au discours (1). Le domaine rhétorique comprend des fonctions méta-

linguistiques ainsi que des relations de cohérence subjectives ou épistémiques (2) ^[1]. Le domaine séquentiel correspond à la structuration des segments du discours tels que les tours de parole ou les sujets thématiques (3). Enfin, le domaine interpersonnel concerne la gestion interactive de l'échange et la relation entre les locuteurs (4).

(1) we don't play any more *because* it gets dark too early (EN-conv-02) ^[2]

(2) they're going to uhm (1.293) they've bought a house *well* he has (0.530) and they're going to run a bookshop (EN-conv-02)

(3) here's a napkin oops (0.280) *by the way* did I mention my dustbin's been blown over (EN-conv-04)

(4) I thought that one *you know* the Brie de Meaux is quite good (EN-conv-07)

Cette typologie est principalement inspirée du *Penn Discourse TreeBank 2.0* (Prasad *et al.*, 2008) et de la proposition de González (2005). Son application au corpus a été réalisée par un seul annotateur, aidé d'un manuel (Crible, 2014) fournissant d'amples définitions, critères et exemples. Elle n'en reste pas moins complexe étant donnée la multifonctionnalité des marqueurs discursifs et leur tendance à la sous-spécification. Une mesure d'accord intra-annotateur a été calculée afin d'évaluer la robustesse de la démarche sur un échantillon de plus de 1000 occurrences : l'analyse a révélé un score de

κ

= 0.779 (84% d'accord) au niveau des quatre domaines et

κ

= 0.74 (75.8% d'accord) au niveau des trente fonctions, ce qui indique une fiabilité satisfaisante du protocole.

4.3 Annotation des disfluences

Les disfluences adjacentes aux marqueurs discursifs ont été identifiées manuellement sur la base d'un protocole élaboré par Crible *et al.* (2016). Les auteurs distinguent les disfluences « simples » des disfluences « composées » selon que leur structure interne est constituée d'un seul ou de plusieurs éléments. Les disfluences simples comprennent les pauses vides et pleines, les termes d'édition (par exemple *oups, je me trompe*), les fragments de mots (par exemple *bat- bateau*), les faux départs et les marqueurs discursifs eux-mêmes. Les disfluences composées comprennent quant à elles les répétitions (par

exemple *le le bateau*) et les substitutions morphosyntaxiques ou propositionnelles (par exemple *le la maison*). Cette annotation utilise un système de crochets ouvrants et fermants pour indiquer le début et la fin d'une disfluece composée, comme dans l'Exemple 5.

(5) uhm so she and she (EN-clas-05)

(0.200)	uhm	so	she	and	she
<UP>	<FP>	<DM>	<RI0	<DM>	RI1>

Dans cet exemple, la pause silencieuse (<UP> pour *unfilled pause*) est suivie d'une pause pleine (<FP> pour *filled pause*), du marqueur discursif *so* 'donc' (<DM>), de la répétition du pronom *she* (<RI> pour *identical repetition*) dans laquelle vient s'insérer le marqueur discursif *and* 'et'. Ces éléments adjacents forment une séquence de disfluences, qui sera par la suite résumée en différentes étiquettes plus ou moins précises sur son contenu et sa structure.

Après l'annotation, toutes ces informations aux niveaux des marqueurs discursifs et des disfluences sont automatiquement extraites et traitées dans une base de données afin d'analyser la distribution en genre et en langue des différentes configurations rencontrées dans le corpus.

5. Résultats quantitatifs sur corpus

L'analyse suivante ne porte que sur les sous-corpus de *DisFrEn* correspondant à des discours intégralement préparés d'une part (discours politique, journal parlé) et complètement spontanés d'autre part (conversation, appel téléphonique, commentaires sportifs). Dans cet échantillon, un total de 532 occurrences de *and* et 490 occurrences de *et* ont été identifiées et annotées d'après la méthode détaillée dans les sections précédentes. En tout, 25 types de fonctions (ou combinaisons de fonctions) différentes ont été attribuées à ces deux marqueurs discursifs, telles que « addition », qui est leur sens de base, mais aussi « conséquence », « spécification » ou encore « commentaire ». Pour faciliter le traitement quantitatif de ces données, nous allons nous concentrer uniquement sur les cas représentant au moins 10 occurrences par type de fonction. Les données ainsi obtenues sont représentées dans le Tableau 3.

Fonction	Préparé		Spontané		Total
	<i>and</i>	<i>et</i>	<i>and</i>	<i>et</i>	
Addition	81	58	255	221	615
Conséquence	6	5	62	28	101
Spécification	5	5	28	42	80
Nouveau sujet	4	6	10	43	63
Ponctuation	0	0	16	19	35
Temporel	0	0	15	10	25
Conclusion	1	5	10	6	22
Contraste	0	0	3	14	17
Énumération	1	2	7	5	15
Concession	0	0	4	9	13
Total	98	81	410	397	986

Tableau 3. Fonctions de *et/ and* en discours préparé vs. spontané

Il apparaît que pour chacun des deux marqueurs discursifs, le sens additif de base est de loin le plus fréquemment exprimé, que ce soit en contexte préparé ou spontané (6). Il est suivi par les fonctions de conséquence (7), de spécification (ajout de précision ou d'exemple, 8) et de changement de sujet ou thème conversationnel (9).

(6) j'ai dormi en bas *et* Jojo avait descendu aussi son matelas (FR-conv-03)

(7) elle avait un problème de boiler *et* j'ai été lui lui remettre un nouveau quoi (FR-conv-05)

(8) <A> tu as couru quoi non non je marchais ah non non j'ai pas couru (0.180) *et* j'ai fait encore un détour pour aller trouver une clope (FR-conv-05)

(9) roulette pied (1.468) d'un pied *et* enchaînement avec passe de l'autre (1.102) *et* le président Chirac il n'a pas revêtu le maillot euh français (FR-spor-04)

On voit que ces différents emplois de *et* ne sont pas tous égaux en termes de valeur subjective. En effet, alors qu'en (6) le locuteur ne fait qu'énoncer deux faits objectifs sans prise de position apparente, les trois autres exemples développent un point de vue plus personnel du locuteur sur son énoncé : en (7), le locuteur souligne le lien de cause à effet entre les deux segments coordonnés par le marqueur *et* ; en (8), il ajoute une précision qui vient soutenir son argument, afin de guider l'interlocuteur vers la conclusion souhaitée (*je n'ai pas couru et en plus j'ai fait un détour, donc ce n'était pas loin*) ; le commentateur sportif en (9) oriente explicitement le discours vers un nouveau sujet qui s'écarte de la description du match pour commenter un détail vestimentaire qu'il estime inattendu ou intéressant. Ces exemples illustrent non seulement la compatibilité de *et* avec un panel d'interprétations variées, mais aussi sa flexibilité sur le continuum objectivité-subjectivité, qui s'applique également à *and* et aux relations discursives en général.

Bien que toujours majoritaire, la fonction additive ne l'est toutefois pas dans la même proportion selon le genre discursif, en particulier en anglais, où 83% des *and* sont additifs dans les discours préparés contre seulement 62% à l'oral spontané. L'écart est moins grand mais reste significatif en français, avec 72% de *et* additifs en discours préparé contre 56% en discours spontané. Ces différences de proportions sont statistiquement significatives ($z = 3.845$, $p < 0.0001$ en anglais ; $z = 2.652$, $p < 0.01$ en français) et confirment notre hypothèse : la proportion d'emplois « spécifiés » (additifs) des marqueurs discursifs augmente à mesure que le discours est plus formel et que le locuteur se préoccupe plus des besoins de compréhension de ses interlocuteurs (principe Q). À l'inverse, les fonctions sous-spécifiées de *et/ and* telles que la conséquence ou le contraste sont nettement plus fréquentes à l'oral spontané, dans un souci d'économie de production (principe R). Sur la base de ces premiers résultats quantitatifs, il semble donc bien que la subjectivité, entendue comme égo-centrisme, prenne le pas sur l'intersubjectivité (ou allo-centrisme) lorsque la pression temporelle et cognitive pèse sur le locuteur en situation de production spontanée.

En examinant de plus près les fonctions non-basiques de *et / and* dans notre échantillon, on peut noter que certaines d'entre elles sont complètement absentes des genres préparés en anglais comme en français. C'est le cas notamment des relations contrastives et concessives, qui demeurent par ailleurs assez rares à l'oral spontané également. Cela ne signifie pas que ces trois types de relations du discours ne sont pas du tout utilisées par les locuteurs à l'oral préparé, mais plutôt qu'elles sont exprimées par d'autres marqueurs discursifs plus spécifiques comme *mais* ou *pourtant*. La valeur négative d'opposition entre deux énoncés, inhérente aux relations de contraste et de concession, semble donc requérir un marquage plus précis qu'un simple coordonnant additif, afin de mieux restituer la perspective du locuteur dans des contextes de parole formelle, comme en (10).

(10) Tripoli n'est pas encore près de tomber pour autant *mais* les insurgés contrôlent à présent (0.431) la côte (FR-news-04)

Dans cet exemple, le journaliste rapporte une concession entre deux constats apparemment factuels et objectifs (la capitale résiste mais les insurgés progressent). Pourtant, la suppression du marqueur *mais* ou sa substitution par un *et* sous-spécifié pourrait nuire à l'interprétation concessive, en faisant apparaître les deux segments comme reliés causalement (*la capitale résiste et/donc les insurgés progressent*), d'autant plus que le discours journalistique traite d'événements parfois compliqués relatés à un public potentiellement non-expert. L'emploi de *mais* plutôt que *et* dans ce type de contexte relève donc d'une stratégie intersubjective visant à faciliter le traitement de l'information en spécifiant la relation négative entre deux énoncés objectifs.

Au niveau de la comparaison anglais-français, les deux langues semblent privilégier des emplois sous-spécifiés différents, avec une nette préférence pour les relations de conséquence marquées par *and* (62 occurrences en anglais spontané), tandis que *et* est plus souvent utilisé pour introduire plus de précisions (« spécification », 42 occurrences en français spontané) ou un nouveau thème conversationnel (« nouveau sujet », 43 occurrences). Ces préférences sont confirmées par l'analyse de tous les marqueurs discursifs exprimant ces différentes fonctions dans les registres spontanés :

- le marqueur le plus fréquent pour exprimer la *conséquence* en anglais est bien *and*, suivi de *so* avec 53 occurrences seulement ; en français, *et* et *donc* se partagent la première place avec 28 et 27 occurrences, respectivement ;
- le marqueur de spécification le plus fréquent est de loin *et* en français, devant *alors* avec seulement 21 occurrences ; *and* n'arrive qu'en deuxième place derrière les 32 occurrences de *I mean* ;

- l'introduction d'un nouveau sujet ou thème est principalement exprimée par *et*, suivi de loin par *mais* (17 occurrences) ; *and* est plutôt rare dans cet emploi, juste derrière *so* (10 et 11 occurrences respectivement).

En somme, les deux langues semblent adopter des stratégies de sous-spécification qui visent des fonctions différentes, pour lesquelles *et* / *and* n'ont peu ou pas de concurrents directs. Dans l'ensemble, le principe égo-centrique d'économie semble prévaloir à l'oral spontané. On pourrait alors suggérer que la création de la cohérence d'un discours y est envisagée comme une activité coopérative qui requiert un effort de reconstruction de la part de l'interlocuteur, forcé de compléter les marques imprécises de la subjectivité du locuteur.

6. Analyse qualitative des stratégies compensatoires

Comme souligné par de nombreux auteurs, les marqueurs discursifs ne sont pas les seuls éléments responsables de la cohérence d'un discours, et leur omission ne « supprime » pas nécessairement la relation qu'ils signalent entre deux énoncés (voir le débat dans la littérature, par exemple Taboada, 2006 vs Hansen, 2006 : 26). De même, une relation causale exprimée par un marqueur discursif sous-spécifié comme *et* n'est pas nécessairement sous-spécifiée elle-même, car d'autres indices linguistiques et contextuels peuvent remplir ce rôle de marquage et renforcer l'information incomplète donnée par la simple conjonction additive. Plusieurs études expérimentales ont même montré que, dans le cas des relations causales, un marquage explicite n'entraîne pas forcément un traitement cognitif plus rapide et plus efficace (Bestgen & Vonk, 1995, Graesser *et al.*, 1994). Ce phénomène est dû au principe bien connu de *post hoc ergo propter hoc* par lequel les lecteurs et auditeurs ont tendance à inférer des relations causales entre deux événements consécutifs, sauf indication contraire. Toute stratégie compensatoire visant à soutenir l'interprétation causale entre deux énoncés est donc potentiellement redondante si le contenu des énoncés eux-mêmes suffit à suggérer cette interprétation. Dans le cas de *et* / *and*, leur valeur temporelle, activée dans un contexte narratif entre deux énoncés adjacents, est donc souvent accompagnée d'une valeur causale, comme dans l'exemple suivant.

(11) so he asks all these well asks around advertises in the paper *and* gets all these weirdos coming to the door (EN-conv-02)

En (11), le fait introduit par *and* (recevoir des gens bizarres à sa porte) est présenté comme la conséquence directe de l'énoncé précédent (faire de la publicité, demander autour de soi) : l'inférence causale ajoutée à la succession temporelle ne requiert ici aucun marquage supplémentaire, ce qui est reflété par l'absence de rupture prosodique ou syntaxique (pas de pause, ellipse du pronom-sujet « *he* »). Parmi toutes les occurrences similaires de *and* exprimant une conséquence, il est remarquable qu'aucune ne corresponde à la séquence *and uh* ou *and uhm*, pourtant si fréquente en anglais spontané : l'absence notable des pauses pleines tend à confirmer la « suffisance » pragmatique du *and* causal.

En revanche, dans les cas de relations négatives comme le contraste ou la concession, mentionnées dans la section précédente, un marqueur additif ne suffit pas à annuler le processus d'inférence temporelle et causale activé automatiquement selon le principe *post hoc ergo propter hoc*. Pourtant, contrairement à notre hypothèse, ces cas sous-spécifiés de *et / and* ne sont pas plus souvent associés à des disfluences que leurs emplois additifs ou causaux. En analysant les occurrences du *et* contrastif, il apparaît que l'opposition est toujours marquée par des éléments lexicaux et syntaxiques. La compensation ne se joue donc pas au niveau du marqueur discursif ou de son entourage direct mais est plutôt encodée dans les énoncés eux-mêmes. L'exemple (12) illustre l'emploi d'indices déictiques (« là-bas » vs « ici ») ; en (13), on peut observer le rôle d'un adverbe temporel (« maintenant »), de la position initiale dans le tour de parole ainsi que la fonction contrastive des temps verbaux.

(12) ils trouvent pas de techniciens là-bas *et* ici ils en ont de trop (FR-conv-05)

(13) <A> il achetait des bagnoles de merde *et* maintenant c'est fini ? <A> ah oui il a acheté une nouvelle camionnette (FR-conv-05)

Dans tous ces exemples où le contraste est rendu explicite par le contenu des énoncés, la fonction du marqueur discursif ne semble plus relever de l'expression de la perspective ou de la subjectivité du locuteur vis-à-vis de son discours, mais consiste seulement en un rôle de liaison, de ponctuation afin d'augmenter la fluidité des enchaînements. Dans une autre perspective, on peut envisager que la valeur sous-spécifiée de *et / and* le rend propice à une fonction de renforcement pragmatique ou d'emphase d'éléments co-textuels. Il est difficile d'établir si ce renforcement est principalement au bénéfice du locuteur (principe R) ou de l'interlocuteur (principe Q). Toutefois, le caractère optionnel du marqueur dans les exemples ci-dessus nous invite à l'envisager comme un effort volontaire à caractère allo-centré et intersubjectif. On peut même se demander si un marqueur plus spécifique de contraste comme *mais* ou *à l'inverse* ne serait pas contre-productif par surcharge informative :

(12 ?) ils trouvent pas de techniciens là-bas *mais* ici ils en ont de trop

(13 ?) <A> il achetait des bagnoles de merde *mais* maintenant c'est fini ?

Bien que possibles, ces énoncés modifiés par le marqueur spécifique *mais* ne sont pas nécessairement plus faciles à interpréter que leurs versions sous-spécifiées originales, contrairement à la concession en (10) qui, elle, n'est pas marquée par d'autres éléments lexicaux ou syntaxiques (voir Section 5 ci-dessus). Cette observation évoque la théorie de *Uniform Information Density* (Levy & Jaeger, 2007) qui postule que la quantité d'information contenue dans un énoncé reste plus ou moins stable d'un mot à l'autre. Il semble donc que la sous-spécification ne relève pas tant d'une incomplétude mais plutôt d'une dynamique informationnelle rythmée par les capacités de production du locuteur d'une part et les besoins interprétatifs de l'interlocuteur d'autre part.

Conclusion

Dans cette étude de cas, nous avons étudié les usages sous-spécifiés du marqueur discursif *et/ and* dans un corpus oral à travers des registres formels et informels de l'anglais et du français. Notre analyse quantitative et qualitative révèle que la sous-spécification existe aussi bien à l'oral préparé que spontané mais qu'elle est plus fréquente dans le deuxième cas. L'association entre sous-spécification, spontanéité et égo-centrisme doit toutefois être nuancée au vu des différentes stratégies de compensation (discursives, syntaxiques) qui permettent de restituer la perspective du locuteur, en combinaison avec le marqueur discursif qui remplit alors un rôle de connexion. Il semble donc que la distinction entre subjectivité et intersubjectivité ne suffise pas à rendre compte de la multifonctionnalité des marqueurs discursifs, qui interagissent avec d'autres catégories linguistiques et forment ainsi une variable complexe dans la production et l'interprétation du discours.

Ces résultats préliminaires nous invitent à adopter une approche systématique des stratégies compensatoires de la sous-spécification afin d'en recenser les différentes formes (lexicales, syntaxiques, prosodiques) à travers les genres et les langues. Nos résultats pour la paire *et/ and* pourraient également être testés sur d'autres marqueurs discursifs exprimant d'autres relations du discours, afin notamment de poursuivre l'analyse du statut cognitif particulier des relations « négatives » ou d'opposition.

Remerciements

Je tiens à remercier Raluca Nita et Ramon Marti-Solano pour leur travail d'édition de ce numéro, ainsi que les deux relecteurs anonymes pour leurs commentaires judicieux. Cette

recherche a bénéficié du soutien du projet ARC « A Multi-Modal Approach to Fluency and Disfluency Markers », financé par la Fédération Wallonie-Bruxelles (bourse n° 12/17-044).

Bibliographie

Andersen, G., 1997, « They like wanna see like how we talk and all that. The use of *like* as a discourse marker in London teenage speech » in Ljung, M. (dir.), *Corpus-based Studies in English*, Amsterdam, Rodopi, p. 37-48.

Anscombre, J.-C. & Ducrot, O., 1983, *L'Argumentation dans la Langue*, Liège-Bruxelles, Mardaga.

Arnold, J., Fagnano, M. & Tanenhaus, M., 2003, « Disfluencies signal thee, um, new information », *Journal of Psycholinguistic Research* 32(1), p. 25-36.

Asr, F. & Demberg, V., 2012, « Measuring the strength of linguistic cues for discourse relations » in Hajicová, E., Poláková, L. & Mírovský, J. (dir.), *Proceedings of ADACA Workshop*, Bombay, Indian Institute of Technology, p. 33-42.

Beeching, K., 2007, « La co-variation des marqueurs discursifs *bon, c'est-à-dire, enfin, hein, quand même, quoi et si vous voulez* : Une question d'identité ? », *Langue française* n° 154(2), p. 78-93.

Bestgen, Y. & Vonk, W., 1995, « The role of temporal segmentation markers in discourse processing », *Discourse Processes* n° 19, p. 385-406.

Bosker, H., Quené, H., Sanders, T. & de Jong, N., 2014, « Native 'um's elicit prediction of low-frequency referents, but non-native 'um's do not », *Journal of Memory and Language* n° 75, p. 104-116.

Bunt, H., 2012, « Multifunctionality in dialogue », *Computer Speech and Language* n° 25, p. 222-245.

Candea, M., 2000, *Contribution à l'étude des pauses silencieuses et des phénomènes dits "d'hésitation" en français oral spontané*, Thèse de doctorat, Université Paris III.

Carlier, A. & De Mulder, W., 2010, « The emergence of the definite article: *ille* in competition with *ipse* in Late Latin » in Davidse, K., Vandelanotte, L., Cuyckens, H. (dir.), *Subjectification*,

Intersubjectification and Grammaticalization. Topics in English Linguistics, Berlin, Mouton De Gruyter, p. 241-275.

Crible, L., 2014, *Identifying and describing discourse markers in spoken corpora. Annotation protocol*, Version 8. Rapport technique, Université catholique de Louvain.

Crible, L., 2017a, *Discourse Markers and (Dis)fluency across Registers: A Contrastive Usage-Based Study in English and French*, Thèse de doctorat, Université catholique de Louvain.

Crible, L., 2017b, « Towards an operational category of discourse markers: A definition and its model » in Fedriani, C. & Sanso, A. (dir.), *Discourse Markers, Pragmatic Markers and Modal Particles: New Perspectives*, Amsterdam, John Benjamins, p. 101-126.

Crible, L., Dumont, A., Grosman, I. & Notarrigo, I., 2016, *Annotation manual of fluency and disfluency markers in multilingual, multimodal, native and learner corpora*, Version 2.0. Rapport technique, Université catholique de Louvain et Université de Namur.

Degand, L. & Fagard, B., 2012, « Competing connectives in the causal domain: French *car* and *parce que* », *Journal of Pragmatics* n° 44(2), p. 154-168.

Degand, L., Martin, L. & Simon, A.-C., 2014, « LOCAS-F : un corpus oral multigenres annoté », *Actes du 4^e Congrès Mondial de Linguistique Française*, EDP Sciences, Berlin.

Dister, A., Francard, M., Hambye, P. & Simon, A.-C., 2009, « Du corpus à la banque de données. Du son, des textes et des métadonnées. L'évolution de la banque de données textuelles orales VALIBEL (1989–2009) », *Cahiers de l'Institut de Linguistique de Louvain* n° 33(2), p. 113-129.

Fetzer, A. & Fischer, K., 2007, « Introduction » in Fetzer, A. & Fischer K., (dir.), *Lexical Markers of Common Grounds*, Amsterdam, Elsevier, p. 1-13.

Frisson, S., 2009, « Semantic underspecification in language processing », *Language and Linguistics Compass* n° 3(1), p. 111-127.

Goldman, J.-P., Prsir, T. & Auchlin, A., 2014, « C-PhonoGenre: A 7-hour corpus of 7 speaking styles in French: relations between situational features and prosodic properties », *Proceedings of the 9th Language Resources and Evaluation Conference (LREC'14)*, p. 302-305.

González, M., 2005, « Pragmatic markers and discourse coherence relations in English and Catalan oral narrative », *Discourse Studies* n° 77(1), p. 53-86.

- Götz, S., 2013, *Fluency in Native and Nonnative English Speech*, Amsterdam, John Benjamins.
- Graesser, A., Singer, M. & Trabasso, T., 1994, « Constructing inferences during narrative text comprehension », *Psychological Review* n° 101, p. 371-395.
- Grice, P., 1981, « Presupposition and conversational implicature » in Cole, P. (dir.), *Radical Pragmatics*, New York, Academic Press, p. 183-198.
- Grosman, I., 2016, « How do French humorists manage their persona across situations? A corpus study on their prosodic variation » in Ruiz-Gurillo, L. (dir.), *Metapragmatics of Humor: Current Research Trends*, Amsterdam, John Benjamins, p. 147-175.
- Hancil, S., 2011, « Avant-propos » in Hancil, S. (dir.), *Marqueurs Discursifs et Subjectivité*, Rouen, Publications des universités de Rouen et du Havre, p. 7-11.
- Hansen Mosegaard, M.-B., 2006, « A dynamic polysemy approach to the lexical semantics of discourse markers (with an exemplary analysis of French *toujours*) » in Fischer, K. (dir.), *Approaches to Discourse Particles*, Amsterdam, Elsevier, p. 21-41.
- Haselow, A., 2012, « Subjectivity, intersubjectivity and the negotiation of common ground in spoken discourse: Final particles in English », *Language and Communication* n° 32, p. 182-204.
- Horn, L. R., 1984, « Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature » in Schiffrin, D. (dir.), *Meaning, Form and Use in Context: Linguistic Implications*, Washington DC, Georgetown University Press, p. 11-42.
- House, J., 2013, « Developing pragmatic competence in English as a *lingua franca*: using discourse markers to express (inter)subjectivity and connectivity », *Journal of Pragmatics* n° 59, p. 57-67.
- Kohn, K., 2012, « Pedagogic corpora for content and language integrated learning. Insights from the BACKBONE project », *The Eurocall Review* n° 20(2), p. 3-22.
- Lacheret, A., Kahane, S. & Pietrandrea, P., 2014, *Rhapsodie: A Prosodic and Syntactic Treebank for Spoken French*, Amsterdam, John Benjamins.
- Lenk, U., 1998, *Marking Discourse Coherence. Functions of Discourse Markers in Spoken English*, Tübingen, Gunter Narr Verlag.

- Levy, R. & Jaeger, F., 2007, « Speakers optimize information density through syntactic reduction » in Schölkopf, B. Platt, J. & Hoffman, T. (dir.), *Advances in Neural Information Processing Systems (NIPS) 19*, Cambridge, MA, MIT Press, p. 849-856.
- Meteor, M., Taylor, A., MacIntyre, R. & Iyer, R., 1995, *Dysfluency annotation stylebook for the Switchboard Corpus*, rapport technique, University of Pennsylvania.
- Nelson, G., Wallis, S. & Aarts, B., 2002, *Exploring Natural Language: Working with the British Component of the International Corpus of English*, Amsterdam, John Benjamins.
- Paillard, D., 2011, « Marqueurs discursifs et scène énonciative » in Hancil, S. (dir.), *Marqueurs Discursifs et Subjectivité*, Rouen, Publications des universités de Rouen et du Havre, p. 13-32.
- Palisse, S., 1997, « Artisans », « Assureurs », *Conversations Téléphoniques en Entreprise*. Téléchargé sur <<http://clapi-univ.lyon2.fr>> (dernier accès en mars 2014).
- Pallaud, B., Rauzy, S. & Blâche, P., 2013, « Auto-interruptions et disfluences en français parlé dans quatre corpus du CID », *TIPA Travaux Interdisciplinaires sur la Parole et le Langage* [en ligne] n°29.
- Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A. & Webber, B., 2008, « The Penn Discourse TreeBank 2.0 », *Proceedings of LREC, June 2008*, p. 2961-2968.
- Romano, M. & Cuenca, M. J., 2013, « Discourse markers, structure and emotionality in oral narratives », *Narrative Inquiry* n° 23, p. 344-370.
- Schiffrin, D., 1987, *Discourse Markers*, Cambridge, Cambridge University Press.
- Schmidt, T. & Wörner, K., 2012, « EXMARaLDA » in Durand, J., Ulrike, G. & Kristoffersen, G. (dir.), *Handbook on Corpus Phonology*, Oxford, Oxford University Press, p. 402-419.
- Shriberg, E., 1994, *Preliminaries to a Theory of Speech Disfluencies*, Thèse de doctorat, University of California at Berkeley, CA.
- Sperber, D. & Wilson, D., 1989, *La Pertinence : communication et cognition*, Paris, Les Éditions de Minuit.
- Spooren, W., 1997, « The processing of underspecified coherence relations », *Discourse Processes*, p. 149-168.
- Sweetser, E., 1990, *From Etymology to Pragmatics*, Cambridge, Cambridge University Press.

Taboada, M., 2006, « Discourse markers as signals (or not) of rhetorical relations », *Journal of Pragmatics* n° 38, p. 567-592.

Traugott, E., 2007, « (Inter)subjectification and unidirectionality », *Journal of Historical Pragmatics* n° 8, p. 295-309.

Webber, B., 2013, « What excludes an alternative in coherence relations? », *Proceedings of International Workshop on Computational Semantics*.

Notes

[1] Cf. la distinction entre relations de contenu (*content*), épistémiques (*epistemic*) et entre des actes de parole (*speech-act*) introduite par Sweetser (1990).

[2] Les références aux exemples correspondent au code de la transcription dont ils sont extraits (ici, la deuxième conversation du corpus anglais).

Pour citer ce document

Par Ludivine Crible, « Emplois sous-spécifiés des marqueurs discursifs et / and à l'oral : stratégie (inter)subjective et variation en genre », *Cahiers FoReLLIS - Formes et Représentations en Linguistique, Littérature et dans les arts de l'Image et de la Scène* [En ligne]. Traces de subjectivité et corpus multilingues, III. Les données des corpus comparables : comparaison des langues dans des genres ciblés, mis à jour le : 06/09/2019, URL : <https://cahiersforell.edel.univ-poitiers.fr/index.php?id=695>

Quelques mots à propos de : [Ludivine Crible](#)

Université Catholique de Louvain, Belgique

Cet article est mis à disposition selon les termes de la Licence Creative Commons : Attribution - Pas d'Utilisation Commerciale 3.0 France. CC BY-NC 3.0 FR

