



HAL
open science

Les régulateurs verbaux de connivence : proposition de typologie fonctionnelle et réflexion sur l'influence de la proximité communicative sur la variation observée

Aziliz Philippot

► To cite this version:

Aziliz Philippot. Les régulateurs verbaux de connivence : proposition de typologie fonctionnelle et réflexion sur l'influence de la proximité communicative sur la variation observée. Sciences de l'Homme et Société. 2021. dumas-03516685

HAL Id: dumas-03516685

<https://dumas.ccsd.cnrs.fr/dumas-03516685>

Submitted on 7 Jan 2022

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.



Les régulateurs verbaux de connivence :
proposition de typologie fonctionnelle et
réflexion sur l'influence de la proximité
communicative sur la variation observée

**Aziliz
PHILIPPOT**

Sous la direction de CYRIL TRIMAILLE & d'AGNÈS TUTIN

Laboratoire : LIDILEM

UFR LLASIC

Département Sciences du langage et Français Langue Etrangère
Section Sciences du langage

Mémoire de master 2 mention Sciences du langage - 20 crédits

Parcours : Linguistique

Année universitaire 2020-2021

Les régulateurs verbaux de connivence :
proposition de typologie fonctionnelle et
réflexion sur l'influence de la proximité
communicative sur la variation observée

**Aziliz
PHILIPPOT**

Sous la direction de CYRIL TRIMAILLE & d'AGNÈS TUTIN

Laboratoire : LIDILEM

UFR LLASIC

Département Sciences du langage et Français Langue Etrangère
Section Sciences du langage

Mémoire de master 2 mention sciences du langage - 20 crédits

Parcours : Linguistique, orientation recherche

Année universitaire 2020-2021

Remerciements

Je tiens à remercier Alexis Ladreyt pour toute l'aide qu'il m'a apporté et pour sa réactivité. Tous ses conseils, et sa présence même, m'ont beaucoup confortée et m'ont soutenue dans ce long travail de mémoire. Je tiens aussi à remercier Isabelle Rousset, qui m'a aidée lors de mes problèmes techniques avec le logiciel ELAN en s'occupant elle-même de mes fichiers.

Je souhaite également remercier mes directeurs de mémoire, Cyril Trimaille et Agnès Tutin, qui ont été très présents au début de ce travail et qui m'ont accompagnée dans ma prise en main de mon objet d'étude. Je les remercie de plus de m'avoir laissé le champ libre quant à l'objet d'étude en question et à la façon dont je souhaitais le traiter, tout en ayant su m'avertir des écueils éventuels lorsqu'ils apparaissaient.

Pour terminer je souhaiterais remercier toutes les personnes avec lesquelles j'ai pu avoir des conversations, qu'elles tournent autour de mon sujet ou non, pendant ce travail de mémoire. Ces conversations m'ont permis, consciemment ou non, de tester ma typologie et de réfléchir à mon objet d'étude en temps réel, comme c'est toujours le cas pour un travail sur des phénomènes linguistiques inscrits au cœur des interactions.

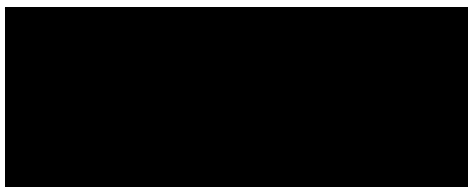
DÉclaration ANTI-PLAGIAT

1. Ce travail est le fruit d'un travail personnel et constitue un document original.
2. Je sais que prétendre être l'auteur d'un travail écrit par une autre personne est une pratique sévèrement sanctionnée par la loi.
3. Personne d'autre que moi n'a le droit de faire valoir ce travail, en totalité ou en partie, comme le sien.
4. Les propos repris mot à mot à d'autres auteurs figurent entre guillemets (citations).
5. Les écrits sur lesquels je m'appuie dans ce mémoire sont systématiquement référencés selon un système de renvoi bibliographique clair et précis.

PRENOM : AZILIZ.....

NOM : PHILIPPOT.....

DATE : 02/09/2021.....



Sommaire

Remerciements	3
Sommaire	5
Introduction	7
Partie 1 - État de l'art	9
CHAPITRE 1. APPROCHES THEORIQUES ET METHODOLOGIQUES	10
CHAPITRE 2. QUELQUES FONDEMENTS NECESSAIRES A L'APPREHENSION DU FONCTIONNEMENT DES REGULATEURS : L'ORGANISATION DE LA PAROLE DANS LES CONVERSATIONS ET LA DIMENSION CONCEPTIONNELLE DU LANGAGE	14
1. L'ORGANISATION DES CONVERSATIONS EN TOURS DE PAROLES ET LES REGLES DE REPARTITION DES TOURS (SACKS, SCHEGLOFF & JEFFERSON, 1974)	14
2. KOCH & CESTERREICHER (2001) ET LA DIMENSION CONCEPTIONNELLE DU LANGAGE	18
CHAPITRE 3. APPROCHE ET DEFINITION DES REGULATEURS VERBAUX DE CONNIVENCE	23
1. LA NOTION GENERALE DE BACKCHANNEL : UNE MULTITUDE D'ETUDES, MAIS PEU DE CONSENSUS	23
2. DE L'IMPORTANCE DES REGULATEURS DANS L'INTERACTION	29
3. LES REGULATEURS VERBAUX DE CONNIVENCE	33
CHAPITRE 4. LA TYPOLOGIE DES REGULATEURS VERBAUX DE CONNIVENCE	37
1. INFLUENCE DES CLASSIFICATIONS FONCTIONNELLES EXISTANTES SUR LA TYPOLOGIE DES RVC : PRISE EN COMPTE DE CERTAINS POINTS COMMUNS ET PROBLEMATIQUES	37
2. LES REGULATEURS COMME MARQUES D'ENREGISTREMENT ET DE CONFIRMATION	43
3. LA FONCTION EXPRESSIVE	45
4. QUELQUES DIFFICULTES LIEES A L'ELABORATION D'UNE TYPOLOGIE POUR LES RVC, ET AU TRAITEMENT DES REGULATEURS EN GENERAL	49
Partie 2 - Méthodologie	53
CHAPITRE 5. CHOIX ET ELABORATION DU CORPUS	54
CHAPITRE 6. ÉLABORATION DE LA TYPOLOGIE	59
CHAPITRE 7. ANNOTATION DU CORPUS	64
1. L'ETABLISSEMENT DE SCORES DE PROXIMITE DANS LE CORPUS	64
2. TRAITEMENT DES DONNEES SOUS ELAN	74
CHAPITRE 8. TRAITEMENT DES DONNEES OBTENUES	79
Partie 3 - Analyse et discussion des résultats	82
CHAPITRE 9. RESULTATS DE L'ANALYSE DU CORPUS	83

1. OBSERVATIONS GENERALES	83
2. FREQUENCES SELON LE DEGRE DE PROXIMITE	86
3. INVENTAIRE DES FORMES DES RVC DANS LE CORPUS.....	94
4. ANALYSE DETAILLEE DE LA VARIATION FORMELLE DES RVC SELON LES TYPES ET SOUS- TYPES FONCTIONNELS.....	105
CHAPITRE 10. DISCUSSION	114
Conclusion.....	117
Bibliographie.....	119
Table des exemples	123
Table des illustrations et tableaux	124
Table des annexes.....	128
Table des matières	136

Introduction

Un point de vue naïf sur les interactions orales considérerait qu'elles font alterner deux rôles distincts : un rôle actif de parole et un rôle passif d'écoute. Or, à y regarder de plus près, il est clair pour quiconque s'intéresse aux interactions que les rôles joués par les locuteurs sont plus complexes que cela. En particulier, un fait primordial est qu'il n'y a pas de rôle passif dans une interaction (c'est d'ailleurs bien pour cela qu'il s'agit d'une *interaction*). Une preuve de cette co-activité permanente des locuteurs est la présence de comportements, gestuels ou linguistiques (contenu vocal ou lexical), que l'on nomme communément *backchannel* dans la littérature, et auxquels nous nous référerons en tant que *régulateurs*. Ces comportements permettent à un locuteur en position d'écoute de signaler sa bonne réception du discours de son interlocuteur, et son engagement dans la conversation.

(1)

Exemple 1 *d'accord* et *hum* en tant que régulateurs (extrait de coree_ghu_14, TCOF)

- | | | |
|---|----------------|--|
| 1 | -10_TCOF_02_01 | elle avait pas le droit alors que moi au final je n'ai pas un seul ami blanc |
| 2 | | enfin blanc mauricien m- mes amis blancs sont tous des expats |
| 3 | -09_TCOF_02_02 | d'accord |
| 4 | -10_TCOF_02_01 | et j'ai que des amis de couleurs après |
| 5 | -09_TCOF_02_02 | hum |
| 6 | -10_TCOF_02_01 | donc ça c'est (0.2) que le temps d'une génération |

Au sein de ces comportements, les éléments linguistiques possédant un contenu lexical, tels que *d'accord* ou encore *c'est chaud*, sont ceux dont la variation est la plus manifeste. Malgré cela, leurs différentes micro-fonctions ont peu été étudiées. Ce fait est aussi valable pour les autres types de régulateurs : gestuels (hochements de tête, regards, etc.) et vocaux (*hum* et assimilés). Il semble cependant qu'un locuteur produisant un *d'accord* n'a pas tout à fait le même comportement qu'un locuteur produisant un *tu m'étonnes* ou encore un *putain*. Cette variation est-elle le reflet d'une variation fonctionnelle ? Le contexte, et en particulier l'aspect conceptionnel de l'interaction donnée, entre proximité et distance (Koch & Esterreicher, 2001), influence-t-il le type de régulateurs produit ?

Cette variation fonctionnelle des régulateurs, et plus précisément d'un sous-ensemble de régulateurs que nous avons appelés *régulateurs verbaux de connivence*, a constitué l'objet de ce mémoire, d'après la problématique suivante : Comment définir les *régulateurs verbaux*

de connivence et quelles sont les différentes micro-fonctions que ce sous-ensemble de régulateurs peut remplir dans les conversations orales en français hexagonal ?

Afin de répondre à cette problématique, une typologie fonctionnelle des *régulateurs verbaux de connivence* sera proposée, établie grâce à une approche guidée par les données, à partir d'un corpus de conversations de deux à trois locuteurs en français hexagonal. Un intérêt particulier sera de plus porté à la recherche d'un lien possible entre le degré de proximité dans une interaction donnée et la fréquence des différents types envisagés, ainsi qu'à l'évaluation du lien entre le contenu lexical des *régulateurs verbaux de connivence* et leurs fonctions.

Ce mémoire sera constitué de trois parties. La première partie, dédiée à l'état de l'art, permettra de définir dans un premier temps les différentes approches théoriques derrière ce travail de mémoire, puis quelques notions théoriques telles de l'organisation de la parole dans les conversations (Sacks et al., 1974) et l'aspect conceptionnel du langage (Koch & Esterreicher, 2001) dans un deuxième temps. Le cœur du sujet, les régulateurs, sera ensuite abordé, en commençant par une définition de ce phénomène en lui-même, suivie d'une définition des *régulateurs verbaux de connivence* en particulier. La dernière partie de l'état de l'art sera consacrée à la description de la typologie mise en place dans ce mémoire.

Une deuxième partie de ce mémoire se concentrera sur la méthodologie adoptée, de l'établissement du corpus puis de la typologie en passant par l'annotation du corpus et le traitement des données finales.

Pour terminer, la troisième et dernière partie présentera les résultats obtenus (fréquences dans le corpus, en général et selon la proximité, et analyse des formes observées) dans un premier temps. Ces résultats seront suivis d'une discussion des limites de ce mémoire, puis, en dernier lieu, de la conclusion.

Partie 1

-

État de l'art

Chapitre 1. Approches théoriques et méthodologiques

Le présent mémoire ne se réclame pas d'une théorie linguistique précise, mais s'inscrit plutôt dans une certaine approche de la langue : l'approche interactionniste. Ainsi, l'objet d'étude de ce mémoire, les régulateurs (dont nous étudions un sous-ensemble), a surtout été étudié dans le cadre de la pragmatique interactionniste et de l'Analyse Conversationnelle, deux courants relevant d'une approche interactionniste du langage. Nous nous intéresserons de plus à la variation d'usage des régulateurs verbaux¹. Nous en ferons l'analyse en intégrant un aspect sociolinguistique à notre approche interactionniste.

L'approche interactionniste du langage suppose de considérer tout discours oral comme une *interaction*, comme un moyen pour les participants de s'influencer mutuellement. Ainsi, pour Nuchèze & Coletta (2002, p. 2) : « tout discours est co-construit à des degrés divers, adressé, médiatisé par une relation, inscrit dans un contexte ».

Dans le champ de la pragmatique (qui s'intéresse globalement au lien entre le contenu linguistique et ses conditions de production), on retrouve cette approche interactionniste, avec de plus un intérêt majeur pour les données empiriques, dans le cadre de la pragmatique des interactions, ou pragmatique interactionniste (de Nuchèze & Coletta, 2002). Cette orientation interactionniste de la pragmatique se base sur les principes et méthodes de l'analyse conversationnelle, courant linguistique dans lequel s'inscrivent les premiers travaux réalisés sur les régulateurs.

Le courant de l'analyse conversationnelle (désormais AC) est né dans les années 70 aux États-Unis, dans la mouvance ethnométhodologique de la sociologie américaine (de Nuchèze & Coletta, 2002). Si les travaux effectués au sein de ce courant ont d'abord porté sur les conversations quotidiennes, l'analyse conversationnelle s'intéresse à tous types d'interactions. Ainsi, selon Sacks (1984, p. 24) « [son] objet d'étude n'est pas la langue en tant que telle, mais ce que la langue peut faire ou plutôt ce que la langue permet de faire »

¹ C'est-à-dire des régulateurs avec un contenu lexical. Le terme *verbal* dans ce cas est le terme utilisé dans les travaux de de Gaulmyn (1987) entre autres, et correspond au canal de communication.

(Cité par Persson, 2014, p. 8). L'analyse conversationnelle implique une approche inductive, non dictée par des catégories et principes préétablis, fondée sur des données empiriques, et authentiques (Persson, 2014).

Du fait de l'intérêt de l'AC pour l'analyse du fonctionnement des interactions, beaucoup de travaux de ce courant ont porté sur la description de l'organisation des conversations. La description des procédés conversationnels permettant cette organisation, cette structuration mais aussi la synchronisation des participants a aussi été une source importante d'études. Parmi les travaux fondateurs de l'AC référencés dans ce mémoire, on retrouve ainsi le travail de Sacks, Schegloff & Jefferson en 1974 intitulé *A simplest systematics for the organization of turn-taking for conversation*, mais aussi le travail fondateur de Yngve en 1970 (cité par de nombreux auteurs dont Schegloff, 1982 ; Goodwin, 1986 ; Maynard, 1990 et bien d'autres) au sujet de notre objet d'étude, les régulateurs².

Dans le prolongement de l'analyse conversationnelle et de la pragmatique interactionniste, ce travail de mémoire portera donc sur l'étude de certains régulateurs verbaux, avec un objectif de description de leur fonctionnement et rôle(s) dans les interactions que sont les conversations orales privées. La méthodologie adoptée sera inductive, basée sur des données empiriques tirées de corpus oraux authentiques (voir partie III. sur la méthodologie).

L'analyse des régulateurs proposée dans ce mémoire prendra de plus en compte leur variation, dans la mesure du possible. Cet intérêt pour la variation relève du domaine de la sociolinguistique, avec encore une fois une approche interactionniste de cette variation (sociolinguistique interactionnelle). Tout comme la pragmatique, la sociolinguistique s'intéresse au langage en contexte, en se plaçant plutôt au niveau du contexte des locuteurs (variables sociales). Elle cherche ainsi, dans son approche générale, à établir des liens entre les variations observées et des variables sociales. En s'intéressant à des interactions authentiques, elle cherche à analyser à la fois les variations internes (qui dépendent du type de relation interpersonnelle et des rôles liés à une situation donnée), mais aussi les variations

² Nommés alors *backchannel* dans ce premier travail de référence

externes (qui dépendent des valeurs sociales et culturelles revendiquées par les participants) (de Nuchèze & Coletta, 2002).

L'aspect sociolinguistique développé dans ce mémoire est en lien avec les paramètres situationnels décrits par Koch & Eesterreicher (2001), et l'aspect conceptionnel du langage. Ces différents paramètres ainsi que la définition de l'aspect conceptionnel du langage (proximité versus distance) seront revus dans le chapitre 2. La prise en compte de ces paramètres nous permettra d'évaluer une possible influence du degré de proximité, et donc de connivence entre les interactants, sur les régulateurs dans notre corpus.

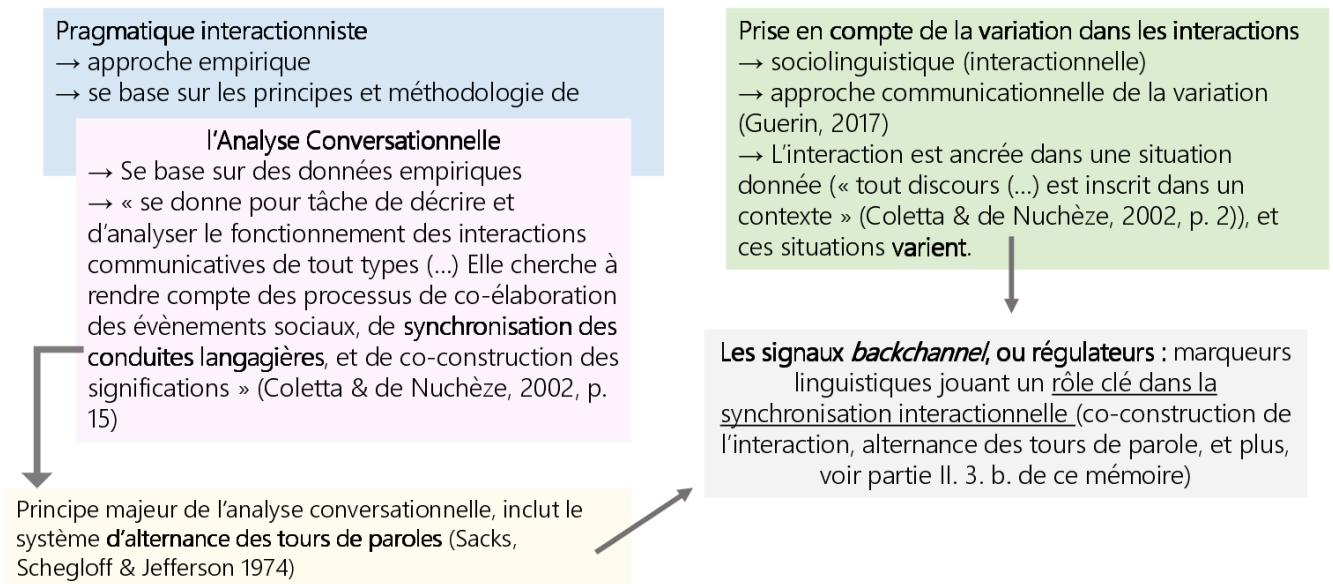
Un fort intérêt sera plus généralement porté sur la variation fonctionnelle des régulateurs verbaux étudiés (établissement d'une typologie de leurs différentes micro-fonctions pragmatiques). Cette variation fonctionnelle sera ensuite mise en lien avec la variation formelle (contenu lexical et structures syntaxiques) observée, afin de réfléchir à un lien possible entre forme et fonction pour les *régulateurs verbaux de connivence*.

L'approche interactionniste adoptée dans ce mémoire implique un cadre méthodologique précis, comme l'indiquent les descriptions des différents courants et approches qui y sont liées. Cette méthodologie est ainsi caractérisée par l'utilisation de données empiriques, authentiques, sur lesquelles fonder les observations réalisées. De plus, l'approche des données est une approche inductive, guidée par le corpus plutôt que basée sur des catégories et formes préexistantes. Les outils d'analyses en lien avec cette méthodologie sont ceux de la linguistique de corpus, outils qui seront détaillés dans la partie 2 sur la méthodologie.

Pour résumer, nous allons analyser dans ce mémoire un sous-ensemble de régulateurs verbaux, que nous appellerons *régulateurs verbaux de connivence*. Ces derniers seront analysés avec une approche interactionniste, dans le prolongement des travaux existants sur les régulateurs en analyse conversationnelle et pragmatique des interactions. En lien avec cette approche, un intérêt pour la variation (typiquement étudiée dans le cadre de la sociolinguistique) de ces régulateurs selon des critères liés à l'interaction sera aussi exploré dans l'analyse des données.

Le schéma ci-dessous récapitule les principes majeurs de ces différents courants et approches interactionnistes, et la façon dont ils peuvent être liés à l'étude des régulateurs.

Vue d'ensemble des approches interactionnistes du langage permettant d'appréhender et d'étudier les signaux *backchannel*, ou régulateurs



Aziliz Philippot – Université Grenoble Alpes | Mémoire de Master 2 – 2020-2021

Figure 1 Vue d'ensemble des approches interactionnistes du langage permettant l'appréhension et l'étude des régulateurs

Ce cadre méthodologique et l'ancrage de notre objet d'étude dans une lignée de travaux relevant d'une approche interactionniste du langage étant fixé, certains éléments restent cependant à éclaircir avant d'entrer dans le vif du sujet et la définition de notre objet d'étude : les *régulateurs verbaux de connivence*. En effet, deux études précédemment mentionnées portent sur des notions et principes dont la compréhension est nécessaire à l'appréhension et l'examen des régulateurs : l'article de Sacks, Schegloff & Jefferson (1974) sur les mécanismes de répartition des tours de parole dans une conversation, ainsi que celui de Koch & Cesterreicher (2001) sur l'ordre conceptionnel (vs medial) du langage (paramètres situationnels et continuum proximité/distance). Nous expliciterons ainsi dans un premier temps ces deux éléments, avant de pouvoir avancer sur la définition générale des régulateurs.

Chapitre 2. Quelques fondements nécessaires à l'appréhension du fonctionnement des régulateurs : l'organisation de la parole dans les conversations et la dimension conceptionnelle du langage

Les régulateurs sont des éléments linguistiques dont le rôle est nécessairement ancré dans l'interaction, comme nous le verrons dans le chapitre 3. Il est donc nécessaire de comprendre pour les étudier tout d'abord le fonctionnement général et les principes de la répartition des tours de paroles dans l'interaction, mais aussi la différence entre l'ordre médial et conceptionnel de la parole, et les détails et implications de l'ordre conceptionnel en particulier. La discussion de ces deux notions sera basée principalement sur les définitions établies dans les travaux fondateurs de Sacks, Schegloff & Jefferson (1974, répartition des tours de parole) et de Koch & Cesterreicher (2001, dimension médiale vs conceptionnelle du langage et continuum proximité vs distance).

1. L'organisation des conversations en tours de paroles et les règles de répartition des tours (Sacks, Schegloff & Jefferson, 1974)

Les interactions étant par nature co-construites par leurs participants, il est nécessaire pour les étudier de comprendre les règles et principes fondamentaux qui assurent une bonne répartition de la parole et donc une bonne co-construction par les participants, règles basées entre autres sur l'absence de pauses dans la parole et l'absence (ou quasi-absence) de moments de négociations explicite de la parole.

Sacks, Schegloff & Jefferson font partie des premiers chercheurs à s'être penchés sur l'aspect systématique de l'organisation de la parole dans la conversation³, et à avoir tenté de décrire cette organisation avec des règles et des principes, qu'ils décrivent dans leur article fondateur de 1974 : *A simplest systematics for the organisation of turn-taking for*

³ Il s'agit bien de l'organisation de la parole dans la *conversation* uniquement, puisqu'il s'agit du seul type d'interaction auquel Sacks et al. se sont intéressés et sur lequel se base leur argumentation. Des règles de distribution de la parole sont aussi observées dans d'autres types d'interaction, mais elles pourraient être différentes (Sacks et al., 1974).

conversation. Cet article, établi dans le cadre de l'analyse conversationnelle, a ensuite été repris dans de nombreux travaux.

Une conversation est constituée de différents tours de paroles, répartis et administrés entre eux par les participants de manière localisée. Le tour de parole est ainsi une unité purement interactionnelle, avec une forme et un contenu variable, négociés par les participants dans une conversation donnée (Traverso, 2016 ; Sacks et al., 1974). Dans ce mémoire, nous définirons le tour de parole comme une prise de parole par un locuteur se terminant par un changement de locuteur. Cette définition correspond à la définition initiale d'un tour de parole (de Nuchèze & Coletta, 2002, p. 186). Elle rejoint de plus la définition de Sacks et al. (1974), qui parlent bien d'un système d'*alternance* des tours de parole.

Selon Sacks et al. (1974), chaque tour de parole est constitué d'une ou de plusieurs *turn-constructive unit* (TCU, ou « unités de composition de tours »). La taille et le contenu des TCU sont variables, allant d'un seul mot à des constructions complexes. Persson explique dans ce sens que les TCU ne sont pas des unités formelles « pré-fabriquées », mais peuvent plutôt être appréhendées comme des épiphénomènes des différentes pratiques des interactants (2014, p. 23). Nous suivrons la définition de Selting (2000) dans ce mémoire, qui fixe les TCU comme « les plus petites unités linguistiquement complètes, pertinentes au niveau interactionnel » (cité par Bertand et al., 2009, p. 6). On peut donc considérer les TCU comme des segments de tours de parole.

Chaque TCU se termine par une *transition-relevance place* (TRP, « place de transition possible » en français). Ces places de transitions correspondent aux points d'achèvement possible des segments que sont les TCU. Elles peuvent être projetées et anticipées par les participants, comme le pointent Sacks et al. :

We may note that it is empirically evident, from sequential materials, that projectability is the case; i.e., we find sequentially appropriate starts by next speakers after turns composed of single-word, single-phrase, or single-clause constructions, with no gap-i.e. with no waiting for possible sentence. (Sacks, Schegloff & Jefferson, 1974, p. 702)

La fin d'une TCU, et donc l'emplacement d'un TRP, peut être identifiée par une corrélation de différents facteurs tels que la syntaxe, la prosodie et la pragmatique. Il est important de noter cependant que ces différents facteurs et leur importance pour une TCU donnée varient selon les participants et la conversation en cours (Ford et Thompson 1996,

cité par Persson, 2014, p.22 ; Eiswirth, 2020b, p. 176). L'importance de la structure syntaxique et de la prosodie dans l'identification d'unités interactionnelles est aussi notée par Degand et Simon (2011) dans leur travail sur la définition d'une « unité discursive de base ».

L'exemple ci-dessous (2) présente un extrait de conversation segmentée en TCU⁴ :

(2)

Exemple 2 Extrait de la conversation politique_ham_15 du corpus TCOF

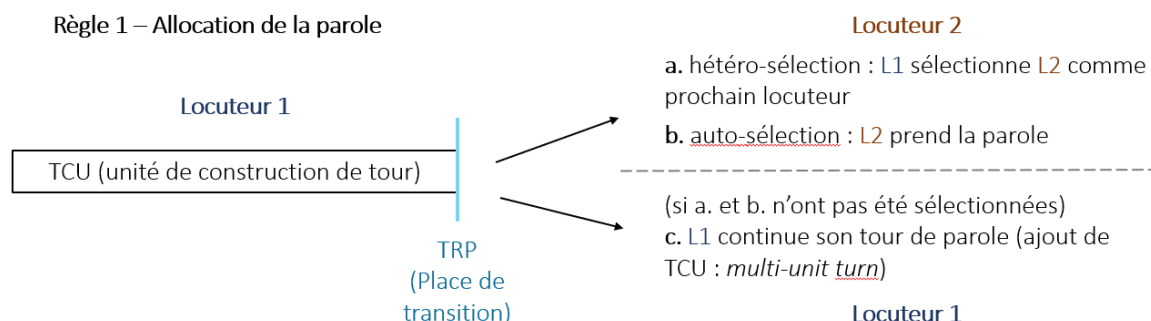
1	-15_TCOF_08_01	je sais pas où j'ai mis mon sac attends
2	-15_TCOF_08_02	tout là-bas
3	-15_TCOF_08_01	tout est dans mon sac ((rires))
4	-15_TCOF_08_02	ouais comme d'ha[bitude]
5	-15_TCOF_08_01	[ma] maison (0.6)
6	-15_TCOF_08_02	ta maison
7	-15_TCOF_08_01	((cherche un flacon de parfum dans son sac)) (4.0) amande douce
8	-15_TCOF_08_02	ah oui je l'avais déjà celui-là il sent trop bon
9	-15_TCOF_08_01	((met du parfum)) ((rires)) (4.0) je l'ai piqué à ma mère
10		parce que le mien il était vide alors euh
11		je me suis dit bon ben je vais prendre celui de ma mère

Les lignes 1 jusqu'à 8 correspondent à des tours de parole (changement de locuteur) constitué d'une seule TCU. En revanche le tour de parole allant des lignes 9 à 11 est constitué de plusieurs TCU. Selon les termes de Sacks et al., on parle dans ce cas de *multi-unit turn*, ou tour à unités multiples ou encore à segments multiples, en Français.

L'aspect systématique et dynamique de la construction et de la répartition des tours de paroles repose sur cette construction en TCU terminées par des TRP. En effet, c'est à chaque TRP que la distribution de la parole est (re-)négociée par les participants : le locuteur qui détenait le tour peut désigner un locuteur pour reprendre la parole, ou bien un interlocuteur peut s'auto-sélectionner pour le faire. Si ces deux premières options n'ont pas été choisies, le locuteur qui détenait la parole peut aussi la garder et continuer son tour de parole (*multi-unit turn*). Ces trois possibilités d'allocation de la parole constituent l'une des

⁴ Segmentation originale du corpus TCOF. Les indications entre double parenthèses, parenthèses simples et les chevauchements (entre crochets) apparaissent sous une forme différente dans la transcription originale.

deux règles de répartition de la parole définies par Sacks et al. dans leur article de 1974. La deuxième règle stipule que ces trois possibilités se renouvellent à chaque TRP, donc à la fin de chaque segment de tour de parole.



Règle 2 : Les trois options de la règle 1 se renouvèlent à chaque TRP (place de transition)

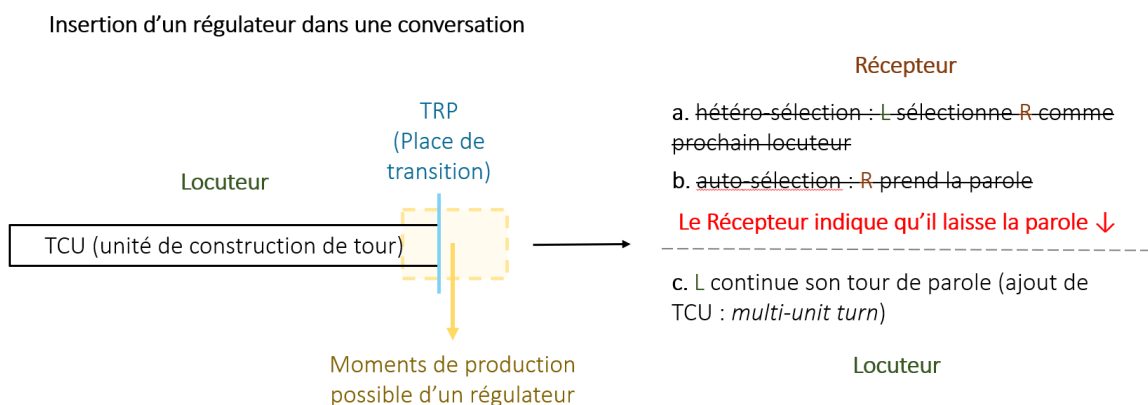
Un tour de parole peut être constitué de une ou plusieurs unités de construction de tour (TCU), et se termine lorsqu'un autre locuteur prend la parole. Les régulateurs ne sont pas considérés comme des tours de parole (ils indiquent au contraire le souhait de laisser le Locuteur continuer son tour)

Figure 2 Schéma récapitulatif des règles d'organisation de la parole dans une conversation (Sacks, Schegloff & Jefferson, 1974)

La prise en compte des dynamiques de répartition de la parole dans une conversation, et l'importance des TRP en particulier, est nécessaire pour étudier les régulateurs. En effet, ceux-ci ne forment pas des tours de parole en eux-mêmes, mais viennent plutôt accompagner le tour de parole du Locuteur⁵ qui détenait jusqu'ici la parole, comme nous allons le voir (de Gaulmyn, 1987, entre autres). Pour ce faire, les régulateurs sont utilisés par le Récepteur là où l'allocation de la parole est négociée, c'est-à-dire aux places de transition possibles (TRP) : c'est précisément en intervenant là où la parole pourrait être redistribuée que les régulateurs confirment au locuteur que la parole lui est laissée (de Nuchèze & Coletta, 2002 ; Traverso, 2016). L'exemple (1) (p. 7) illustre bien ce fait, les deux régulateurs *d'accord* et *hum* (lignes 3 et 5) venant s'insérer après des énoncés (TCU) complets de la locutrice - 10_TCOF_02_01, donc après des TRP. Cette locutrice indique bien, par la suite, qu'elle

⁵ Nous utilisons dans ce mémoire les termes de Locuteur et Récepteur pour désigner les rôles interactionnels des interactants lors de l'émission de régulateurs, ceux-ci étant toujours émis par un Récepteur en lien avec le discours d'un Locuteur détenant la parole (voir chapitre 3. 1).

continue son tour de parole en commençant ses TCU par les marqueurs logiques *et* (ligne 4) et *donc* (ligne 6).



Note : Si le Locuteur ne choisit pas de garder la parole même après la production d'un régulateur par le Récepteur, le Récepteur peut finalement commencer un tour de parole.

Figure 3 Schéma de l'utilisation d'un régulateur dans une conversation

L'importance du placement séquentiel des régulateurs et le rôle que ces marqueurs jouent dans la répartition des tours de paroles sera discutée plus en détail dans le chapitre 3.

2. Avant ceci la différence entre la dimension médiale et la dimension conceptionnelle du langage, ainsi que les caractéristiques de cette dimension conceptionnelle, seront explicitées. En effet, notre discussion s'est jusqu'ici axée sur les interactions en tant que phénomène oral, parlé. Cependant, comme argumentent Koch & Cœsterreicher (2001), les interactions ne s'inscrivent pas uniquement dans ce *medium* qu'est l'oral (une interaction pouvant aussi avoir lieu à l'écrit), mais possèdent également une dimension *conceptionnelle* bien particulière, qui peut être située sur un continuum allant de la proximité à la distance.

2. Koch & Cœsterreicher (2001) et la dimension conceptionnelle du langage

La distinction entre un aspect médial de la communication (oral *versus* écrit) et un aspect conceptionnel (langage parlé, langage écrit) est souvent associée en premier lieu à l'article de Koch & Cœsterreicher (2001), justement intitulé *Gesprochene Sprache und geschriebene Sprache : langage parlé et langage écrit*. Dans cet article, les deux auteurs établissent une différence entre le médium utilisé pour communiquer (l'oral ou bien l'écrit) d'une part, et d'autre part, la conception de la communication s'établissant cette fois-ci sous la forme d'un continuum (le « continuum communicatif »). Cette aspect conceptionnel de la communication correspond à l'« allure linguistique » de l'interaction (Koch & Cœsterreicher, 2001, p. 585), à son style. Le « continuum communicatif » de cet aspect conceptionnel est

délimité par les deux extrêmes que sont le domaine de l'immédiat⁶ et le domaine de la distance, que nous allons détailler par la suite. S'il existe une certaine affinité entre le médium écrit et le domaine de la distance d'une part, et le médium oral et le domaine de l'immédiat d'autre part, cette affinité n'est pas exclusive, et une communication inscrite dans l'immédiat peut très bien se réaliser sous forme écrite, comme en témoignent les SMS entre proches par exemple (Koch & Esterreicher, 2001 ; Mahrer, 2019).

À quoi réfèrent alors précisément les notions d'immédiat/proximité et de distance, et comment les appréhender ? Tout d'abord ces deux notions constituent, on l'a dit, les deux extrêmes d'un continuum communicatif. Une communication peut ainsi relever plus ou moins de la distance ou de l'immédiat selon sa position vis-à-vis de certains paramètres, appelés paramètres situationnels. Ensuite, les notions d'immédiat et de distance sont motivées par des « impératifs communicatifs » distincts, comme le résume très bien Mahrer (2019), reprenant Koch & Esterreicher (2001). Ainsi le domaine de la proximité :

(...) répond aux données anthropologiques suivantes : le besoin d'interagir avec des congénères présents (à la fois ici et maintenant) souvent connus de l'énonciateur et partageant avec lui une même situation, où l'émotion est souvent présente, dans un cadre discursif et un contexte culturel inférables de données partagées. (Mahrer, 2019, p. 2)

Et le domaine de la distance :

(...) répond à la nécessité d'échanger, « la tête froide », avec des interlocuteurs absents (futurs ou ailleurs), souvent inconnus, et suppose donc une autonomie sémantique et une cohérence textuelle supérieures, ayant à comporter lui-même les indications situationnelles et contextuelles nécessaires à l'effectuation de son projet communicatif. (Mahrer 2019 : 2)

On peut ainsi symboliser cette différence avec l'exemple de la conversation ordinaire entre proche (où la proximité est à son maximum), contre l'exemple du texte de loi (ou la distance est à son maximum) (Mahrer, 2019).

⁶ Le terme de proximité est préféré par certains auteurs, dont Gadet (2017), que nous suivrons sur ce point.

Koch & Oesterreicher détaillent de plus 10 paramètres situationnels permettant de « caractériser le comportement communicatif des interlocuteurs par rapport aux déterminants situationnels et contextuels » (ibid., p. 586) :

Tableau 1 Les paramètres situationnels de Koch & Oesterreicher (2001)

1) communication privée	communication publique (1)
2) interlocuteur intime	interlocuteur inconnu (2)
3) émotionnalité forte	émotionnalité faible (3)
4) ancrage actionnel et situationnel	détachement actionnel et situationnel (4)
5) ancrage référentiel dans la situation	détachement référentiel de la situation (5)
6) coprésence spatio-temporelle	séparation spatio-temporelle (6)
7) coopération communicative intense	coopération communicative minimale (7)
8) dialogue	monologue (8)
9) communication spontanée	communication préparée (9)
10) liberté thématique	fixation thématique (10)
Etc.	Etc.

Comme les deux auteurs l'indiquent, le caractère scalaire du continuum communicatif proximité / distance est donc à la fois lié à la gradation propre à chacun des paramètres ainsi qu'à la combinaison de leurs valeurs (Koch & Oesterreicher, 2001). La place d'une situation de communication donnée sur le continuum communicatif doit donc être calculée pour chaque paramètre, puis en faisant la somme des valeurs obtenues par paramètres. Il est de plus important de noter que la valeur assignée à chaque paramètre peut évoluer au cours d'une situation de communication donnée, au fur et à mesure des thèmes abordés, des séquences communicatives et du rapprochement des participants (Guerin, 2017).

Malgré ce caractère variable y compris au cours d'une même conversation, ces paramètres situationnels restent un très bon moyen d'évaluer l'ancrage d'une situation de communication donnée dans le domaine de la proximité ou de la distance. À ce titre, ils seront mobilisés lors de l'analyse de notre corpus afin d'établir la valeur de proximité/distance par locuteur dans chaque interaction. Ce point sera discuté plus en détail dans le chapitre 7. 1.

Il est à noter cependant que le corpus utilisé pour ce mémoire étant uniquement constitué de conversations orales privées, le degré de proximité/distance varie peu. De plus,

toutes les conversations sont orientées vers la proximité (même si certains écarts peuvent être observés, comme nous le verrons par la suite).

La revue de la littérature au sujet de la proximité nous fait émettre l'hypothèse que les régulateurs, que nous étudions dans ce mémoire, font partie des éléments linguistiques typiquement liés à la proximité. En effet, Koch & Œsterreicher mentionnent les « signaux de l'auditeur qui accompagnent régulièrement la performance du locuteur, sans, pour autant, introduire un tour de parole »⁷ (2001, p. 594) comme des particules liées à la proximité. Le fait que les régulateurs permettent entre autres d'assurer le lien entre les interlocuteurs, participant ainsi la co-construction du discours, vient renforcer cette hypothèse d'un ancrage de ces comportements linguistiques dans le domaine de la proximité communicative.

Il semble donc possible de considérer la présence de régulateurs dans une communication comme une marque de proximité et pas uniquement d'oralité. Ce fait peut être mis en lien avec une réflexion de Gadet à propos de phénomènes tels que les appellatifs (comme *mec, meuf, (mon) gars*, etc.)⁸ et certains types d'interrogatives :

(...) Ces phénomènes, reposant cruciallement sur l'interaction et l'ajustement entre les interactants, nous semblent tout aussi bien caractérisés par la proximité que par l'oralité : c'est du fait d'un flou autour de la conception de l'oralité que l'on a pu les confondre ou réduire un ordre à l'autre. (Gadet, 2017, p. 124)

Le présent mémoire portant uniquement sur les conversations privées *orales*, nous n'irons pas plus loin dans la discussion de la primauté de l'oral (média) ou bien de la proximité (conception) dans l'usage des régulateurs, et les considérerons comme des marques à la fois d'oralité et de proximité.

Cette hypothèse d'un lien entre la proximité et la présence de régulateurs sera toutefois testée et discutée dans notre analyse (chapitre 9. 2), avec un examen de l'influence possible du taux de proximité (général et pour certains paramètres) sur la fréquence et le type de régulateurs observés.

⁷ C'est-à-dire nos régulateurs

⁸ Exemples de Gadet (2017, p. 123), qui se base sur la définition donnée par Cappeau & Schnedecker (2017).

Pour résumer, il existe un ensemble de règles dans les conversations qui permet la répartition des tours de paroles, et qui repose sur des unités de constructions de tours (TCU) dont les places possibles de transition (TRP) peuvent être anticipées par les participants. La connaissance de ces règles et principes est donc très importante dans le cadre d'un travail sur des unités interactionnelles telles que les régulateurs, tout comme la connaissance et la prise en compte de la différence entre l'aspect médial et conceptionnel du langage. Ces deux notions sont nécessaires pour expliquer et comprendre les régulateurs dans les conversations orales. En effet, les régulateurs participent fortement à l'organisation et à la co-construction de l'interaction, et nous avons aussi pu voir qu'ils paraissent relever à la fois de l'oralité et de la proximité.

Ces deux notions étant fixées, il conviendra désormais de s'intéresser au vif du sujet : les régulateurs verbaux de connivence. Afin de les appréhender au mieux, il s'agira dans un premier de définir les régulateurs de manière globale, en insistant dans un deuxième temps sur leur importance dans l'interaction. Ces informations permettront d'établir ensuite les raisons pour lesquelles l'objet d'étude de ce mémoire s'est porté sur les *régulateurs verbaux de connivence*, et d'expliquer et définir les particularités de ce sous-ensemble de régulateurs.

Chapitre 3. Approche et définition des *régulateurs verbaux de connivence*.

Les régulateurs forment un objet d'étude particulièrement complexe, du fait de la quantité d'études dédiées à ce sujet d'une part, mais aussi du fait de leur importance et de leur ancrage dans une situation d'énonciation donnée d'autre part. Ce mémoire portant uniquement sur un sous-ensemble de régulateurs, il sera donc primordial de définir tout d'abord la notion générale de *backchannel* et le fonctionnement des régulateurs, avant, dans un deuxième temps, d'insister sur la manière dont ils s'ancrent, tant d'un point de vue interactionnel que pragmatique, dans les conversations. Enfin, ces informations assimilées, nous pourrions définir notre objet d'étude précis pour ce mémoire : les *régulateurs verbaux de connivence*.

1. La notion générale de *backchannel* : une multitude d'études, mais peu de consensus

Quel que soit le nom donné à ce phénomène, ce qu'on pourrait résumer comme les comportements (verbaux et non verbaux) d'écoute sans reprise de parole d'un locuteur dans une interaction ont fait couler beaucoup d'encre, et sont étudiés depuis plus de quatre décennies dans différents courants de la linguistique (Eiswirth 2020b). En effet, le travail à l'origine du terme *backchannel* et marquant les débuts de la recherche sur ce phénomène (cité par de nombreux auteurs, dont Schegloff, 1982 ; Goodwin, 1986 ; Maynard, 1990 ; O'Keeffe & Adolphs, 2008 ; Bertrand et al., 2009, ou encore Eiswirth, 2020b) a été publié par Victor Yngve en 1970.

Outre le terme populaire de *backchannel*, on retrouve aussi les termes *listener responses* (Eiswirth, 2020a, 2020b), *response tokens* (O'Keeffe & Adolphs, 2008), *minimal responses* (Fellegy, 1995, cité par O'Keeffe & Adolphs, 2008 et Eiswirth, 2020b), ou encore *régulateurs* en français (de Gaulmyn, 1987 ; Persson, 2014). À la suite de ces derniers, nous utiliserons dans ce mémoire ce terme de « régulateurs », tout d'abord parce qu'il présente l'intérêt d'être français et d'être utilisé dans certaines études francophones sur ce phénomène (de Gaulmyn, 1987 ; Cosnier, 1988 cité par Bertrand et al., 2009), mais aussi parce qu'il nous semble plus adéquat pour décrire le phénomène étudié et le comportement d'un

récepteur dans une conversation. Nous qualifierons de plus le locuteur produisant un régulateur et manifestant ce comportement d'écoute de « Récepteur », et son interlocuteur détenant la parole de « Locuteur » afin de suivre la terminologie adoptée par de Gaulmyn (1987), encore une fois, mais aussi afin d'éviter les confusions.

La multitude de termes employés pour ce phénomène reflète la variabilité des définitions et la dispersion des études au sujet des régulateurs. En effet, si ceux-ci ont d'abord été étudiés principalement dans le domaine de l'Analyse Conversationnelle (*Conversational Analysis* en anglais, voir entre autres les travaux de Schegloff, 1982 ou bien Goodwin, 1986), ils ont ensuite intéressé différentes communautés, avec des travaux portant sur la prosodie par exemple, ou bien sur les différences interculturelles⁹ (Bertrand et al., 2009, p. 3-4).

Il existe cependant un consensus quant à la définition minimale de ces régulateurs, qu'on pourrait résumer de la manière suivante : il s'agit des comportements (verbaux, vocaux ou gestuels) utilisés par un Récepteur dans une interaction afin de manifester son écoute et son engagement sans reprendre la parole au Locuteur. Comme nous le disions, en revanche, il y a beaucoup moins de consensus pour ce qui est des détails et de la portée de cette définition, celle-ci étant très variable selon les auteurs. Ainsi, comme le pointe Eiswirth (2020a), certains auteurs basent leurs définitions sur la forme des régulateurs (la quantité et qualité des formes retenues selon les auteurs étant variable¹⁰), d'autres sur la position séquentielle de ceux-ci dans l'interaction, ou encore sur un mélange de la forme et de la position des régulateurs. Nous allons de notre côté nous intéresser particulièrement à la position séquentielle et à la fonction des régulateurs dans notre définition. En effet, les régulateurs sont des éléments purement interactionnels, et ne peuvent donc être conçus et définis qu'en prenant en compte cet ancrage dans l'interaction.

⁹ Voir par exemple le travail de Maynard, 1990, sur la différence dans l'utilisation des *backchannels* dans des conversations entre des Japonais et des Américains (cadre de l'Analyse Contrastive).

¹⁰ Ainsi certains auteurs ne s'intéressent qu'à certains sous-ensembles formels de régulateurs, et d'autres se basent sur la longueur des segments pour les considérer ou non comme régulateurs (avec des longueurs variables selon les auteurs) (Eiswirth, 2020b : 174).

En nous basant sur les travaux de de Gaulmyn (1987), Bertrand et al. (2009) et Eiswirth (2020a) en particulier, nous définirons les régulateurs de la manière suivante :

1. D'un point de vue formel, les régulateurs sont de brefs éléments de nature verbale (c'est-à-dire possédant un contenu lexical, i.e. *d'accord, c'est sûr, tu m'étonnes*, etc. ; le terme de verbal renvoie ici au canal de communication), vocale (i.e. *hum, ah, oh, ouais*¹¹, etc.) et/ou co-verbaux (i.e. hochements de tête, regards, etc.). Ils sont souvent réalisés en multimodalité, et peuvent soit s'ajouter, soit être utilisés seuls. Les régulateurs verbaux sont généralement constitués d'une à quatre ou cinq unités lexicales. Tous sont indépendants syntaxiquement de l'énoncé du Locuteur.
2. D'un point de vue de la structure interactionnelle, ils sont produits *pendant* le tour de parole d'un Locuteur, ou bien à sa fin si le Locuteur décide de ne pas reprendre la parole. Ils sont réalisés après ou en superposition avec des places de transition possibles (TRP)¹², MAIS ne *doivent* pas forcément être réalisés à chaque TRP : ils sont donc optionnels. Ils n'ont pas un statut de tour de parole, comme nous allons le préciser, et viennent soit se superposer soit s'insérer avec le tour de parole du locuteur.

Eiswirth (2020b, p. 175) schématise leur ancrage structurel dans l'interaction de la manière suivante :

1 *Speaker: ongoing multi-unit-turn*

2 *Listener: Listener Response (= régulateur)*

3 *Speaker: continuation of ongoing multi-unit-turn*

Nous ajouterons ici, comme dit précédemment, que le troisième tour est optionnel lorsque le locuteur décide de ne pas continuer sur un autre thème malgré la production d'un régulateur par le récepteur (voir figure 3, p. 18).

¹¹ Les formes *oui* et *ouais* sont considérées comme des régulateurs vocaux dans ce mémoire, du fait de leur forte désémantisation.

¹² Cette superposition possible est à mettre en lien avec le phénomène de projection des TRP, qui peuvent être anticipées par les locuteurs (voir chapitre 2. 1 et la figure 3 pour plus de détails).

3. D'un point de vue pragmatique, leur fonction majeure est de permettre au récepteur de manifester son écoute sans interrompre le flux de parole du locuteur, et d'encourager celui-ci à continuer de parler. Ils marquent ainsi la connivence et viennent accompagner le locuteur dans son discours.

Si nous considérons les régulateurs comme des non-tours de parole dans notre définition, ce statut reste assez débattu dans la littérature, et dépend fortement de la définition d'un tour de parole adoptée. De plus, si la plupart des auteurs s'accorde à dire que les régulateurs ne sont *pas* des tours de parole, le statut précis à leur accorder au-delà de cette affirmation reste flou (de Gaulmyn, 1987 ; de Nuchèze & Coletta, 2002). Comme nous l'avons déjà dit au chapitre 2. 1, nous considérons dans ce mémoire qu'un tour de parole est une prise de parole par un locuteur engendrant un changement de locuteur. Cette définition est compatible avec le statut de non-tour des régulateurs. En effet, même si la production d'un régulateur correspond bien à une émission de parole, cette production accompagne et se superpose à la parole du Locuteur plutôt que de la reprendre : il n'y a donc pas réellement de changement de locuteur (voir figure 3).

Si tous les régulateurs ont la même macro-fonction pragmatique (point 3 de la définition), ils peuvent aussi remplir différentes micro-fonctions spécifiques. Celles-ci sont cependant beaucoup moins représentées dans la littérature (O'Keeffe & Adolphs, 2008 ; Persson, 2014), et encore une fois la dispersion de la littérature existante rend difficile une vue d'ensemble de ces particularités. Plusieurs auteurs ont ainsi pu analyser des fonctions spécifiques individuellement (Schegloff, 1982 ; Goodwin, 1986), ou bien ont tenté de recenser les différentes micro-fonctions principales des régulateurs (O'Keeffe & Adolphs, 2008 ; Bertrand et al., 2009 ; Eiswirth, 2020a), en se basant sur l'organisation séquentielle des régulateurs, sur leurs formes, sur leurs effets pragmatiques ou bien sur un ensemble d'indices (pragmatiques, séquentiels, prosodiques, etc.). Certaines de ces classifications sont résumées dans le tableau 2 ci-dessous.

Si la variation terminologique est encore une fois très importante dans ces différentes classifications, certains éléments sont référencés de manière récurrente, et il est ainsi possible d'extraire quelques généralités quant aux micro-fonctions des régulateurs.

Tableau 2 Résumé de quelques classifications fonctionnelles des régulateurs trouvées dans la littérature

Goodwin, 1986	de Gaulmyn, 1987	Maynard, 1989 (cité par O’Keeffe & Adolphs, 2008 et Bertrand et al. 2009)
<ul style="list-style-type: none"> • <i>Continuers</i> (Schegloff 1982): “treat [the current] unit as preliminary to further talk” (1986 : 210) • <i>Assessments</i>: “operate on the particulars of a current unit” (1986 : 210) 	<ul style="list-style-type: none"> • « <i>Le récepteur enregistre l’énonciation et/ou l’énoncé du locuteur (...)</i> » • « <i>Le récepteur évalue positivement l’énoncé du locuteur (...)</i> » • « <i>Le récepteur répète en écho un segment de l’énoncé du locuteur</i> » • « <i>Le récepteur collabore à l’énoncé en cours du locuteur (...)</i> » <p style="text-align: center;">(1987 : 207)</p>	<ul style="list-style-type: none"> • <i>Display of understanding of content</i> • <i>Support towards the speaker’s judgment</i> • <i>Agreement</i> • <i>Strong emotional response</i> • <i>Minor addition, correction or request for information</i>
O’Keeffe & Adolphs, 2008	Bertrand et al., 2009	Eiswirth, 2020a
<ul style="list-style-type: none"> • <i>Continuer tokens</i> (after Schegloff, 1982): “maintain the flow of talk” • <i>Convergence tokens</i>: “markers of agreement/convergence” • <i>Engagement tokens</i>: “markers of high engagement where addressee(s) respond on an affective level to the content of the message” • (minor function) <i>Information receipt tokens</i>: “markers of points in the discourse where adequate information has been received” <p style="text-align: center;">(2008 : 16-17)</p>	<ul style="list-style-type: none"> • <i>Continuer</i> (prendre note minimalement) • <i>Understanding</i> (j’ai bien compris mais sans degré d’adhésion, degré supérieur au <i>continuer</i>) • <i>Acknowledgment</i> (support, adhésion à un propos) • <i>Assessment</i> (évaluation, -rire par exemple-, jugement, déclaration d’attitude) • <i>Request / Confirmation request</i> • <i>Complex</i> <p style="text-align: center;">(2009 : 187)</p>	<ul style="list-style-type: none"> • <i>Acknowledgment</i>: “both backward- and forward-looking in that they acknowledge the preceding talk and orient to the relevance of the current Speaker continuing without making any particular kind of continuation relevant” (2020a: 120) • <i>Surprise mark</i>: “show [the listener’s] appreciation of the content of a previous TCU as new, surprising, or otherwise unexpected” (2020a: 122) • <i>First Assessment</i>: “treat some aspect of the preceding talk as assessable” (2020a: 126) • <i>Second Assessment</i>: occur after “the Speaker makes a first assessment of some aspect of their own talk” (2020a: 129) • <i>Self-initiated other-repair</i> • <i>Other-initiation</i> • <i>Voicing</i> • <i>Collaborative completion</i>

Ainsi la mention des fonctions de *continuer* et d'*assessment*¹³ (Goodwin, 1986, entre autres) revient régulièrement dans la littérature. Ces deux fonctions correspondent à deux utilisations différentes des régulateurs sur le plan séquentiel.

En tant que *continuer* (ou *continueur* en français), le régulateur fonctionne comme un pont entre les segments du tour de parole du Locuteur, et permet au Récepteur d'indiquer qu'il est bien conscient que le Locuteur n'a pas fini son tour de parole, et qu'il attend la suite. Il s'agit donc, comme son nom l'indique, d'une fonction de *continuation*. Cette fonction a été identifiée pour la première fois par Schegloff (1982). En tant qu'*assessment* (ou *évaluation* en français), en revanche, le régulateur porte sur le(s) segments précédents ou un aspect de ceux-ci. Il vient donc ratifier (avec ou sans degré d'adhésion et/ou d'affect) l'énonciation/l'énoncé en cours. Les *assessments*, s'ils encouragent toujours le locuteur à continuer de parler, ne portent donc pas sur le discours du locuteur comme un tout en cours de construction mais bien sur des aspects spécifiques de ce discours (ils ont pour fonction d'évaluer le(s) segment(s) précédents). Les exemples (3) et (4), p. 44 (chapitre 4. 2), présentent ces deux cas de figure.

Il apparaît ensuite que plusieurs types de régulateurs puissent être dégagés au niveau de la forme. À la suite de Laforest (1992, citée par Bertrand et al., 2009), nous retiendrons la distinction entre les régulateurs simples et les régulateurs complexes :

[Laforest] oppose en outre les régulateurs simples, c'est-à-dire qui ne constituent pas un réel tour de parole, aux régulateurs complexes, qui renvoient aux divers cas de reformulation, complétion, répétition et metaquestion dont le statut de non tour s'avère plus délicat à établir. (Laforest 1992, citée par Bertrand et al., 2009, p. 5).

La différence entre ces deux types de régulateurs sera discutée plus en détail dans la partie 3 de ce chapitre, puisque seuls les régulateurs verbaux *simples* seront étudiés dans ce mémoire. La façon dont ces typologies fonctionnelles préexistantes seront intégrées dans notre propre typologie sera quant à elle détaillée dans le chapitre 4. 1.

¹³ Ces deux fonctions ne seront pas étudiées en détail dans notre analyse, mais elles seront en revanche reprises dans notre typologie à titre indicatif, et restent dans tous les cas deux fonctions, ou deux organisations séquentielles des régulateurs, qu'il est nécessaire de bien comprendre.

Pour finir, il semble qu'il existe une distinction entre la portée du régulateur sur l'acte d'énonciation d'un côté et l'énoncé lui-même du locuteur de l'autre, distinction mentionnée dans le travail de de Gaulmyn :

Les régulateurs verbaux sont orientés positivement et expriment une convergence avec l'acte d'énonciation et, éventuellement, avec l'énoncé du locuteur principal : ce sont des facteurs de consensus. (de Gaulmyn, 1987 : 205)

Pour résumer, le terme de régulateurs, ou plus communément *backchannel* dans la littérature internationale, est un terme générique désignant un ensemble de comportements verbaux, vocaux ou bien co-verbaux, pouvant être réalisés seuls ou combinés, qui permettent à un locuteur se positionnant comme Récepteur de faire part de son écoute vis-à-vis du discours d'un Locuteur détenant la parole. Ces régulateurs peuvent être définis selon un ensemble de critères formels, mais aussi interactionnels (leur placement séquentiel dans l'interaction) et pragmatiques (leurs différentes fonctions dans le discours, c'est-à-dire les actions qu'ils permettent de réaliser). Ce sont des marqueurs interactionnels, qui sont toujours orientés vers le lien entre le Récepteur et le Locuteur, et permettent de marquer la connivence et le consensus.

Les régulateurs peuvent avoir différentes micro-fonctions, qui semblent pouvoir se combiner de façon assez complexe, et sur lesquelles nous reviendrons lors de la description de notre typologie dans le chapitre 4. Dans un premier temps, il conviendra cependant de donner plus de détails sur les différents éléments qui font des régulateurs des marqueurs-clés de l'interaction. En effet, ces marqueurs sont très souvent mentionnés, y compris dans des travaux dont ce n'est pas le propos principal, lorsqu'il s'agit de décrire le système d'alternance des tours de paroles ou de manière plus générale la co-construction de l'interaction (comme c'est le cas dans les ouvrages de de Nuchèze & Coletta, 2001 et de Traverso, 2016, par exemple). Les régulateurs semblent ainsi être désignés comme des éléments incontournables de la description des interactions.

2. De l'importance des régulateurs dans l'interaction

Comme nous l'avons précisé plus haut, les régulateurs ne peuvent pas être définis sans prendre en compte leur ancrage dans l'interaction. Leur brièveté, leur forme, et même leur fonction pragmatique d'accompagnement ne sont pas des critères suffisants pour les définir (de Gaulmyn, 1987). Eiswirth (2020b) insiste aussi sur ce point, en précisant que les

régulateurs, puisqu'ils sont ancrés dans l'interaction, doivent être définis en prenant ce critère en compte.

Nous l'avons fait dans notre définition (partie 1 de ce chapitre) en mentionnant leur position vis-à-vis du tour de parole du locuteur et leur statut de non-tour de parole. Nous allons maintenant aller plus loin en détaillant les mécaniques de ce rôle interactionnel fort des régulateurs, puis leurs fonctions pragmatiques/interactionnelles qui en découlent, afin d'illustrer leur rôle clé dans l'interaction.

On peut distinguer deux aspects principaux du rôle interactionnel des régulateurs, qui expliquent tous les deux leur importance dans une interaction.

Du fait de leur placement séquentiel vis-à-vis du tour de parole du locuteur, tout d'abord, les régulateurs viennent ponctuer les propos de celui-ci, contribuant ainsi à la structuration de son discours :

Les régulateurs soulignent, ponctuent les articulations syntaxiques et les mouvements argumentatifs du locuteur. Les régulateurs contribuent à l'amplification du discours et signalent la structuration progressive de ce discours, le travail d'élaboration du locuteur, incité ou stimulé par le récepteur.
(de Gaulmyn, 1987, p. 209-210)

Ce fait est aussi souligné par Bertrand et al., qui précisent de plus que les régulateurs donnent ainsi des indications à la fois sur le « processus de production des discours des locuteurs », mais aussi sur le « processus d'écoute des interlocuteurs » (2009, p. 4). De Gaulmyn rejoint encore une fois cette idée, en émettant l'hypothèse que les régulateurs constituent des « trace(s) des opérations de compréhension et de stockage en mémoire chez l'auditeur » (1987, p. 210). Cette hypothèse expliquerait leur présence derrière ou autour de pauses logiques, ménagées par le locuteur, c'est-à-dire des moments où l'énoncé du locuteur est complet¹⁴.

Ensuite, les régulateurs jouent aussi un rôle très important de réaffirmation des rôles dans l'interaction, qui explique aussi en partie leur placement séquentiel. En effet, comme

¹⁴ Les exemples (4), (5) et (6), p. 44, 45 et 47, permettent en particulier d'observer ce phénomène, étant de plus situés en fin de thème.

nous l'avons vu dans le chapitre 2. 1, au sujet des règles de répartition de la parole dans les conversations, c'est précisément derrière les TRP (ou places de transition possibles), que la parole peut possiblement être redistribuée. Or, les régulateurs, qui apparaissent effectivement autour de TRP, ne constituent pas un tour de parole, et ne correspondent donc pas à une tentative de reprise de parole. Ils permettent de ce fait au Récepteur de réaffirmer son rôle de Récepteur, et de confirmer au Locuteur qu'il peut conserver la parole et donc son rôle de locuteur principal. Ce fait est détaillé dans le travail de Eiswirth (2020a), entre autres, qui explicite de plus les différentes manières dont les régulateurs jouent sur la distribution des rôles de récepteur et de locuteur :

(...) [Listener Response] orient to the current speaker Speaker-Listener role distribution in three different ways : (a) They do not claim the floor (...) (b) They can be produced after a short lapse, and usually, the Speaker then picks up the thread (c) If there is a lapse following the LR, the Listener generally waits until the Speaker continues rather than taking over the floor. If they have waited and signalled in other ways that they expect the Speaker to continue, the Listener may eventually take the floor. (Eiswirth, 2020a, p. 83-84)

Pour résumer, les régulateurs ont deux fonctions principales d'un point de vue interactionnel : ils permettent tous d'abord de ponctuer et d'accompagner la structuration du discours du locuteur, en se plaçant derrière des points possibles de transition et donc des unités de construction de tours complètes (d'un point de vue syntaxique, pragmatique et prosodique). Ce positionnement séquentiel leur permet de réaffirmer les rôles de récepteur et de locuteur, puisque les règles de la répartition des tours de paroles impliquent que ces rôles peuvent être renégociés à chaque TRP. En émettant un régulateur derrière une TRP, le Récepteur appuie donc le discours du Locuteur, lui signale le bon enregistrement de ce discours, et lui réaffirme son souhait de rester Récepteur uniquement, l'encourageant de cette manière à continuer de parler.

En plus de la fonction pragmatique d'accompagnement et d'encouragement à garder la parole déjà mentionnée dans la définition des régulateurs, ces fonctions interactionnelles des régulateurs peuvent être mises en lien avec d'autres fonctions pragmatiques/interactionnelles complémentaires : la co-construction du discours, et la confirmation tacite de bonne réception.

Ainsi si les régulateurs dépendent nécessairement des segments qui les précèdent, comme nous avons pu le voir, du point de vue de leur organisation séquentielle, ils y sont

aussi liés d'un point de vue pragmatique/interactionnel. En effet, comme le souligne Persson (2014), les régulateurs sont à la fois orientés vers le tour précédent, mais aussi vers le tour suivant : le type de régulateur réalisé par un Récepteur dépend de son interprétation du tour précédent, et façonne aussi le tour subséquent.

Les régulateurs ne fonctionnent ainsi donc pas comme de simples marques de réceptions vides de sens, mais permettent au contraire au Récepteur d'agir, même minimalement, sur la conversation, de rester acteur de celle-ci. Ils lui permettent de continuer à *co-construire* l'interaction (de Gaulmyn, 1987 ; Traverso, 2016 ; Eiswirth, 2020b entre autres). C'est d'ailleurs en partie pour cette raison que le terme « régulateurs » a été préféré dans ce mémoire à celui de *backchannel*, pourtant majoritaire dans la littérature sur le sujet. En effet, la notion de *backchannel* ne met pas suffisamment l'accent sur le rôle actif du récepteur (Eiswirth, 2020b). Le terme de *régulateur* permet en revanche de laisser sous-entendre ce rôle (l'acte de régulation étant bel et bien effectué par le Récepteur). De Gaulmyn indique en ce sens que l'activité du Récepteur, par le biais des régulateurs, est une activité de « contrôle » lui permettant entre autres « d'orienter l'énoncé au fur et à mesure qu'il écoute » (1987, p. 213)¹⁵.

Une deuxième fonction pragmatique/interactionnelle des régulateurs à mettre en lien, encore une fois, avec leur organisation séquentielle, est une fonction de confirmation tacite de bonne réception et d'absence de problème dans le discours du Locuteur. En effet, comme il a été répété plusieurs fois, les régulateurs sont majoritairement produits derrière des TRP (places de transition possibles). Ces TRP sont aussi les moments où les actes de réparation¹⁶ peuvent avoir lieu (Sacks, Schegloff & Jefferson, 1974). De ce fait, comme le pointe Schegloff (1982), en émettant un régulateur, le Récepteur laisse passer l'opportunité non seulement de reprendre la parole, mais aussi de produire un acte de réparation. Les régulateurs ne sont donc pas des signaux de compréhension en eux-mêmes, mais peuvent

¹⁵ Il semble cependant que l'utilisation de régulateurs n'est pas toujours consciente, et peut être une utilisation de routine, comme nous le verrons dans notre analyse (chapitre 9, point 3. 3 en particulier).

¹⁶ Les réparations étant des actes soit *auto-initiés* (hésitations, reprises, etc. par le locuteur détenant la parole) soit *hétéro-initiés* (demandes de clarification, collaborations avec proposition de mots/phrases, etc. par l'interlocuteur) qui permettent de régler une perturbation dans l'interaction (Traverso, 2016).

bien être vus comme confirmant la compréhension et de manière générale l'absence de problème dans la réception du discours du Locuteur, précisément parce qu'ils laissent passer l'opportunité de réparer de possibles problèmes (Schegloff, 1982, p. 88).

Pour conclure, les régulateurs permettent donc, d'un point de vue pragmatique/interactionnel, non seulement de marquer la connivence, de manifester son écoute et d'encourager le locuteur à continuer sur sa lancée, mais aussi de co-construire l'interaction et de confirmer implicitement l'absence de problème à la fois lors de l'émission du discours par le Locuteur et lors de sa réception. Ces différentes fonctions sont fortement liées à leur organisation et rôle interactionnel, et viennent donc confirmer le fait qu'il est impossible d'étudier les régulateurs sans prendre cet ancrage interactionnel en compte.

Toutes ces informations nous permettent ainsi de conclure, et ce, malgré le caractère optionnel des régulateurs, qu'ils constituent des éléments incontournables des interactions en règle générale.

Afin de pouvoir rendre compte au mieux de ces différentes fonctions et rôles des régulateurs ainsi que de leur variation, une seule catégorie de régulateurs sera étudiée dans ce mémoire, celle des *régulateurs verbaux de connivence*. En effet, cette catégorie présente plusieurs avantages qui justifient son éléction pour cette étude exploratoire visant à examiner les micro-fonctions des régulateurs et leur variation.

3. *Les régulateurs verbaux de connivence*

Comme nous avons pu le constater dans la partie 1 de ce chapitre, la littérature existante ne permet pas de dégager une typologie fonctionnelle unifiée des régulateurs. De plus, comme le souligne Persson « [d]ans le domaine de la linguistique française, il manque largement des études systématiques de régulateurs particuliers » (2014, p. 115). Le choix a donc été fait, pour ce mémoire, de travailler sur une catégorie précise de régulateurs, afin de pouvoir en apporter une description plus fine en s'intéressant notamment à leurs micro-fonctions. Le nom choisi pour désigner cette catégorie de régulateurs est celui de *régulateurs verbaux de connivence*.

Les régulateurs verbaux correspondent aux formes lexicalisées des régulateurs, c'est-à-dire aux formes constituées de mot(s)/unité(s) identifiables de la langue donnée (de Gaulmyn, 1987). Dans le travail de O'Keeffe & Adolphs, ils correspondent globalement aux *non-minimal responses*, qui sont décrites comme étant souvent des « adverbes ou adjectifs

ayant fonction de marqueur pragmatique (e.g. *good, really, great, absolutely*) ou bien de courts syntagmes ou propositions minimales (e.g. *you're not serious, 's that so?, by all means, fair enough, that's true, not all*) ». À l'inverse, les *minimal responses* sont définies comme de « courts énoncés ou vocalisations non verbales » (O'Keeffe & Adolphs, 2008, p. 6, traduction libre).

Cette catégorie de régulateurs est rarement étudiée seule de façon systématique. Seules deux études parmi celles référencées dans ce mémoire se sont concentrées sur les régulateurs verbaux : celle de de Gaulmyn en 1987 (sur le français) et celle de O'Keeffe & Adolphs en 2008 (sur l'anglais Britannique et Irlandais).

Du fait de leur contenu lexical, les variations formelles de ces régulateurs sont distinctives et aisées à établir, puisqu'elles correspondent, comme nous avons pu le voir, à des formes lexicales identifiables de la langue donnée. L'identification d'un lien possible entre leurs formes et leurs fonctions en est ainsi facilitée (même s'il faut rappeler, comme le font Bertrand et al. (2009) et Eiswirth (2020a), que les régulateurs sont fortement polyfonctionnels et qu'une même forme, y compris lexicale, peut réaliser différentes fonctions selon le contexte). Ce contexte facilite ainsi l'étude de ces régulateurs particuliers pour une étude exploratoire visant à examiner leurs différentes micro-fonctions et leur variation.

Il existe néanmoins beaucoup de régulateurs verbaux différents, avec des fonctions très variables. Pour un travail de mémoire, il aurait été trop coûteux en temps d'analyser toutes ces fonctions pragmatiques possibles de régulateurs verbaux. Le choix a donc été fait de se concentrer sur un sous-ensemble de ces régulateurs, dont la fonction pragmatique correspond au rôle minimal et emblématique des régulateurs, tel que nous l'avons vu dans la définition : permettre au récepteur de manifester son écoute sans interrompre le flux de parole du locuteur ; encourager le locuteur à continuer de parler ; marquer la connivence et accompagner le locuteur dans son discours.

Afin de rappeler ce rôle canonique, nous avons décidé, dans le cadre de ce mémoire, de nommer ce sous-ensemble de régulateurs les *régulateurs verbaux de connivence* (désormais RVC).

S'il est bien évident que *tous* les régulateurs relèvent de la connivence, quoique selon des degrés variables, cette dénomination permet selon nous d'accentuer le fait qu'il s'agit du sous-ensemble de régulateurs verbaux dont le rôle est consigné à l'expression de cette

connivence : les RVC se limitent à l'accompagnement du discours du locuteur, sans se placer en rupture vis-à-vis de celui-ci. Reprenons une partie du tableau des classifications existantes afin d'explicitier cette idée :

Tableau 3 Classifications fonctionnelles des régulateurs par Maynard (1989) et Bertrand et al. (2009)

	Maynard, 1989 (cité par O’Keeffe & Adolphs, 2008 et Bertrand et al. 2009)	Bertrand et al., 2009
Différentes sous-fonction permettant l’accomplissement de la fonction pragmatique prototypique des régulateurs (voir point 3. de la définition dans la partie 1 de ce chapitre). Pas d’ajout ni de question de la part du récepteur	<ul style="list-style-type: none"> • <i>Display of understanding of content</i> • <i>Support towards the speaker’s judgment</i> • <i>Agreement</i> • <i>Strong emotional response</i> 	<ul style="list-style-type: none"> • <i>Continuer</i> (prendre note minimalement) • <i>Understanding</i> (j’ai bien compris, mais sans degré d’adhésion, degré supérieur au <i>continuer</i>) • <i>Acknowledgment</i> (support, adhésion à un propos) • <i>Assessment</i> (évaluation, -rire par exemple-, jugement, déclaration d’attitude) <p>(2009, p. 5)</p>
Divers cas permettant au récepteur d’aller plus loin que l’expression de la simple connivence avec le discours du locuteur.	<ul style="list-style-type: none"> • <i>Minor addition, correction or request for information</i> 	<ul style="list-style-type: none"> • <i>Request/Confirmation request</i> • <i>Complex</i> <p>(2009, p. 5)</p>

Les classifications fonctionnelles des deux auteurs ci-dessus détaillent plusieurs micro-fonctions des régulateurs, et se terminent dans les deux cas par la description de catégories plus spécifiques (d’ailleurs significativement placées en fin de classification) : les divers cas de collaborations entre le Récepteur et le Locuteur (ajout et correction chez Maynard, 1989 (cité par O’Keeffe & Adolphs, 2008 et Bertrand et al., 2009) et les cas complexes chez Bertrand et al., 2009¹⁷) et les requêtes/demandes d’informations.

Ces divers cas de collaborations et de demandes d’informations semblent de fait aller plus loin que l’expression minimale de la connivence, puisque le Récepteur ne va pas simplement dans le sens du Locuteur, mais participe plus activement à la construction du discours du Locuteur, soit en proposant des éléments pour le compléter, l’enrichir, etc., soit en demandant de telles informations. Contrairement aux cas autrement cités dans ces deux

¹⁷ Voir la différence entre régulateurs simples et complexes, dans la partie 1 de ce chapitre.

classifications, les régulateurs complexes et les demandes (ou ce qu'on pourrait désigner comme les régulateurs interrogatifs) donnent une légère impression de rupture vis-à-vis du discours du Locuteur, en ce qu'ils accompagnent celui-ci de manière plus personnelle, sans aller directement et uniquement dans le sens du Locuteur.

En excluant les régulateurs verbaux interrogatifs et complexes pour former le sous-ensemble des RVC, seuls les cas de régulateurs simples, répondant minimalement à la macro-fonction des régulateurs (accompagnement et encouragement de la parole du Locuteur) sont analysés. Ce sous-ensemble de régulateurs correspond ainsi aux formes prototypiques des régulateurs verbaux, ce qui justifie encore une fois leur sélection pour une étude exploratoire, plutôt que des formes plus complexes et moins emblématiques.

Pour conclure, nous définirons donc les régulateurs verbaux de connivence de la manière suivante :

Signaux d'écoute lexicaux (oui/ouais et non exclus) d'un interlocuteur se plaçant dans le rôle de Récepteur dans une conversation. Ces marqueurs sont indépendants du tour de parole du Locuteur possédant la parole (syntaxiquement), sont constitués d'une à environ 4-5 unités lexicales et ne constituent pas un tour de parole. Ils font partie des régulateurs simples et non complexes. Leur fonction pragmatique générale est d'accompagner le tour de parole du Locuteur, en allant dans le sens de ses propos et en l'encourageant à continuer à parler, exprimant ainsi la connivence. Ces régulateurs sont assertifs, et viennent confirmer tacitement au Locuteur que ses propos ont bien été enregistrés et qu'il n'y a pas de problème de compréhension. D'un point de vue interactionnel, ils apparaissent comme tous les régulateurs autour de points possibles de transition (TRP ; Sacks, Schegloff & Jefferson, 1974) dans la parole du Locuteur, et permettent de cette façon de ponctuer son discours et surtout d'indiquer à celui-ci qu'il peut continuer à parler (réaffirmation des rôles de Locuteur et de Récepteur) (de Gaulmyn, 1987 ; Bertand et al., 2009 ; Eiswirth 2020).

Cette définition étant établie, il s'agira désormais de détailler la typologie fonctionnelle de ces RVC telle que nous avons pu la développer dans ce mémoire. Comme nous allons le voir, cette typologie se situe dans la continuité de la littérature existante, avec à la fois une prise en compte des travaux précédents selon une volonté de synthétisation, mais aussi avec l'ajout et la description plus avancée de micro-fonctions spécifiques, le tout se basant sur les données empiriques observées dans le corpus.

Chapitre 4. La typologie des Régulateurs Verbaux de Connivence

Après avoir établi l'orientation majoritairement interactionniste de ce mémoire (Chapitre 1) et fait un point sur les notions d'organisation de la parole dans la conversation et de dimension médiale versus conceptionnelle de la parole (Chapitre 2), la définition générale des régulateurs ainsi que leur importance dans l'interaction ont pu être dégagées. Cela étant, l'objet d'étude de ce mémoire porte non pas sur tous les régulateurs possibles mais bien sur un sous-ensemble de régulateurs particuliers : les *régulateurs verbaux de connivence*, ou RVC, pour lesquels une définition a été proposée. Le présent chapitre aura pour but d'aller plus loin dans cet effort de définition, en détaillant la typologie fonctionnelle de ces RVC envisagée dans ce mémoire. Il s'agira tout d'abord de revenir sur les classifications fonctionnelles existantes, en identifiant leurs points communs et leurs éléments problématiques. Cette première étape permettra d'expliquer la typologie adoptée dans ce mémoire. Par la suite, il conviendra de définir et de donner des détails sur les types fonctionnels envisagés dans le cadre de ce mémoire (parties 2 et 3). Enfin, il sera question dans le quatrième point de ce chapitre des difficultés à garder en tête lors d'un travail de typologisation et d'analyse d'éléments linguistiques tels que les RVC, comme nous le proposons dans ce mémoire.

1. Influence des classifications fonctionnelles existantes sur la typologie des RVC : prise en compte de certains points communs et problématiques

La littérature au sujet des régulateurs est, nous avons déjà pu le voir, très diversifiée et foisonnante. De ce fait, quelques typologies fonctionnelles telles que celle que nous avons réalisée pour ce mémoire existent d'ores et déjà. Afin d'établir notre typologie des RVC, il a donc été nécessaire dans un premier temps de s'intéresser aux différentes micro-fonctions déjà détaillées dans ces classifications existantes, afin d'établir des points communs, mais aussi des problématiques à prendre dans compte dans l'élaboration de notre typologie.

Reprenons dans un premier temps le tableau des classifications fonctionnelles existantes (sans les régulateurs complexes, puisque les RVC sont des régulateurs simples uniquement).

Tableau 4 Résumé de quelques classifications fonctionnelles des régulateurs trouvées dans la littérature
(régulateurs complexes exclus)

<p>Goodwin, 1986</p>	<p>de Gaulmyn, 1987</p>	<p>Maynard, 1989 (cité par O’Keeffe & Adolphs, 2008 et Bertrand et al. 2009)</p>
<ul style="list-style-type: none"> • <i>Continuers</i> (Schegloff 1982): “treat [the current] unit as preliminary to further talk” (1986 : 210) • <i>Assessments</i>: “operate on the particulars of a current unit” (1986 : 210) ● ● ● 	<ul style="list-style-type: none"> • « <i>Le récepteur enregistre l’énonciation et/ou l’énoncé du locuteur (...)</i> » • « <i>Le récepteur évalue positivement l’énoncé du locuteur (...)</i> » ● ; ● ? (1987 : 207) 	<ul style="list-style-type: none"> • <i>Display of understanding of content</i> • <i>Support towards the speaker’s judgment</i> • <i>Agreement</i> • <i>Strong emotional response</i>
<p>O’Keeffe & Adolphs, 2008</p>	<p>Bertrand et al., 2009</p>	<p>Eiswirth, 2020a</p>
<ul style="list-style-type: none"> • <i>Continuer tokens</i> (after Schegloff, 1982): “maintain the flow of talk” • <i>Convergence tokens</i>: “markers of agreement/convergence” ● ; ● ? • <i>Engagement tokens</i>: “markers of high engagement where addressee(s) respond on an affective level to the content of the message” • (minor function) <i>Information receipt tokens</i>: “markers of points in the discourse where adequate information has been received” ● ? (2008: 16-17) 	<ul style="list-style-type: none"> • <i>Continuer</i> (prendre note minimalement) • <i>Understanding</i> (j’ai bien compris, mais sans degré d’adhésion, degré supérieur au <i>continuer</i>) ● ● • <i>Acknowledgment</i> (support, adhésion à un propos) • <i>Assessment</i> (évaluation, -rire par exemple-, jugement, déclaration d’attitude) (2009 : 187) 	<ul style="list-style-type: none"> • <i>Acknowledgment</i>: “both backward- and forward-looking in that they acknowledge the preceding talk and orient to the relevance of the current Speaker continuing without making any particular kind of continuation relevant” (2020a: 120) ● • <i>Surprise mark</i>: “show [the listener’s] appreciation of the content of a previous TCU as new, surprising, or otherwise unexpected” (2020a: 122) ● ? • <i>First Assessment</i>: “treat some aspect of the preceding talk as assessable” (2020a: 126) ● ● ● • <i>Second Assessment</i>: occur after “the Speaker makes a first assessment of some aspect of their own talk” (2020a: 129)

En observant les micros-fonctions proposées de plus près, il est possible de dégager quelques similarités, symbolisées par les différentes couleurs de police et points de couleurs (● ● ● ●). Il semble donc que la plupart des auteurs s’accordent, quoiqu’avec des portées et des termes différents, sur quatre micro-fonctions que nous pouvons synthétiser dans le tableau 5 ci-dessous :

Tableau 5 Tableau synthétique des fonctions majoritairement retrouvées dans la littérature

« prise en compte » (Bertrand et al., 2009)/engagement personnel Minimal ↓ Maximal	Simple enregistrement : le récepteur est conscient qu'un tour de parole est en cours et montre qu'il a suivi ce tour jusqu'ici et laisse le locuteur continuer.
	Soutien/ratification de l'acte d'énonciation du locuteur sans adhésion
	Soutien de l'énoncé du locuteur avec adhésion
	Évaluation affective de l'énoncé du locuteur

Il semble de plus que ces différentes fonctions puissent être associées à des organisations séquentielles spécifiques (*continuer* et *assessment*). En effet, la majorité des auteurs paraît s'accorder sur le fait que les régulateurs indiquant un enregistrement simple fonctionnent sur le modèle des *continuers*. Pour le reste des fonctions en revanche, leur organisation séquentielle entre *continuer* et *assessment* n'est pas souvent précisée¹⁸. Le consensus sur ces différentes fonctions est moindre, de plus, puisqu'elles ne sont pas toutes représentées ou « découpées » de la même manière dans les typologies fonctionnelles existantes. Ce fait est d'autant plus vrai dans le cas de la micro-fonction que nous avons résumée en tant que « évaluation affective » dans le tableau 5. Cette idée d'implication affective/émotionnelle est rendue dans les travaux de Maynard (1989, cité par O'Keeffe & Adolphs, 2008 et Bertrand et al. 2009), O'Keeffe & Adolphs, (2008) et Bertrand et al. (2009), toujours en des termes différents.

Un autre élément problématique de ces classifications fonctionnelles est la composante que nous avons appelée « engagement personnel du récepteur » et que Bertrand et al. (2009, p. 5) résument comme le « degré de « prise en compte » par l'interlocuteur »¹⁹. Il semble en effet que, dans la majorité des typologies extraites de la littérature, il existe une certaine gradation entre les différentes fonctions des régulateurs, du moins pour les régulateurs simples. Cependant, mis à part Bertrand et al. précédemment cités, aucun auteur n'explique cette gradation. Nous avons ainsi de notre côté souhaité remédier à ce problème

¹⁸ Il semble cependant que ces trois micros-fonctions relèvent plus de l'*assessment*, étant donné leur portée sur le discours précédent du Locuteur (idée de soutien ou d'évaluation).

¹⁹ Symbolisé dans la colonne de gauche du tableau 5. Le code couleur, du rouge au bleu dans les tableaux 4 et 5, reprend aussi cette idée de gradation.

en l'incluant dans notre typologie, puisque cette notion de gradation semblait pertinente pour rendre compte à la fois des spécificités, mais aussi des liens existants entre ces différentes micros-fonctions.

Il est aussi important de noter que les différentes études résumées dans le tableau 4 ci-dessus portent sur des langues différentes. Or, étant donné l'ancrage fortement interactionnel des régulateurs, il est très probable que leurs micro-fonctions varient en réalité selon la langue parlée. Les quatre micro-fonctions détaillées dans le tableau 5 paraissent ainsi être des fonctions générales, mais dont il est nécessaire de ne pas présumer l'existence dans chaque langue. La typologie fonctionnelle que nous proposons dans ce mémoire est de cette manière uniquement valable pour le français hexagonal, en l'absence d'étude comparative sur d'autres langues.

Pour terminer, il existe un dernier problème spécifique à la littérature francophone au sujet des régulateurs qui est que, comme l'indique Persson (2014, p. 115), les études francophones « n'ont pas attaché d'intérêt particulier au caractère systématique et ordonné de l'usage contextualisé de formes particulières, en tant que régulateurs ». Il y a donc très peu de prise en compte et de détails sur les formes et fonctions spécifiques des régulateurs dans la littérature française. Or, il est établi dans la littérature internationale que les régulateurs sont sujets à la variation culturelle, comme nous l'avons aussi rappelé dans le paragraphe précédent. Si ce fait a surtout été démontré au niveau des formes et des fréquences d'utilisation (voir Maynard, 1990 pour une étude comparative anglais américain/japonais ou encore O'Keeffe & Adolphs, 2008, pour une étude comparative anglais britannique/irlandais), il n'est pas exclu que cette variation existe aussi au niveau des fonctions des régulateurs. L'établissement d'une typologie fonctionnelle des régulateurs verbaux de connivence dans le cadre de ce mémoire, basée sur leur usage observé dans un corpus de conversations entre locuteurs du français hexagonal, serait ainsi un premier pas vers une meilleure compréhension des usages particuliers des régulateurs en français.

Avant de passer à la description de la typologie établie dans ce mémoire, qui a pour objectif d'apporter un début de réponse à ces différents problèmes, il conviendra cependant de revenir dans un premier temps sur une notion qui devrait nous permettre de mieux appréhender le cas des régulateurs véhiculant de l'affect, la notion d'expressivité.

Afin de mieux comprendre le fonctionnement des régulateurs tels que les « évaluations affectives », nous avons ainsi décidé de nous tourner vers d'autres travaux

proposant des typologies d'unités pragmatiques de la langue (marqueurs discursifs, phraséologismes, etc.). Notre souhait était de vérifier si d'autres typologies envisageaient l'expression de l'affect comme une fonction langagière spécifique. Cela a effectivement été le cas, tout particulièrement dans l'étude de Gharbi (2020), proposant une typologie de phraséologismes justement appelés *formules expressives de la conversation* (d'après les travaux de Krzyżanowska & Grossman²⁰). Cette étude nous a permis de mieux comprendre la notion d'expressivité, mais aussi de confirmer en partie notre typologie dans un deuxième temps (voir partie 3).

Revenons donc tout d'abord sur la notion d'expressivité. Comme le pointe Gharbi (2020), mais aussi Kleiber (2006), l'expressivité, tout comme l'exclamativité, représente de manière indexicale une émotion ou un affect du locuteur. Tout énoncé véhiculant de l'affect et permettant d'exprimer des sentiments pourra ainsi être considéré comme expressif. Il est important de noter de plus que ces sentiments peuvent être plus ou moins forts, voire peuvent être mis en scène dans certains cas²¹.

Le choix a donc été fait, dans le cadre de ce mémoire, de garder le terme d'*expressivité* pour parler des régulateurs véhiculant de l'affect. Ce choix permet d'une part de faire écho à la littérature existante sur les phraséologismes, et d'autre part de proposer une typologie plus nuancée des RVC.

Cette précision étant établie, passons désormais à la description de la typologie que nous avons établie. Il est à noter que cette typologie n'est pas définitive, et pourra être revue après la première tentative d'annotation que constitue ce mémoire.

Trois types fonctionnels principaux ont pu être dégagés dans cette typologie : les marques d'enregistrements, les marques de confirmation ainsi que les expressifs. Les expressifs se divisent ensuite en deux sous-types fonctionnels, nommés marques de convergence et marques d'évaluation. Tous ces différents types peuvent être envisagés selon une gradation dans ce que nous avons appelé l'« engagement personnel du récepteur », à son niveau minimal lorsque le récepteur utilise une marque d'enregistrement, et à un niveau


²⁰ Voir par exemple l'ouvrage récent édité par Krzyżanowska, Grossman & Kwapisz-Osadnik : *Les formules expressives de la conversation : Analyse contrastive : Français-Polonais-Italien* (2021).

²¹ Voir partie 3 de ce chapitre.

beaucoup plus important lorsque le récepteur utilise un RVC expressif. Nous avons aussi inclus à titre indicatif dans cette typologie un lien avec les types *continuers* et *assessments* (Goodwin, 1986), qui reflètent la portée du régulateur (attente d'une suite *versus* évaluation/traitement du segment précédent).

Sous forme de tableau, cette typologie se présente de la manière suivante :

Tableau 6 Présentation synthétique de la typologie des régulateurs verbaux de connivence

Degré d'engagement personnel du récepteur	Organisation séquentielle du RVC	Type	Sous-type	Exemples de formes récurrentes ²²	
Minimal  Maximal	<i>Continuer</i> <i>Assessment</i>	Marque d'enregistrement ●		<i>d'accord, OK</i>	
		Marque de confirmation ●		<i>c'est clair, ouais c'est vrai, c'est sûr</i>	
	<i>Assessment</i>	Expressif	Marque de convergence ●		<i>tu m'étonnes, grave, c'est normal</i>
			Marque d'évaluation ●		Formes composées : <i>c'est chaud, c'est fou ça, c'est cool</i>
				Interjections : <i>putain, merde, la vache, punaise</i>	

Comme l'indiquent les points de couleurs compris dans le tableau 6, les différentes fonctions que nous avons dégagées peuvent être mises en lien avec les quatre fonctions principalement retrouvées dans la littérature existante (voir tableau 5). Les types fonctionnels envisagés ici ne sont pas pour autant des copies conformes des fonctions retrouvées dans la littérature, mais peuvent s'en éloigner un peu.

²² Pour un inventaire plus complet des formes observées par type, voir le chapitre 9. 3. .2.

Ces différents types et sous-types fonctionnels seront discutés en détail dans les parties suivantes, et des exemples seront proposés pour chacun d'entre eux. Nous détaillerons tout d'abord le fonctionnement des marques d'enregistrement et marques de confirmation (partie 2), puis le fonctionnement des RVC expressifs et des sous-types correspondants (marques de convergence et marques d'évaluation) (partie 3). Pour terminer sur cette typologie, nous reviendrons dans la quatrième partie de ce chapitre sur quelques difficultés qu'il a été nécessaire de prendre en compte pour ce travail de typologisation des *régulateurs verbaux de connivence*.

2. Les régulateurs comme marques d'enregistrement et de confirmation

Les deux premiers types fonctionnels de régulateurs verbaux de connivence envisagés dans notre typologie, les marques d'enregistrements et les marques de confirmation, possèdent de nombreux points communs. Tout d'abord il s'agit de types fonctionnels non expressifs, puisqu'ils ne reflètent pas d'affect particulier de la part du récepteur, fait qui est aussi à mettre en lien avec leur degré d'engagement personnel faible. Ces deux types sont donc axés principalement sur le Locuteur, et permettent de l'accompagner dans son discours et de lui signaler le bon déroulement de la conversation, sans pour autant faire intervenir la subjectivité du Récepteur.

2.1. Marques d'enregistrement

Les RVC fonctionnant comme marques d'enregistrement (désormais ME) permettent au Récepteur de signaler sa bonne réception du propos du Locuteur, mais sans lui apporter d'informations sur son interprétation de ce même propos. Ils correspondent en ce sens à un engagement personnel minimal du Récepteur : leur rôle est de montrer que les propos du Locuteur ont bien été « enregistrés » dans la mémoire du Récepteur (voir l'hypothèse de de Gaulmyn stipulant que les régulateurs constituent des « trace(s) des opérations de compréhension et de stockage en mémoire chez l'auditeur » (1987, p. 210)). Ils ne font cependant rien de plus que de signaler cet enregistrement : en émettant une ME, le Récepteur ne laisse rien paraître de son interprétation du discours du Locuteur, il n'y a aucune ratification de ce discours. Cet enregistrement peut donc correspondre au stockage en mémoire/à la prise en compte simple de l'information ou bien à la compréhension de l'information mise en mémoire.

D'un point de vue théorique, l'hypothèse soulevée initialement était que les RVC ayant fonction de marque d'enregistrement étaient de type *continuer*, c'est-à-dire qu'ils permettaient de montrer au Locuteur que la suite de son discours était attendue (voir Chapitre 3.1). En revanche, les données semblaient dès le départ aller à l'encontre de cette hypothèse, puisque les RVC utilisés pour cette fonction minimale d'enregistrement pouvaient aussi se trouver en clôture de thème par exemple. Il semble donc que les régulateurs fonctionnant comme ME puissent être utilisés à la fois pour signifier au locuteur que la suite de son discours est attendue (type *continuer - continueur*), mais aussi pour indiquer que le discours du locuteur jusqu'ici a bien été pris en compte et enregistré/compris par le récepteur (type *assessment - évaluation*).

Exemples :

(3)

Exemple 3 *OK* comme marque d'enregistrement (type *continuer*) (extrait de voyage_hab_14, TCOF)

- | | | |
|---|----------------|--|
| 1 | -11_TCOF_11_01 | bah le régime s'est fait renversé et la chute du Shah d'Iran (0.5) |
| 2 | | et mon père il est venu quand les islamistes sont venus au pouvoir en fait (0.3) |
| 3 | -12_TCOF_11_02 | OK (1.0) |
| 4 | -11_TCOF_11_01 | et vu qu'il y avait déjà mon oncle qui habitait ici |

(4)

Exemple 4 *d'accord* comme marque d'enregistrement (type *assessment*) (extrait de voyage_hab_14, TCOF)

- | | | |
|----|----------------|---|
| 1 | -11_TCOF_11_01 | euh oui je suis allée voir toute la famille du côté de mon papa |
| 2 | -12_TCOF_11_02 | oh |
| 3 | -11_TCOF_11_01 | et du coup euh |
| 4 | -12_TCOF_11_02 | ((rires)) |
| 5 | -11_TCOF_11_01 | du coup hum ((sonnerie de téléphone)) |
| 6 | | donc voilà [quoi] |
| 7 | -12_TCOF_11_02 | [donc] tu as vu la famille de ton papa |
| 8 | -11_TCOF_11_01 | voilà |
| 9 | -12_TCOF_11_02 | d'accord (1.0) |
| 10 | | et c- c- c'est vraiment différent là-bas d'ici ou |

2.2. Marques de confirmation

Les RVC ayant fonction de marque de confirmation permettent au Récepteur de montrer au Locuteur qu'il soutient son discours en venant lui apporter une ratification (toujours sans expressivité et sans notion d'adhésion cependant). Ainsi la marque de confirmation, sans être un réel accord entre les deux locuteurs, est un moyen pour le

Récepteur d'aller dans le sens du Locuteur, et de lui montrer son soutien. Le nom donné à ces marqueurs (« de confirmation ») a été choisi pour indiquer qu'ils viennent à la fois confirmer la prise en compte du discours du Locuteur, mais aussi confirmer le Locuteur dans son propos²³.

Si le Récepteur ne s'engage pas complètement dans la communication lorsqu'il émet une marque de confirmation, ce type de RVC indique toutefois un engagement personnel du Récepteur plus important que les marques d'enregistrement, qui correspondent à un degré d'engagement minimal.

Exemple :

(5)

Exemple 5 *c'est clair* comme marque de confirmation (extrait de apéritif entre ami(e)s – pois, CLAPI)

- | | | |
|---|-----------------|---|
| 1 | -15_CLAPI_02_01 | ben on y go enfin de toute façon au ninkasi en fait on va y aller à pied je pense |
| 2 | -15_CLAPI_02_02 | ouais |
| 3 | -17_CLAPI_02_03 | ouais |
| 4 | -15_CLAPI_02_01 | comme ça on pourra regarder ce qu'y a (0.3) |
| 5 | -17_CLAPI_02_03 | ouais (.) c'est clair |

Ces deux premiers types fonctionnels étant détaillés, il s'agira désormais de préciser la définition du type fonctionnel expressif et de ses deux sous-types, les marques de convergence et marques d'évaluation.

3. *La fonction expressive*

Contrairement aux marques d'enregistrement et marques de confirmation, les RVC expressifs permettent au Récepteur d'être plus engagé dans la conversation, et de faire part de sa subjectivité face au discours du Locuteur. Le Récepteur peut cependant faire porter cette subjectivité sur des aspects différents du discours du Locuteur, aspects qui correspondent aux deux sous-types expressifs envisagés dans la typologie présentée ici : les marques de convergence et les marques d'évaluation. Ces deux sous-types seront traités dans l'ordre mentionné ci-dessus, mais avant cela, nous reviendrons sur certaines caractéristiques

²³ Toujours sans correspondre à un réel accord cependant.

du type fonctionnel expressif en général, qui permettront de mieux appréhender ce type fonctionnel et ses sous-types.

Les RVC de type expressif permettent au Récepteur d'exprimer un fort engagement personnel dans la conversation. Leur qualité d'expressif est liée à leur association, quoique plus ou moins importante, à une certaine émotion, un certain affect. Ce lien avec l'émotion repose en partie sur la prosodie, mais aussi et surtout sur la sémantique des RVC. En effet, les RVC expressifs sont nécessairement associés, par leur sémantique, à l'émotion et/ou à la subjectivité. Ainsi les lexèmes les plus représentés pour ce type de RVC sont par exemple *bien*, *cool*, *chaud* ou encore *putain*. Ce lien plus ou moins fort avec l'émotion et la subjectivité du Récepteur, non exprimé dans le cas des marques d'enregistrement ou de confirmation, explique le fait que la fonction expressive permette au Récepteur de manifester un engagement personnel plus important.

Au niveau interactionnel, les RVC expressifs interviennent souvent au niveau de frontières thématiques et permettent ainsi de ponctuer la fin d'un thème (comme c'est le cas dans les exemples (6) et (7) présentés plus bas, tous deux suivis par le segment *voilà*). Les RVC expressifs portent sur l'énoncé les précédant (type *assessment*), et permettent au Récepteur d'apporter une réaction affective à cet énoncé. Nous avons pu dégager deux types de RVC expressifs avec des fonctions légèrement différentes, les marques de convergences et les marques d'évaluation, que nous allons désormais nous atteler à détailler. Il reste cependant important de noter que ces deux fonctions reposent fortement sur la connivence, l'intersubjectivité entre le Récepteur et le Locuteur : le Récepteur, fortement engagé dans la conversation, exprime sa propre subjectivité tout en continuant à aller le plus possible dans le sens du Locuteur.

Les deux sous-types fonctionnels envisagés ci-dessous peuvent être retrouvés de façon assez similaire dans la typologie des formules expressives de la conversation établie par Gharbi (2020), qui distingue une catégorie de formules dites interpersonnelles réactionnelles et une autre dite évaluative.

3.1. Marques de convergence

Les RVC ayant fonction de marques de convergence sont des RVC expressifs axés sur l'intersubjectivité : ils permettent au Récepteur d'indiquer au Locuteur qu'il approuve sa parole, et que leurs subjectivités se rejoignent (forte convergence). Si ces RVC ne renvoient pas à une réelle prise de position de la part du Récepteur (ils restent avant tout des régulateurs

et ne sont donc pas argumentatifs, ne correspondent pas à une réelle prise de parole), ils sous-entendent bien un accord du Récepteur avec les propos du Locuteur ou avec le point de vue du locuteur.

Leur aspect expressif est lié au fait qu'ils permettent d'exprimer un fort sentiment de convergence, d'accord. De plus, ils sont souvent utilisés pour manifester une réaction empathique face à un sentiment donné ou des propos mêlés d'affect de la part du Locuteur.

Exemple :

(6)

Exemple 6 *tu m'étonnes* comme marque de convergence (expressif) (extrait de voyage_hab_14, TCOF)

- | | | |
|---|----------------|---|
| 1 | -11_TCOF_11_01 | ben ça ça faisait vraiment un changement avec ma vie de tous les jours |
| 2 | | et tout c'est [pas du tout] |
| 3 | -12_TCOF_11_02 | [bah oui] |
| 4 | -11_TCOF_11_01 | le même train de vie il y a plein d'autres contraintes que il y avait pas en France et tout |
| 4 | -12_TCOF_11_02 | tu m'étonnes |
| 5 | -11_TCOF_11_01 | du coup voilà + |

3.2. Marques d'évaluation

Les RVC expressifs fonctionnant comme marques d'évaluation permettent au Récepteur de laisser transparaître leur sentiment appréciatif ou dépréciatif vis-à-vis du propos du Locuteur, en apportant une évaluation positive ou négative sur celui-ci. Cette évaluation reste toutefois relativement implicite, puisqu'encore une fois les régulateurs ne sont pas argumentatifs et ne sont pas de réelles prises de parole.

Les marques d'évaluation possèdent aussi un fort aspect intersubjectif, puisque l'évaluation apportée reflète dans une grande partie des cas à la fois les sentiments du Récepteur face au discours du Locuteur, mais aussi les sentiments supposés/la réaction du locuteur vis-à-vis de la situation décrite. Ils permettent ainsi de tester la bonne compréhension du discours du Locuteur, et la connivence entre les participants.

Les marques d'évaluation constituent le type de RVC pouvant prendre les formes les plus variées, comme nous le reverrons dans l'analyse. Parmi les formes que peuvent prendre ces marques d'évaluation, nous avons intégré dans le cadre de ce mémoire les interjections (*putain, merde, la vache*, etc.), pour les raisons suivantes :

- Les interjections²⁴ sont fortement expressives, puisqu'elles sont « des indices ou index en ce qu'au moment de leur émission, elles sont reliées de manière indexicale à l'émotion éprouvée par le sujet qui les prononce » (Kleiber, 2006, p. 18).
- En émettant une interjection, le Récepteur se présente donc comme ressentant un certain affect, une certaine émotion, et l'on peut considérer que cette émotion est éprouvée vis-à-vis du discours du Locuteur, en réaction à ce discours. Il s'agit donc d'une évaluation implicite²⁵, pouvant être appréciative ou bien dépréciative (l'émotion étant liée à une situation positive ou bien négative).

Selon Kleiber (2006), les interjections peuvent aussi être utilisées volontairement, et permettre une certaine mise en scène de l'émotion même si celle-ci n'est pas réellement ressentie par le locuteur. Ce fait est particulièrement important dans le cas des régulateurs, puisqu'il permet d'expliquer les situations où la prosodie ne reflète pas l'expressivité de l'interjection (par exemple un *putain* prononcé sur un ton factuel). Il permet de plus de justifier, pour les interjections en particulier, mais aussi pour certaines formes lexicales avec un sémantisme renvoyant à une certaine émotion ou une forte subjectivité, le statut de RVC expressif y compris dans des situations où la prosodie ne renvoie pas ou très peu d'émotion²⁶.

Exemple :

(7)

²⁴ Nous traitons ici uniquement des interjections secondaires et non primaires (Halté, 2013), puisque seules les interjections possédant un contenu lexical peuvent être considérées comme des RVC.

²⁵ Cette réaction pouvant encore une fois posséder à la fois un aspect très personnel et correspondre à la réaction du récepteur seul, et aussi un aspect intersubjectif et correspondre à une estimation des sentiments du locuteur.

²⁶ Nous reviendrons sur ce fait et sur la façon dont il a été appliqué lors du traitement des données dans la partie méthodologie (chapitre 6)

Exemple 7 *oh putain* et *oh c'est une horreur* comme marques d'évaluation (expressifs) (extrait de tourisme_arn_15, TCOF)

- 1 -10_TCOF_10_02 bon ça à la limite ça passe mais quand ça fait deux semaines qu'ils traînent là euh
2 ça va et le pire je crois c'est quand elle a eu ses règles et qu'elle a laissé son tampon plein de sang
3 -13_TCOF_10_01 **oh putain**
4 -13_TCOF_10_03 **oh c'est une horreur** (0.7)
5 -10_TCOF_10_02 [voilà]
6 -13_TCOF_10_01 [mais] elle a été élevée chez des cochons elle ou quoi

Pour conclure, la typologie dégagée dans ce mémoire peut être rapprochée des différentes micro-fonctions généralement retrouvées dans les classifications fonctionnelles existant dans la littérature. Quelques différences vouées à régler les problèmes trouvés dans ces classifications ont toutefois été adoptées, notamment la prise en compte d'une fonction expressive ainsi que d'une gradation dans le degré d'engagement personnel du Récepteur. Trois fonctions majeures ont ainsi pu être envisagée dans notre typologie : une fonction de marque d'enregistrement, de marque de confirmation ainsi qu'une fonction expressive. Deux sous-fonctions (marque de convergence et marque d'évaluation) peuvent de plus être retrouvées au sein de cette fonction expressive.

Afin de réaliser la typologie présentée ci-dessus, un travail important de reconnaissance et de sensibilisation à diverses difficultés a été effectué. Ces difficultés sont liées à la nature des régulateurs verbaux, qui possèdent un contenu lexical pouvant aussi être utilisé de façon référentielle, en tant que réelle prise de parole. Un autre danger lié à ce contenu lexical est l'écueil de l'assimilation stricte entre forme et fonction, d'autant plus que le contenu lexical retrouvé dans certains régulateurs peut souvent être mobilisé pour remplir d'autres fonctions dans les interactions. Ces différents problèmes et la façon dont ils ont été pris en compte sera donc détaillée ci-dessous.

4. Quelques difficultés liées à l'élaboration d'une typologie pour les RVC, et au traitement des régulateurs en général

Avant de passer à l'analyse du corpus, il convient de revenir sur quelques difficultés liées à la nature même des régulateurs verbaux, qui ont un fort impact sur notre étude et sur toute étude traitant de ce type d'élément linguistique. La première difficulté ainsi discutée concernera l'identification des régulateurs en eux-mêmes, puis un deuxième problème plus

large concernera le lien entre la forme et la fonction d'unités pragmatiques tels que les régulateurs.

La première étape de tout travail d'analyse sur des unités linguistiques précises est d'identifier et de segmenter les unités en question. Dans le cas des régulateurs, cette segmentation peut cependant être très difficile dans certains contextes, car il n'est pas toujours aisé, pour le/la chercheur.se, de déterminer la nature de certaines interventions du Récepteur, entre régulation et prise de parole (Kerbrat-Orecchioni, 2001 ; de Nuchèze & Coletta, 2002). Cela est d'autant plus vrai pour les régulateurs verbaux, qui possèdent un contenu lexical et sémantique pouvant être référentiel (de Gaulmyn, 1987). Par exemple, un *c'est vrai* en réponse peut être un régulateur et venir confirmer le Locuteur dans sa position sans correspondre à une prise de position du Récepteur. Mais ce même *c'est vrai* peut aussi être une intervention référentielle, et correspondre à une prise de position du locuteur produisant ce segment (réel accord). Plusieurs solutions doivent donc être envisagées afin de faire face à cette difficulté, telles que le recours à plusieurs annotateurs (par exemple dans le travail de Bertrand et al., 2009) ou à la *next-turn proof-procedure* (Eiswirth, 2020b), et de manière générale au contexte interactionnel dans lequel le segment analysé est ancré. Cette prise en compte du contexte est d'autant plus importante que l'interprétation du Locuteur ne concorde pas toujours avec l'intention du Récepteur, comme le rappelle de Gaulmyn (1987, p. 211) : « l'hésitation de l'interprète reflète l'incertitude même des choix opérés par les locuteurs (...) L'interprétation de l'allocutaire peut être concordante ou discordante avec l'intention du locuteur ». Les différentes solutions adoptées dans ce mémoire (dont la *next-turn proof-procedure*) seront discutées dans le chapitre 6.

D'autres problèmes se posent aussi pour le/la chercheur.se s'intéressant aux régulateurs, notamment au niveau du lien entre la forme et la fonction de ces unités linguistiques.

Ainsi certaines formes lexicales pouvant fonctionner comme régulateurs sont fortement multifonctionnelles, et peuvent être utilisées, outre comme régulateurs, comme marqueurs discursifs et remplir une autre fonction pragmatique ou bien être référentielles et correspondre à une réelle prise de parole. Certaines formes lexicales peuvent aussi être utilisées en tant que RVC, mais selon des types fonctionnels différents. C'est par exemple le cas de *OK*, qui est fortement multifonctionnel. On le retrouve, lorsqu'il est utilisé comme RVC, majoritairement en tant que marque d'enregistrement, mais aussi en tant que marque

d'évaluation dans certains cas, avec la prosodie correspondante. Ainsi dans l'exemple ci-dessous, *OK* est utilisé comme marque d'évaluation, sur un ton moqueur et ironique.

(8)

Exemple 8 *OK* comme marque d'évaluation (expressif) (extrait de tourisme_arn_15, TCOF)

1	-10_TCOF_10_02	ce midi je lui ai dit euh
2		bon appétit elle me fait ((imite la personne)) merci toi aussi
3	-13_TCOF_10_03	((rire)) OK
4	-13_TCOF_10_01	la bucheronne
5	-13_TCOF_10_03	gra- ((rire))
6	-10_TCOF_10_02	j'ai eu trop peur en plus et cet après-midi j'ai dû faire euh

Pour terminer, il est aussi très important de noter, comme le fait de Gaulmyn (1987), que l'interprétation sémantique d'un régulateur verbal n'est pas égale à son interprétation fonctionnelle, et qu'il est donc nécessaire de faire une différence entre les deux. Ce fait explique d'ailleurs qu'une même forme lexicale puisse remplir différentes fonctions en tant que RVC : ce n'est pas la sémantique d'un RVC qui détermine sa fonction, mais bien la façon dont il est utilisé dans la situation d'énonciation. L'exemple ci-dessus (8) explique très bien cette différence entre l'interprétation sémantique et l'interprétation fonctionnelle, puisque la sémantique de *OK* indique plutôt l'accord, l'assentiment. Or, dans cet exemple, *OK* est utilisé en tant que régulateur (associé à un ton moqueur), et a pour fonction, pour la Réceptrice (-13_TCOF_10_03), de signaler à la Locutrice (-10_TCOF_10_02) une évaluation moqueuse d'un élément de son énoncé (l'évaluation porte ici sur la personne imitée par la Locutrice).

Pour conclure, le/la chercheur.se travaillant sur les régulateurs, et en particulier les régulateurs verbaux de connivence, doit toujours prêter une attention particulière à la situation d'énonciation dans laquelle le RVC intervient, car ces unités peuvent aussi correspondre à de réelles prises de parole dans certains cas d'une part, et d'autre part car les formes lexicales pouvant être utilisées comme RVC sont fortement multifonctionnelles et ne sont pas associées à une seule fonction. Le travail d'analyse, à l'aide de la typologie, doit être effectué en prenant en compte ces différentes difficultés et en vérifiant systématiquement le contexte d'occurrence du segment étudié.

La façon dont ces différentes difficultés ont été abordées lors de l'élaboration de notre typologie relève de la méthodologie adoptée dans le cadre de ce mémoire, méthodologie qui sera détaillée dans la partie suivante. Suivant les étapes de travail et

d'analyse, il s'agira tout d'abord de décrire la façon dont le corpus utilisé dans ce mémoire a été choisi et élaboré (chapitre 5), puis de décrire la méthodologie de l'élaboration de la typologie (chapitre 6). Le chapitre 7 sera consacré à l'annotation du corpus (élaboration d'une grille de score permettant de calculer le degré de proximité des interactants et annotation sous ELAN). Enfin, le dernier chapitre de cette partie consacrée à la méthodologie (chapitre 8) portera sur le traitement final des données issues de l'annotation.

Partie 2

-

Méthodologie

Chapitre 5. Choix et élaboration du corpus

La toute première étape de ce travail de mémoire a été l'élaboration d'un corpus de données orales sur lesquelles baser nos observations.

En lien avec l'approche *corpus-driven* de ce travail de mémoire, des corpus oraux tirés d'une seule base de données de corpus oraux, le TCOF²⁷, avaient été initialement sélectionnés comme base de travail. L'écoute de corpus tirés de cette base de données a permis de choisir l'objet d'étude pour ce mémoire, et des enregistrements tirés d'autres corpus ont ensuite été ajoutés à notre base de données pour ce mémoire.

L'objectif primaire de ce travail de mémoire, même sans objet d'étude délimité, était d'étudier des phénomènes apparaissant dans les conversations quotidiennes du français hexagonal, c'est-à-dire dans le vernaculaire. À ce titre, et étant donné les changements rapides du langage vernaculaire, un sous-ensemble de corpus tirés du TCOF avait été sélectionné comme base pour le corpus de ce mémoire. En effet, le TCOF présente les avantages suivants :

- Il s'agit d'un des corpus français proposant les données orales les plus récentes²⁸.
- Le site hébergeant le corpus offre la possibilité de chercher des enregistrements selon des critères spécifiques, et ce d'une manière simple et efficace.
- Les enregistrements présentés dans ce corpus ont été pris dans des situations assez variées, par exemple au niveau du type de conversation ou du lien entre les interactants.

²⁷ Pour « Traitement de Corpus Oraux en Français », développé au sein du laboratoire ATILF (UMR CNRS 7118) à l'université de Nancy 2. Disponible aux URL suivants : <https://tcof.atilf.fr/index.php> ou <https://hdl.handle.net/11403/tcof/v2.1>

²⁸ Les enregistrements les plus récents au sein de ce corpus dataient de 2014 et de 2015 au moment du téléchargement du corpus, durant le mois de février 2021.

Afin de sélectionner uniquement des enregistrements présentant un parler le plus ordinaire possible, seuls les enregistrements répondant aux critères suivants (obtenus grâce à une recherche sur le site) ont été sélectionnés :

- Type de corpus : entre adulte
- Cadre situationnel : privé
- Genre de discours : conversation

Un total de vingt-trois conversations privées avait ainsi été téléchargé initialement. Certaines ont cependant dû être écartées, pour des raisons que nous détaillerons un peu plus bas. Au total seules onze de ces conversations ont été conservées dans le corpus final. De ce fait, afin d'avoir un corpus plus conséquent (et donc la possibilité de faire des analyses plus étoffées), des corpus oraux extraits d'autres bases de données (MPF et CLAPI) ont aussi été téléchargés pour constituer le corpus final de ce mémoire. Le but était d'avoir des conversations avec des contenus variés (que ce soit au niveau du type de conversation ou des thèmes abordés), afin de vraiment aborder la variation des RVC. Si la réflexion ne s'est pas réellement posée en termes de volume, le but était tout de même d'obtenir un corpus de plus de 5h minimum d'oral.

Suivant cette réflexion, six conversations privées ont été téléchargées à partir du corpus MPF (*Multicultural Paris French*), disponible sur ORTOLANG (<https://hdl.handle.net/11403/mpf/v3>). Ce corpus porte spécifiquement sur le français vernaculaire de la région parisienne, en particulier le français parlé par les jeunes issus d'un milieu multiculturel. Ces six conversations ont été choisies sur la base de leur classification en tant qu'« événement écologique », classification repérée sur le site <https://mpfvitrine.modyco.fr/vitrine/>. Trois de ces conversations ont été conservées dans le corpus final.

Pour terminer, quatre conversations entre ami(e)s ont été téléchargées à partir de la banque de données de corpus CLAPI (Corpus de Langues Parlées en Interaction) (<http://clapi.icar.cnrs.fr>). Ces quatre conversations correspondent aux interactions privées les plus récentes de CLAPI (2009). Seule une de ces conversations a été écartée du corpus final.

Tous les corpus ainsi sélectionnés présentaient l'intérêt majeur d'être disponibles sous des formats compatibles avec le logiciel ELAN, que nous avons utilisé pour l'annotation et l'analyse de notre objet d'étude, les *régulateurs verbaux de connivence* ou RVC (voir chapitre 3. 3). Ainsi tous les corpus proposaient des fichiers audio au format .wav,

et des transcriptions complètes au format .trs (TCOF et CLAPI) ou .TextGrid (MPF). Ces transcriptions déjà existantes ont dû être très légèrement modifiées dans certains cas lors de l'annotation du corpus, modifications sur lesquelles nous reviendrons plus en détail dans le chapitre 7.

Comme cela a déjà été mentionné, toutes les conversations initialement téléchargées n'ont pas été conservées pour former le corpus final. Certaines conversations ont pu être écartées pour différentes raisons, telles que :

- Des problèmes techniques (absence de métadonnées, de transcriptions, etc.).
- Un nombre trop important de locuteurs : afin de ne pas avoir trop de variables à contrôler et analyser dans notre analyse, seules les conversations à deux ou trois locuteurs ont été conservées.
- Les conversations peu naturelles avec un taux très élevé d'hypercorrection de la part des participants ou les conversations excessivement monologiques ont été écartées.
- Afin de privilégier et d'étudier l'impact de la variation de la proximité dans le corpus, certaines conversations avec des formats très similaires ont été écartées. Beaucoup de conversations tirées du corpus TCOF présentaient ainsi un format de « récit d'expérience », avec un locuteur gardant la parole pour raconter une expérience pendant la majeure partie de la conversation. Une partie de ces conversations a par conséquent été exclue du corpus final, généralement pour cause de manque de naturel ou de brièveté.

Le corpus obtenu présente des avantages non négligeables : tout d'abord il est relativement récent, avec des enregistrements datant au maximum d'une douzaine d'années et une variété de situations de communications, majoritairement naturelles. Toutes les interactions se déroulent en face à face, dans un cadre privé pour la majeure partie d'entre elles. Certains biais existent toutefois dans ce corpus, biais qu'il est nécessaire de mentionner même s'ils n'impactent pas réellement l'étude exploratoire proposée dans ce mémoire. Ainsi le corpus présente un fort déséquilibre dans les tranches d'âge et les genres représentés, avec

32 locuteurs dans la tranche d'âge allant de 18 à 29 ans²⁹, et 29 locutrices sur les 42 locuteurs. Un autre biais est celui de la spontanéité et du naturel des interactions du corpus. En effet, même si les interactions choisies pour faire partie du corpus ont été triées pour être les plus spontanées et naturelles possibles, certaines d'entre elles présentent des moments ou aspects non naturels (hypercorrection : utilisation d'un registre soutenu, contrôle important sur le rythme et le volume utilisé, etc.).

Ces différents biais n'ont cependant qu'un impact limité sur notre étude, puisqu'il s'agit uniquement d'une étude exploratoire. L'objectif primaire est de tenter d'établir une typologie des fonctions généralement observées des régulateurs verbaux de connivence, puis, dans une moindre mesure, de réfléchir à leur variation formelle et variation selon le degré de proximité des interactants.

Le contenu du corpus peut être résumé dans le tableau suivant :

Tableau 7 Contenu du corpus final utilisé dans le cadre du présent mémoire

	TCOF	MPF	CLAPI
17 conversations, 07h24min03s de parole 42 locuteurs	11 conversations	3 conversations	3 conversations
	3h11m58s	1h40m10s	2h31m55s
	26 locuteurs : 3 locuteurs = 4 conv. 2 locuteurs = 7 conv.	8 locuteurs : 3 locuteurs = 4 conv. 2 locuteurs = 7 conv.	8 locuteurs : 3 locuteurs = 4 conv. 2 locuteurs = 7 conv.
	5 conversations de 2014 6 conversations de 2015	2 conversations de 2011 1 conversation de 2012	2009
	Résumé des types de conversation retrouvés	4 récits d'expérience 1 entretien 2 discussions libres 1 échange de conseils 3 récits d'expérience + discussion	1 échange de nouvelles 2 conversations mixtes (repas + échange de nouvelles + débat ; récit + explications + débat)

Pour résumer, le corpus utilisé dans le cadre de ce mémoire a été établi avec un objectif principal de recueil d'une langue vernaculaire, parlée dans des conversations en face

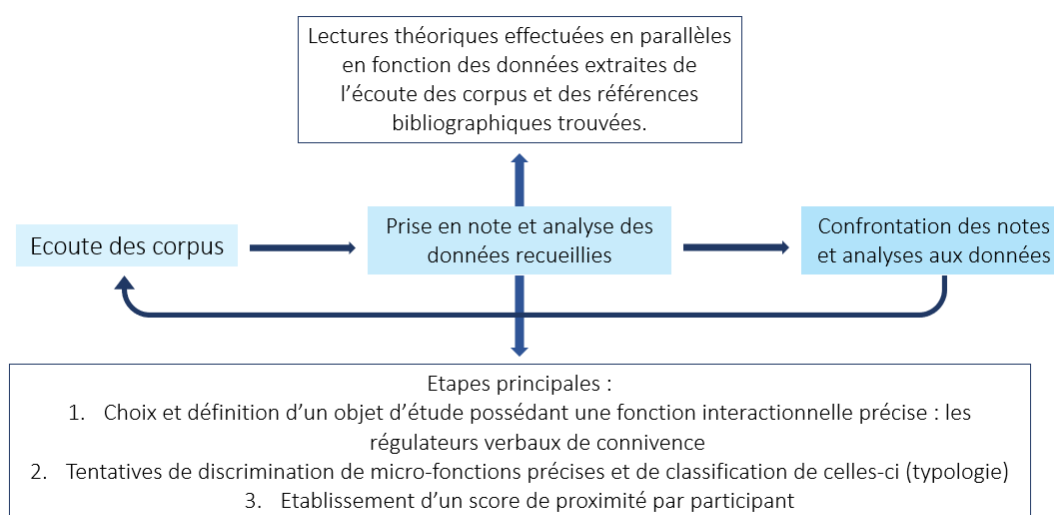
²⁹ Concernant les autres tranches d'âge, 4 locuteurs ont entre 30 et 49 ans, 2 locuteurs ont plus de 50 ans, et l'âge de 4 locuteurs est inconnu.

à face. Cet objectif nous a conduit à construire notre corpus à partir de données extraites de trois corpus différents : le corpus CLAPI, le corpus MPF et le corpus TCOF. Ce corpus a permis par la suite d'élaborer la typologie développée dans ce mémoire, avec une méthodologie que nous allons détailler dans le chapitre suivant.

Chapitre 6. Élaboration de la typologie

La typologie élaborée dans le cadre de ce mémoire, et présentée dans sa version la plus récente dans le chapitre 4, est le fruit d'un long travail analytique découlant d'une approche initialement onomasiologique et *corpus-driven* des données (Tognini-Bonelli, 2001). La typologie, et la définition même des *régulateurs verbaux de connivence*, ont ainsi pu être affinées au gré d'un processus qu'il s'agira désormais de détailler.

Comme cela a pu être précisé précédemment, le présent mémoire ne s'est initialement pas développé à partir d'un objet d'étude précis, mais bien à partir d'une approche méthodologique et d'un choix de langue (la langue vernaculaire) précis. Le choix méthodologique phare de ce mémoire était de se laisser guider par les données pour choisir un objet d'étude, tout d'abord, puis pour l'affiner au fur et à mesure des écoutes. Un deuxième choix méthodologique clé a été de choisir une approche onomasiologique des données, c'est-à-dire une approche partant d'une fonction linguistique pour en analyser les formes (et micro-fonctions dans le cas présent), plutôt qu'une approche sémasiologique partant d'une forme pour en analyser les fonctions. Ces deux choix méthodologiques justifient le processus d'analyse adopté pour ce mémoire : un processus répétitif, alternant entre les étapes suivantes :



Aziliz Philippot – Université Grenoble Alpes | Mémoire de Master 2 – 2020-2021

Figure 4 Schéma des étapes du processus d'analyse des données du corpus, en vue de l'élaboration de la typologie

Il est à noter, en lien avec le chapitre 5, que le corpus a aussi été affiné et augmenté au fur et à mesure de l'avancée analytique. L'idée était de confronter les analyses effectuées non seulement aux corpus déjà analysés, mais aussi à des corpus inconnus. Le fait de tester les définitions avancées dans l'analyse sur de nouveaux corpus permettait en effet de vérifier si elles pouvaient être appliquées dans leur état sur des données nouvelles, mais aussi de renouveler notre approche des données et de minimiser l'influence des analyses déjà effectuées sur notre jugement. Notre approche des données est ainsi passée d'une approche *corpus-driven* uniquement à une approche mixte *corpus-driven* et *corpus-based*, étant donné que nos hypothèses étaient testées sur le corpus au fur et à mesure.

La première étape avant l'élaboration de la typologie a été de sélectionner un phénomène linguistique sur lequel travailler. Après plusieurs écoutes, notre intérêt s'est porté sur les phénomènes lexicaux permettant à une personne ne détenant pas la parole de manifester son écoute et de laisser son interlocuteur continuer de parler, c'est-à-dire aux régulateurs verbaux de connivence tels que nous les avons décrits dans le chapitre 3. 3.

Une fois le phénomène linguistique étudié bien délimité, l'élaboration de la typologie a pu être effectuée suivant ce même processus d'écoute, analyse, et retour à l'écoute. De nombreuses versions de la typologie ont pu être dressées, d'abord très générales, puis plus détaillées.

Étant donné l'approche onomasiologique de ce travail et le souhait de tisser une typologie fonctionnelle, et non formelle, des RVC, un effort particulier a été produit afin de ne pas se laisser entraîner par les formes des RVC³⁰. En plus de l'écoute attentive et répétée des corpus, plusieurs procédures ont été mises en place pour éviter cet écueil, mais aussi pour garantir que les segments analysés avaient bien le statut de régulateurs et non pas de tours de parole.

- Un système de double valeur a été mis en place pour les cas où la fonction précise du RVC était trop difficile à déterminer : deux valeurs différentes (par exemple marque de confirmation et marque d'évaluation (expressif) étaient alors assignées au même segment. Ces doubles valeurs ont permis de rendre compte à la fois des

³⁰ Voir chapitre 4. 4. Pour le détail sur le statut de non-tour des régulateurs, voir chapitre 3. 1.

cas où la forme ne semblait pas corrélérer avec la fonction, mais aussi des cas où la qualité expressive ou non expressive du RVC (et donc sa fonction précise) ne pouvait pas être déterminée. Le but de ce système était donc de réduire ou du moins de proposer une alternative à l'influence de la forme et du sémantisme des RVC sur leur compréhension, mais aussi de rendre compte de la multifonctionnalité de ces segments. Ce système a été mis en place après l'accord inter-annotateurs effectué sur un extrait du corpus, comme nous allons le voir.

- La méthode du *next-turn proof-procedure* (Eiswirth 2020a et 2020b, méthode tirée de l'analyse conversationnelle) a aussi été utilisée tout au long de l'analyse du corpus. La *next-turn proof-procedure* consiste à toujours prendre en compte le tour suivant le segment étudié afin de garantir que l'interprétation du/de la chercheur.se ne diverge pas de l'interprétation en contexte (telle que manifestée par le tour suivant) du segment. Cette méthode s'est révélée particulièrement utile pour faire la différence entre les RVC et les interventions ayant le statut de tour de parole. Nous avons en effet considéré que les RVC n'ayant pas un statut de tour de parole, ils n'étaient pas ratifiés et/ou repris dans le tour suivant.

Après de nombreuses répétitions du processus présenté dans la figure 4, la typologie dégagée semblait prête à être testée dans ce travail de mémoire. Une dernière étape était cependant nécessaire avant de l'utiliser pour annoter le corpus : l'éprouver dans le cadre d'un accord inter-annotateurs.

Lors d'un accord inter-annotateurs, le(s) chercheur(s) ayant établi la typologie à tester (ou autre) ainsi que des individus (généralement d'autres chercheurs mais pas nécessairement) n'ayant pas travaillé sur cette typologie réalisent une même tâche d'annotation à l'aide de la typologie en question (Arstein & Poesio, 2008). Cet exercice permet de vérifier que la typologie établie est bien compréhensible pour des personnes autres que la/les personne(s) l'ayant élaborée, c'est-à-dire qu'elle est bien formulée, avec des critères compréhensibles et qui permettent effectivement de discriminer les catégories dessinées entre elles. La réussite de cet exercice garantit donc la reproductivité de la recherche et sa non-ambiguïté.

L'accord inter-annotateurs a été effectué par deux personnes en plus de l'autrice de ce mémoire : une chercheuse ainsi qu'un doctorant, tous deux travaillant dans le domaine

des sciences du langage. Ce travail d'annotation a été effectué sur quatre extraits du corpus, pour un total de 20 minutes de parole environ et 31 cas présegmentés sur ELAN à annoter³¹.

Une fois les résultats de cet accord inter-annotateurs obtenus, ils ont été compilés dans un fichier excel afin de calculer le score d'accord entre les participants (score γ , établi avec le logiciel GammaSoftware 2.0 (Mathet, Widlöcher & Métivier, 2015)). Ce score γ permet de calculer le taux d'accord entre les participants. On considère que l'accord est suffisant lorsque le score atteint plus de 0.80 γ . Dans notre cas, le score final était de 0.441 γ , score insuffisant pour considérer que la typologie était suffisamment détaillée et claire.

L'examen détaillé des résultats de l'accord inter-annotateurs a révélé que les problèmes se situaient particulièrement au niveau de l'identification et parfois de la différenciation des deux sous-types de RVC expressifs (les marques de convergence et marques d'évaluation), ainsi qu'au niveau de la différenciation entre les RVC et d'autres types de régulateurs non analysés. Afin de régler ces problèmes, différents éléments ont été modifiés dans notre typologie : tout d'abord le système de doubles valeurs pour les cas ambigus a été mis en place, ensuite la définition des RVC a été revue pour insister sur leur caractère de régulateurs simples, ce qui n'était pas précisé dans le fichier original pour l'accord inter-annotateurs³². Pour finir, la définition du type fonctionnel expressif et de ses sous-types a été revue et détaillée, en ajoutant la mention de la prise en compte du sémantisme notamment, au lieu de la prosodie seule, pour juger de l'expressivité ou non d'un segment.

Un deuxième accord inter-annotateurs prenant en compte ces modifications n'a cependant pas pu être effectué par manque de temps. La typologie présentée dans ce mémoire est donc seulement expérimentale, et mériterait d'être testée de nouveau et modifiée si besoin à l'avenir. Malgré ce caractère expérimental, nous avons décidé de continuer avec cette typologie, avec les modifications apportées suite à l'accord inter-annotateurs. En effet, le travail proposé dans ce mémoire était voué à être exploratoire

³¹ Ces cas pouvant être des régulateurs verbaux de connivence mais aussi des tours de parole au contenu lexical pouvant correspondre à un RVC. Le choix d'intégrer ce deuxième cas a été fait afin de vérifier si la définition même des RVC était compréhensible et applicable.

³² Versus les régulateurs complexes, voir chapitre 3 parties 1 et 3.

uniquement, il nous semblait qu'utiliser notre typologie, même dans sa version non-définitive, nous permettrait non seulement de proposer quelques observations générales sur la variation des régulateurs verbaux de connivence, mais aussi de tester et de critiquer au besoin la typologie établie.

Pour résumer, le processus de va-et-vient précisé dans la figure 4 nous a permis de réaliser différentes étapes avant d'arriver à la version finale de la typologie présentée dans ce mémoire : l'identification d'un objet d'étude précis et la mise en place d'une typologie. Un accord inter-annotateur effectué à l'aide de la typologie sur une partie du corpus a de plus permis de tester notre travail, qui s'est avéré insuffisamment clair. Quelques modifications dans la définition des *régulateurs verbaux de connivence* et dans la typologie ont été mises en place à la suite de cette expérience, modifications qui n'ont cependant pas pu être testées dans le cadre d'un deuxième accord inter-annotateurs.

Une fois notre typologie expérimentale obtenue, il a été nécessaire de mettre en place des procédures afin de pouvoir annoter les données du corpus, et obtenir des résultats dans un deuxième temps. Le chapitre suivant sera dédié aux détails de ce processus d'annotation.

Chapitre 7. Annotation du corpus

L'annotation du corpus a nécessité plusieurs étapes de travail préalables. Une première étape a été la mise en place d'une grille de score permettant de calculer la position de chaque participant au sein de chaque interaction sur le continuum proximité/distance (Koch & Esterreicher, 2001). En parallèle de cette première étape, il a été nécessaire de réfléchir à une grille d'annotation adaptée à notre objet d'étude d'une part, et aux analyses que nous souhaitons effectuer d'autre part. L'annotation du corpus et l'extraction de données statistiques n'a été possible qu'une fois ces deux étapes préalables réalisées. Il s'agira donc dans un premier temps de détailler (point 1) la méthodologie mise en place pour l'analyse de la proximité dans le corpus. Les détails de l'annotation (mise en place de la grille, annotation en elle-même et statistiques dérivées) seront ensuite expliqués dans la partie 2 de ce mémoire.

1. L'établissement de scores de proximité dans le corpus.

Comme il a été précisé dans le chapitre 2. 2, l'idée d'une dimension conceptionnelle de la parole (opposée à une dimension modale oral vs écrit) sur un continuum allant de la proximité à la distance est tirée des travaux de Koch & Esterreicher (2001). Afin de caractériser les interactions et de les situer sur ce continuum, ces deux chercheurs ont établi une liste de paramètres situationnels à observer. Nous avons souhaité reprendre cette idée dans notre mémoire, en émettant l'hypothèse que la place d'une interaction, voire d'un locuteur donné sur le continuum proximité/distance, pouvait avoir un impact sur le nombre et le type de RVC retrouvés.

Pour tester cette hypothèse, le choix a été fait de reprendre les 10 paramètres situationnels/paramètres de la communication de Koch & Esterreicher comme outil d'analyse. Un score a ainsi été établi pour chaque interlocuteur dans chaque conversation, basé sur le comportement de ce locuteur vis-à-vis des 10 paramètres³³ dans la conversation

³³ Koch & Esterreicher laissent cependant le champ libre à l'existence de plus de paramètres, comme l'indique le « etc. » aussi présent dans le tableau original de 2001 (p. 586)

donnée. Ces 10 paramètres sont présentés ci-dessous (tableau 8) dans leur version de 2008, « plus claire » que la version de 2001 selon les termes de Modicom (2015, p. 4).

Tableau 8 Les paramètres situationnels de Koch & Esterreicher, version de 2008 (p. 201, cité par Modicom, 2015, p. 4)

1) Parole privée	Parole publique (1)
2) Relation de familiarité entre les participants	Les participants ne se connaissent pas (2)
3) Implication émotionnelle	Faible implication émotionnelle (3)
4) Prévalence de la situation d'énonciation	Effacement de la situation d'énonciation (4)
5) Proximité des référents	Eloignement des référents (5)
6) Proximité spatio-temporelle des participants	Distance spatio-temporelle (6)
7) Coopération entre les participants	Pas de coopération entre les participants (7)
8) Dialogisme	Monologisme (8)
9) Discours spontané	Discours réfléchi (9)
10) Développement thématique libre	Thèmes fixés (10)
Etc.	Etc.

Un score individuel a été assigné pour chaque locuteur au sein de chaque conversation, basé sur la somme de son score par paramètre. Chaque paramètre était évalué avec des valeurs allant de -2 à 2 (un score de -2 indiquant un maximum de proximité et 2 un maximum de distance), pour des scores finaux pouvant varier, en théorie, entre -20 et 20.

Il est à noter que, comme le pointe Gadet (2007), tous les paramètres (sauf le 6^{ème}, qui fonctionne en dichotomie) varient selon un continuum, comme on a pu le voir, mais avec différents degrés. Il était en revanche nécessaire dans le cadre de cette étude d'utiliser des degrés discrets et systématiques, et la même échelle (entre -2 et 2 points) a donc été utilisée pour tous les paramètres (sauf le 6^{ème} encore une fois, pour lequel le score peut uniquement être de -2 ou 2).

Plusieurs étapes ont été suivies afin d'établir des scores les plus pertinents possibles : tout d'abord une grille d'analyse a été établie pour chaque paramètre, ainsi qu'une liste de critères linguistiques et extralinguistiques (informations présentes dans les métadonnées) pouvant être mobilisés pour établir le score. Ensuite, une fois une majorité de scores établis, tous les scores ont été revus afin d'obtenir une vue d'ensemble et de garantir la validité des scores sur la totalité du corpus. Les valeurs obtenues sont donc à minima relatives. Il reste tout de même nécessaire de ne pas les prendre comme des valeurs absolues mais bien comme

des indications, du fait entre autres de la complexité des notions abordées dans les paramètres (en particulier pour les paramètres 3, 4, 5 et 10 par exemple).

Ces précisions étant établies, les sections suivantes détailleront les grilles d'analyse par paramètre, ainsi que les critères linguistiques et extralinguistiques observés et la portée du score (individuel, par conversation ou sur tout le corpus).

1.1. Parole privée/publique

En plus de la prise en compte du contexte (face-à-face ou non des interactants, lieu et conscience de l'enregistrement), l'établissement du score pour ce paramètre se base sur les indices linguistiques suivants :

- Utilisation de phatiques tels que « tu vois », « tu sais », etc.
- Termes d'adresse (quantité et qualité).

Tous les locuteurs ont même le score par conversation pour ce paramètre.

Tableau 9 Détails de la grille de score pour le paramètre 01 (parole privée/publique)

-2	Communication destinée aux interlocuteurs eux-mêmes uniquement : seuls les participants entendent la conversation et le contenu est pertinent uniquement pour eux.
-1	Communication majoritairement destinée aux interlocuteurs eux-mêmes : le contenu est principalement pertinent pour eux, mais aussi pour des personnes extérieures qui pourraient entendre la communication ³⁴ .
0	Communication entre des personnes distinctes mais ouvertes à la participation d'autres personnes et/ou avec un contenu pouvant être apprécié par des auditeurs extérieurs (cas des fêtes publiques par exemple > mélange des groupes, etc.)
1	Communication majoritairement dédiée à un public, certains aspects du contenu restent cependant accessibles uniquement à un sous-ensemble donné de personnes et une certaine prise en compte de l'individualité des interactants persiste.
2	Communication entièrement destinée à un public : le contenu est pensé pour le public et l'auditorat est libre.

³⁴ Comme c'est le cas dans certains des enregistrements de notre corpus, dans lesquels les locuteurs fournissent des récits principalement destinés à leurs interlocuteur mais aussi partiellement aux auditeurs du corpus. Ce cas peut aussi s'appliquer aux conversations privées tenues dans la rue ou autre place publique, par exemple.

1.2. Relation de familiarité entre les participants/les participants ne se connaissent pas

Les participants se connaissent déjà dans la très grande majorité des enregistrements de notre corpus, cependant les degrés de familiarité peuvent varier. La valeur de ce paramètre repose à la fois sur les informations présentes dans les métadonnées des corpus et/ou les informations relationnelles évoquées dans les enregistrements, mais aussi sur certains indices linguistiques, tels que :

- La présence de marqueurs discursifs et phraséologismes rappelant l'histoire conversationnelle des participants (par exemple « tu m'avais dit », « je me souviens », « mais oui », etc.).
- Les phénomènes de co-constructions de récits mentionnant des souvenirs (le fait de compléter les phrases de l'interlocuteur ou bien d'ajouter des informations, par exemple).

Le score par locuteur d'une même conversation peut varier uniquement dans les conversations à trois locuteurs.

Tableau 10 Détails de la grille de score pour le paramètre 02 (relation de familiarité entre les participants/les participants ne se connaissent pas)

-2	Les interlocuteurs se connaissent très bien et ont déjà eu des conversations à de très nombreuses occasions.
-1	Les interlocuteurs se connaissent bien et ont déjà eu des conversations à de nombreuses occasions.
0	Les interlocuteurs sont des connaissances et ont déjà partagé quelques conversations.
1	Les interlocuteurs se connaissent très peu et ne se sont quasiment jamais parlé.
2	Les interlocuteurs ne se connaissent pas du tout et ne se sont jamais parlé.

1.3. (Faible) implication émotionnelle

L'émotionnalité, ou implication émotionnelle, est un des paramètres qui varient le plus dans ce corpus, mais aussi un des plus complexes à évaluer. Un faisceau d'indices linguistiques liés à l'affect sont observés pour établir le score individuel pour ce paramètre :

- La prosodie (débit, amplitude, hauteur de voix).
- La présence de modalisateurs tels que les intensifieurs, ou d'intensité dans le discours.

- La présence d'interjections (hors RVC).
- Le thème abordé.
- L'utilisation de vocabulaire lié aux émotions.

(9)

Exemple 9 Extrait d'un passage avec une très forte implication émotionnelle des participants

- 1 -19_MPF_01_01 elle était **folle** je me suis dit non je- je suis obligé d'aller lui laisser ce bonheur (0.5)
- 2 **waouh** (0.6)
- 3 **c'était beau c'était beau** (0.3)
- 4 -19_MPF_01_02 Hum déjà le coup de la glace là avec les autres **c'était bon c'était drôle** (1.6)
- 5 Mais forcément une petite minette qui vient regarder un match de foot toute seule
comme [ça]
- 6 -19_MPF_01_01 [en robe en robe] ((**rires**))
- 7 -19_MPF_01_02 En robe [lunettes de soleil et tout ça ((**rires**))]

Ce score est individuel.

Tableau 11 Détails de la grille de score pour le paramètre 03 ((faible) implication émotionnelle)

-2	Les locuteurs sont impliqués émotionnellement dans la conversation pour toute sa durée ou presque (indices d'affect omniprésents).
-1	Les locuteurs sont impliqués émotionnellement dans la conversation de manière fréquente (indices d'affect fréquents).
0	Les locuteurs sont moyennement impliqués émotionnellement dans la conversation (quelques indices d'affect).
1	Les locuteurs sont peu impliqués émotionnellement dans la conversation (peu d'indices d'affect).
2	Les locuteurs ne sont pas du tout ou très peu impliqués émotionnellement dans la conversation (pas du tout/très peu d'indices d'affect).

1.4. Prévalence/effacement de la situation d'énonciation

Ce paramètre s'intéresse à la façon dont la situation d'énonciation est référencée dans la conversation. Afin de définir une situation d'énonciation, il est nécessaire de s'intéresser aux éléments suivants : qui parle, à qui, où et quand. Ainsi plus les locuteurs dans une conversation feront des références à ces différents éléments, plus ils seront ancrés dans leur situation d'énonciation. L'établissement du score de proximité ou de distance pour ce paramètre s'effectue ainsi sur la base d'un faisceau d'indices linguistiques, tels que :

- Quantité de termes d'adresse et/ou références aux interlocuteurs présents.
- Comportement adéquat et engagement dans l'activité en cours.

Ce score est individuel.

Tableau 12 Détails de la grille de score pour le paramètre 04 (prévalence/effacement de la situation d'énonciation)

-2	Les participants sont totalement immergés dans la situation d'énonciation, et utilisent de nombreux moyens linguistiques pour actualiser cette situation d'énonciation tout au long de la conversation.
-1	Les participants s'inscrivent dans la situation d'énonciation et l'actualisent régulièrement pendant la conversation.
0	Les participants s'inscrivent dans la situation d'énonciation mais ne l'actualisent que de temps en temps et semblent ainsi s'en détacher parfois.
1	Les participants sont relativement détachés de la situation d'énonciation, et utilisent peu de moyens linguistiques pour l'actualiser.
2	Les participants sont détachés de la situation d'énonciation et n'utilisent pas ou très peu de moyens linguistiques pour l'actualiser.

1.5. Proximité/éloignement des référents

La valeur de ce paramètre semble dépendre moins de la quantité d'indices linguistiques référentiels que de leur qualité. Il est donc nécessaire, pour établir le score de proximité ou de distance pour ce paramètre, de s'intéresser au type d'indices linguistiques suivants :

- Référence à des personnes, objets, etc. présents ou absents et / ou référence à des événements actuels ou distants.
- Utilisation de déictiques.

Ce score est individuel.

Tableau 13 Détails de la grille de score pour le paramètre 05 (proximité/éloignement des référents)

-2	Toutes les références concernent des référents proches, qu'ils soient concrètement présents ou bien connus personnellement de tous les participants.
-1	La majeure partie des références concernent des référents proches.
0	Certaines références concernent des référents proches, mais d'autres concernent aussi des référents plus éloignés, soit d'un point de vue spatio-temporel soit d'un point de vue personnel (personnes / situations inconnues ou abstraites).
1	La majeure partie des références concernent des référents éloignés.
2	Toutes les références concernent des référents éloignés.

1.6. Proximité/distance spatio-temporelle des participants

Ce paramètre est le seul paramètre qui fonctionne en dichotomie (Koch & Esterreicher, 2001). Toutes les conversations de notre corpus ayant lieu en face à face, le score est de -2 (proximité maximale) sur tout le corpus pour ce paramètre.

Tableau 14 Détails de la grille de score pour le paramètre 06 (proximité/distance spatio-temporelle)

-2	Co-présence spatio-temporelle.
2	Séparation spatio-temporelle.

1.7. Présence de coopération entre les participants

La valeur du score de proximité ou de distance pour ce paramètre est établie en se basant sur les indices linguistiques suivants :

- Présence ou non de pauses et/ou ruptures dans le flux de la conversation.
- Présence ou non de chevauchements³⁵.
- Dispute ou non de la prise de parole.

Ce score est individuel.

Tableau 15 Détails de la grille de score pour le paramètre 07 (coopération/pas de coopération entre les participants)

-2	La répartition de la parole entre les participants ne présente aucun problème, et s'effectue de manière fluide et/ou sans négociation.
-1	La répartition de la parole entre les participants présente peu de problèmes, et s'effectue généralement de manière fluide et/ou sans négociation.
0	La répartition de la parole entre les participants peu poser quelques problèmes, avec des moments de pauses dans la conversation et/ou de négociation de la parole.
1	La répartition de la parole entre les participants pose beaucoup de problèmes, et les moments de pauses dans la conversation et/ou de négociation de la parole sont fréquents.
2	La répartition de la parole entre les participants est problématique sur toute la conversation, avec de très nombreux moments de pauses dans la conversation et/ou de négociation de la parole.

³⁵ Pour rappel l'émission d'un régulateur par un récepteur pendant le discours d'un locuteur n'est pas considérée comme un cas de chevauchement, du fait du statut de non-tour du régulateur (voir chapitre 3).

1.8. Dialogisme/monologisme

La valeur de ce paramètre dépend principalement des indices linguistiques et interactionnels suivants :

- Temps de parole équivalent/monopolisation du temps de parole par un participant
- Co-construction de l'interaction/interaction monogérée
- Rôle variables/rôles fixes

Tous les locuteurs ont même le score par conversation pour ce paramètre³⁶.

Tableau 16 Détails de la grille de score pour le paramètre 08 (dialogisme/monologisme)

-2	L'interaction est entièrement dialogale et co-gérée : les participants ont des temps de parole équivalents et des rôles variables.
-1	L'interaction est principalement dialogale et co-gérée.
0	L'interaction est partiellement dialogale et co-gérée, mais contient aussi des moments plutôt monogérés ou un seul participant garde la parole, avec une fixation des rôles.
1	L'interaction est principalement monogérée.
2	L'interaction est entièrement monogérée.

1.9. Discours spontané/réfléchi

Il s'agit d'un des paramètres dont la valeur est la plus complexe à établir. En effet, les métadonnées n'indiquent que très rarement le degré de préparation des enregistrements, et même lorsque l'information est présente, il est difficile de savoir exactement à quel point les interlocuteurs ont réfléchi à leur discours. L'hypercorrection, qui est un des indices linguistiques permettant d'établir un degré moindre de spontanéité, est aussi difficile à établir. En effet, les enregistrements sont souvent courts, ce qui ne permet pas de déterminer si les locuteurs modifient ou non leur style habituel. Il est néanmoins possible d'établir une

³⁶ A l'exception de la conversation intitulée « Aristide5b » du corpus MPF, dans laquelle le troisième participant arrive après les autres, et s'inscrit donc dans une dynamique différente. Ce fait est aussi valable pour le paramètre 10.

valeur indicative en se basant sur les critères suivants, et sur les informations présentes dans les métadonnées et/ou déductibles du contexte linguistique et extralinguistique :

- Conscientisation de la communication : communication prévue à l’avance ou bien entièrement spontanée.
- Dans le cadre d’une communication prévue à l’avance : présence ou non d’un script et importance de ce script.
- Formes utilisées/stylistique : présence ou non d’hyper-correction de la part des participants.

Ce score est individuel.

Tableau 17 Détails de la grille de score pour le paramètre 09 (discours spontané/réfléchi)

-2	La communication est entièrement spontanée : les participants s’expriment librement sans hyper-correction et/ou le contenu de l’interaction n’avait pas été fixé d’avance.
-1	La communication est majoritairement spontanée : les participants s’expriment librement avec peu d’hyper-correction et/ou seule une proportion mineure de l’interaction avait été fixée d’avance.
0	La communication est moyennement spontanée : les participants peuvent user de l’hyper-correction et contrôler leur discours et/ou le contenu de l’interaction était en partie fixé d’avance.
1	La communication est majoritairement préparée : les participants s’expriment globalement de façon contrôlée et réfléchie et/ou une proportion majeure de l’interaction avait été fixée d’avance.
2	La communication est entièrement préparée : les participants contrôlent tout leur discours et/ou l’entièreté de l’interaction avait été fixée d’avance.

1.10. Développement thématique libre/thèmes fixés

En théorie, l’évaluation de ce paramètre s’effectue sur la base de la fluidité des thèmes abordés, à la fois au niveau de la quantité globale de thèmes et au niveau du temps passé à discuter chaque thème. En pratique cependant cette tâche n’est pas si aisée, car la notion de thème est assez abstraite. Pour avoir des critères plus discrets et donc plus simples à quantifier, le choix a été fait de déterminer le score de proximité ou de distance pour ce paramètre en se basant sur :

- Le temps de parole alloué à un thème/aspects d’un thème : les participants discutent-ils pendant longtemps ou passent-ils assez rapidement d’un thème à un

autre, et développent-ils rapidement ou non les différents aspects possibles des thèmes abordés ?

- Le type de marqueurs de structuration du discours utilisés (présence ou non de marqueurs de rupture de thèmes tels que « et du coup » « et sinon », etc., par exemple)

Tous les locuteurs ont même le score par conversation pour ce paramètre.

Tableau 18 Détails de la grille de score pour le paramètre 10 (développement thématique libre/fixe)

-2	Le développement thématique est fluide : de nombreux thèmes différents sont abordés et/ou les participants rebondissent régulièrement sur différents aspects d'un même thème.
-1	Le développement thématique est majoritairement fluide : des thèmes différents sont abordés et/ou les participants rebondissent sur différents aspects d'un même thème.
0	Le développement thématique est plutôt fluide, cependant une certaine fixation des thèmes est aussi observée : quelques thèmes différents sont abordés et/ou les participants rebondissent de temps en temps sur différents aspects d'un même thème.
1	Les thèmes abordés sont majoritairement fixés : il y a peu de changements de thèmes et/ou peu de rebondissements sur différents aspects d'un même thème.
2	Les thèmes abordés sont entièrement fixés : un thème est discuté sans interruption et sans rebondissements jusqu'à sa fin, et uniquement une fois ce thème terminé un autre thème est abordé de la même manière.

Les scores par locuteurs ont été établis à l'aide d'un tableur excel. Une fois ces scores établis, un système de codage des noms des participants a été mis en place afin de faciliter le travail d'analyse. Le codage mis en place est le suivant :

(-)XX¹_CORPUS_XX²_XX³

Le premier chiffre ((-)XX¹) correspond au score de proximité, le deuxième (XX²) à l'indice de la conversation au sein du corpus, et le troisième (XX³) à l'indice du locuteur au sein de cette conversation. Les indices des conversations et des locuteurs ont été établis en suivant l'ordre alphabétique des noms originaux (pour les conversations) et l'ordre d'apparition des locuteurs (pour les locuteurs). Ainsi si l'on prend le premier locuteur (c'est-à-dire le premier à avoir pris la parole) de la première conversation (selon l'ordre alphabétique) du corpus CLAPI, sachant que ce locuteur possède un score de proximité de -16, on obtient le codage suivant : -16_CLAPI_01_01.

Ce code a été utilisé pour tous les traitements de données ainsi que pour les exemples dans ce mémoire. Le tableau contenant les codes et le nom des conversations et locuteur originaux est disponible en annexe 1, p. 129.

En parallèle de ce traitement des scores de proximité par participant, une partie très importante du travail sur les données a été effectuée avec le logiciel ELAN. Les tâches réalisées avec ce logiciel ont été les suivantes : élaboration d'une grille d'annotation, annotation en elle-même et extraction de données statistiques. La façon dont ces différentes tâches ont été effectuées sera détaillée ci-dessous.

2. Traitement des données sous ELAN

Le logiciel ELAN permet de travailler sur de l'annotation de fichiers vidéo ou audio. La flexibilité du logiciel permet d'obtenir des annotations complexes, et surtout de les adapter aux besoins de la recherche en cours. Il est cependant nécessaire de bien se familiariser avec le logiciel et de tester plusieurs techniques d'annotation avant d'arriver à un résultat satisfaisant, comme nous allons le voir. Après l'obtention d'un système d'annotation adéquat, nos corpus ont pu être annotés de manière systématique, ce qui nous a permis d'obtenir des résultats statistiques.

2.1. Mise en place d'une grille d'annotation

La grille d'annotation établie dans le cadre de ce mémoire a été élaborée petit à petit, et améliorée plusieurs fois à la suite d'une meilleure compréhension et prise en main du logiciel ELAN d'une part et de l'avancée de la typologie d'autre part. La grille d'annotation présentée ici est la version finale qui a servi à l'annotation et l'extraction des données pour l'analyse.

Le choix a été fait d'annoter les RVC par locuteur, afin de pouvoir traiter les données individuelles. De ce fait, trois acteurs « enfants » (c'est-à-dire dépendant hiérarchiquement d'un acteur « parent », en l'occurrence les locuteurs) ont été créés pour chaque locuteur, correspondant aux trois types majeurs de RVC : les marques d'enregistrements, les marques de confirmation et les expressifs. Un acteur « enfant » intitulé Sous-type expressif.NOM³⁷ a

³⁷ NOM correspondant au nom original du locuteur.

aussi été créé pour les expressifs, afin de pouvoir renseigner la fonction exacte du RVC annoté, entre marque de convergence et marque d'évaluation. La capture d'écran ci-dessous présente ces différents acteurs et leurs relations hiérarchiques :

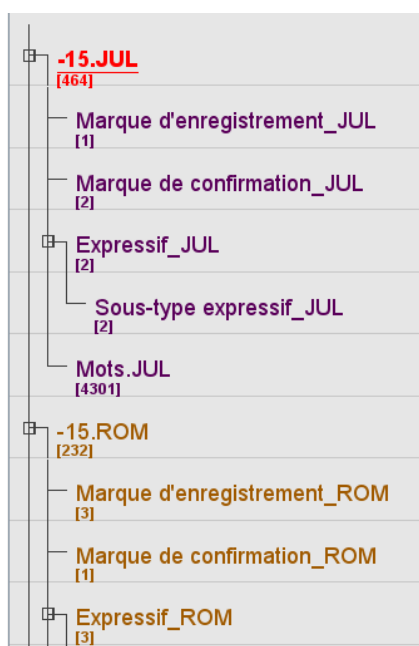


Figure 5 Acteurs de la grille d'annotation selon leur relation hiérarchique (extrait du fichier ELAN traitant le corpus apéritif entre ami(es) - pois de CLAPI)

Afin de faciliter le traitement des RVC annotés et de les différencier de la transcription originale, des types et stéréotypes particuliers ont été appliqués aux acteurs sous lesquels les RVC ont été annotés. Les types permettent de préciser les relations entre les acteurs (au moyen des stéréotypes), mais aussi de permettre l'utilisation d'un Vocabulaire Contrôlé, sur lequel nous reviendrons sous peu.

Un premier type nommé « RVC » a été assigné aux acteurs Marque d'enregistrement, Marque de confirmation et Expressifs. Ce type a été associé au stéréotype « included in » (ou inclus dans), qui empêche toute annotation créée sous l'acteur enfant de dépasser les frontières de l'annotation d'origine sous l'acteur parent³⁸.

Ce fonctionnement a permis de réduire le temps passé à établir les frontières des segments annotés, puisque la majorité des corpus proposaient déjà une annotation originale à la seconde près. Les RVC apparaissent dans leur grande majorité entre deux tours de parole

³⁸ Les annotations enfants peuvent en revanche être plus courtes que les annotations parents.

d'un interlocuteur (ou d'un Locuteur, face au Récepteur), la segmentation en tant que RVC suivait généralement cette segmentation originale. Pour certains corpus en revanche, en particulier les corpus tirés du MPF et de TCOF, la transcription originale ne rendait pas compte des phénomènes de superposition de la parole (deux acteurs parlant en même temps), ou alors ceux-ci étaient signalés par des signes orthographiques tels que < > dans le TCOF par exemple. Ces corpus ont donc dû être légèrement modifiés au niveau des RVC, afin d'en déplacer les limites temporelles pour faire apparaître la superposition de la parole des interactants. Une fois les annotations segmentées pour les acteurs Marque d'enregistrement, Marque de confirmation et Expressifs (types RVC), leur valeur lexicale a pu être utilisée pour remplir l'annotation. Dans le cas d'une double valeur, le segment était annoté sur les deux lignes correspondantes (par exemple sous Marque d'enregistrement et sous Expressif), et la même valeur lexicale était indiquée dans les deux cas.

Un second type a été créé pour l'acteur Sous-type expressif, intitulé Sous-type de RVC. Le stéréotype « Symbolic association » (ou association symbolique), qui permet de créer des annotations enfants strictement de la même durée que les annotations parents, a été utilisé dans ce cas. Un Vocabulaire Contrôlé intitulé « RVC Expressifs » a été associé à ce type de piste, afin d'automatiser l'annotation des sous-types d'expressifs. Les entrées de ce Vocabulaire Contrôlé étaient les suivantes³⁹ :

- Marque de convergence
- Marque d'évaluation – forme composée
- Marque d'évaluation – interjection
- Marque de convergence et/ou d'évaluation⁴⁰

La valeur lexicale du type expressif étant déjà indiquée dans l'annotation parent « Expressif », seul ce VC a été utilisé pour remplir les annotations une fois obtenues⁴¹.

³⁹ Les définitions de ces entrées ainsi que du nom général « RVC Expressifs » correspondent aux définitions des ces types, données au chapitre 4. 4.

⁴⁰ En cas de double valeur.

⁴¹ Les annotations associées à un stéréotype « symbolic association » devait être créées à partir du menu Acteurs > Create annotation on dependent tiers.

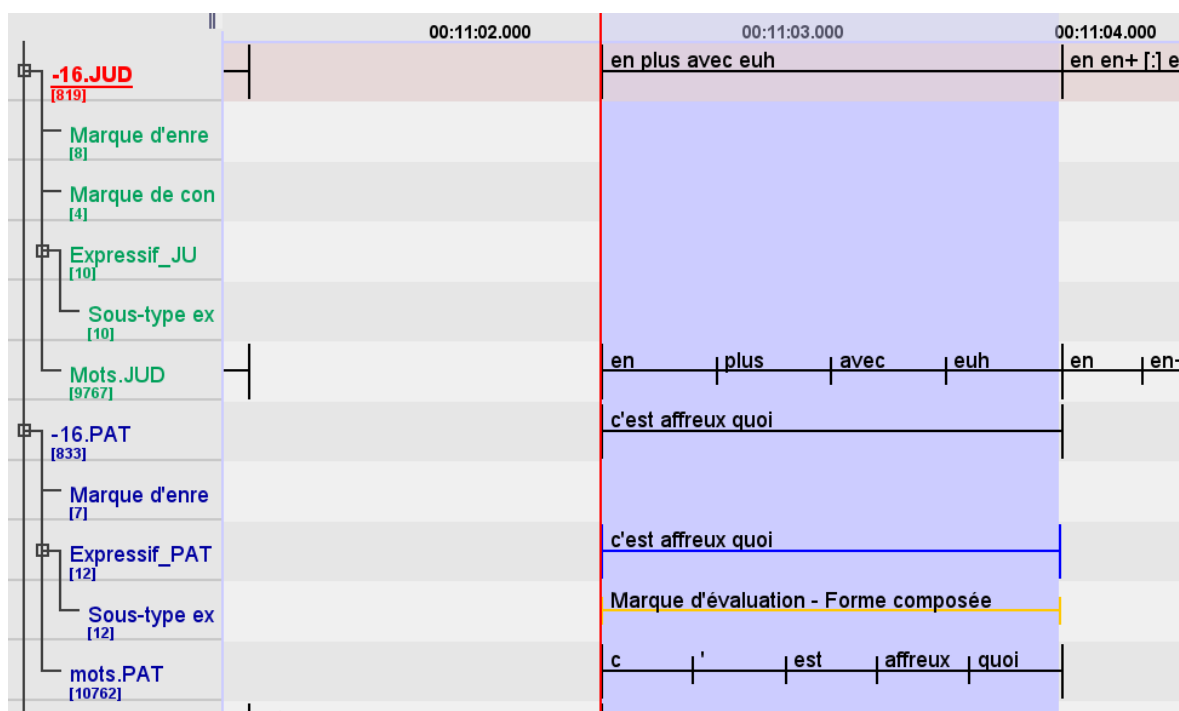


Figure 6 Annotation finalisée d'un RVC expressif de type marque d'évaluation (extrait du fichier ELAN traitant le corpus apéritif entre ami(e)s – glasgow de CLAPI)

Ce fonctionnement a permis d'obtenir des données concernant la variation individuelle (par locuteur) des participants, des données concernant la variation selon le degré de proximité (indiqué par locuteur encore une fois), mais aussi des données générales sur les types fonctionnels retrouvés dans le corpus. De plus, la création d'un acteur par type fonctionnel, associée à une annotation du contenu lexical des RVC trouvés, a rendu possible des analyses basées à la fois sur les fonctions et sur les formes des RVC.

Les détails de cette grille d'annotation étant établis, passons désormais à la description des statistiques et données obtenues une fois toutes les annotations effectuées.

2.2. Statistiques et données obtenues sous ELAN

Afin de pouvoir traiter les annotations obtenues, nous les avons tout d'abord exportées du logiciel afin de les reprendre dans un tableur excel, plus aisé à manipuler et modifiable (voir chapitre 8). Afin d'exporter ces données, nous sommes passés par le menu Fichier > Multiple File Processing > Statistics for multiple files. Cela nous a permis d'obtenir une interface de statistiques compilant tous les fichiers.

Tous les acteurs correspondant à nos RVC (marques d'enregistrement, marques de confirmation, expressifs et sous-types d'expressifs) ont ensuite été sélectionnés avant de mettre à jour les statistiques, et le résultat obtenu a été enregistré au format .txt.

Afin de calculer la fréquence des RVC dans le corpus, il était nécessaire de connaître le nombre de mots par participants. Cette information n'étant pas disponible dans les métadonnées des corpus extraits du MPF et de CLAPI, une autre fonction de ELAN a été utilisée afin d'obtenir cette information. Ainsi les mots par participant ont été extraits automatiquement sur une ligne dédiée (intitulée Mots.NOM⁴²) grâce à la fonction Tokenizer l'acteur dans le menu Acteur. Un inconvénient de cette fonction cependant était qu'elle permettait de séparer les mots en se basant sur les espaces, et considérant donc certaines ponctuations (l'apostrophe notamment) comme tokens distincts. Les apostrophes et autres informations non verbales telles que les temps de pause ont ainsi dû être retirées du total de tokens pour obtenir le véritable nombre de mots⁴³ par locuteur. En revanche les productions vocales telles que *hum, oh, ah*, etc., n'ont pas été retirées du nombre total de mots (la tâche aurait été trop chronophage pour ce mémoire). Ces productions vocales comptent par conséquent comme des « mots » dans notre analyse, y compris dans le décompte des mots formant des régulateurs verbaux de connivence. Par exemple, le RVC *d'accord* est composé de deux « mots » avec ce système, et le RVC *ah c'est chaud* de quatre mots.

Le reste du traitement des données a été effectué sur des tableurs excels ou bien sous forme de prise de notes pour l'analyse qualitative, comme il sera expliqué dans le chapitre suivant.

⁴² NOM correspondant encore une fois au nom des participants.

⁴³ Ce nombre de mots reste cependant approximatif, les calculs ayant été faits à la calculatrice sans automatisation possible.

Chapitre 8. Traitement des données obtenues

La première étape du traitement des données a été la création de plusieurs fichiers excel, établis à partir des données obtenues suite à l'annotation du corpus. Ces fichiers ont ensuite été enrichis afin de permettre la réalisation de calculs statistiques et l'extraction de tableaux de données selon divers critères, tels que la proximité par exemple.

Les données recueillies dans le corpus ont permis de créer un premier fichier excel. Celui-ci ne contenait cependant pas toutes les informations nécessaires pour les analyses que nous avons en tête, et présentait de plus quelques catégories superflues dans notre cas (durée de l'annotation en seconde par exemple). En effet, les analyses que nous souhaitions effectuer étaient les suivantes :

- Calcul de la fréquence des RVC de manière générale et par type fonctionnel.
- Travail sur les formes (contenu lexical et structures syntaxiques) des RVC selon différents critères (type fonctionnel, score de proximité).
- Calcul de la fréquence des RVC (général et par type fonctionnel encore une fois) selon le score de proximité des locuteurs, avec une prise en compte non seulement du score total mais aussi du score pour certains paramètres particuliers.
- Travail sur les doubles valeurs présentes dans le corpus.

Au sujet de ce dernier critère, il paraissait aussi nécessaire de pouvoir traiter ces doubles valeurs de deux manières différentes : en les comptant une seule fois comme des cas distincts, ou en les comptant deux fois (une fois par type fonctionnel représenté).

Pour remplir ces objectifs, quatre fichiers excel distincts ont été créés :

- 1- Un fichier renseignant, par annotation, son type, son contenu, le nombre d'occurrences de cette annotation dans le corpus, son association ou non à une double valeur⁴⁴, sa structure syntaxique et l'unité lexicale centrale sur lequel ce

⁴⁴ Un « Oui » dans cette colonne indiquant la présence d'une double valeur, avec l'indication du nombre d'annotation concernées au besoin. Par exemple, 4 *cool* sont évalués à la fois en tant que marques de

RVC est construit. La manipulation des données de ce fichier a permis d'extraire des tableaux indiquant les structures syntaxiques ainsi que les lexèmes retrouvés selon les différents types fonctionnels de RVC.

- 2- Une deuxième version de ce premier fichier nous a permis de travailler sur les formes uniquement, sans les doubles valeurs. Pour ce faire, les formes pour lesquelles une double valeur existait ont été identifiées et synthétisées sur une seule ligne pour supprimer les répétitions. La manipulation de ce deuxième fichier a rendu possible l'extraction de tableaux détaillant les formes des RVC selon le nombre d'unités linguistiques les composant et selon leur type fonctionnel, ainsi que d'informations sur le nombre de formes uniques retrouvées dans le corpus.
- 3- Un troisième tableur renseignait les annotations par locuteur (une ligne par annotation), ainsi que des informations sur le total de RVC produits pour chaque locuteur (nombre de RVC, nombre de « mots » les composant, nombre de mots total, pourcentage des mots formants des RVC dans le discours du locuteur). Afin de traiter le paramètre de la proximité, le nom codé des participants a été utilisé, et deux colonnes contenant les scores de paramètres précis (08 + 10 et 02) ont été ajoutées à ce tableur.
- 4- Un dernier tableur reprenait le contenu lexical des RVC, par type fonctionnel, selon le score de proximité auquel ils étaient associés : score faible (de -02 à -05 points), assez faible (de -06 à -10 points), assez fort (de -11 à -15 points) ou fort (de -16 à -19 points). Ce tableur a permis d'obtenir des données pour l'analyse du contenu lexical selon le taux de proximité dans le corpus.

En parallèle avec ce travail informatisé, quelques observations ont pu être notées tout au long de notre travail d'analyse du corpus dans des carnets papiers. Ces observations ont notamment permis de sélectionner des exemples de cas spécifiques afin d'illustrer nos propos, exemples qui ont ensuite été informatisés.

confirmation et en tant que marques d'évaluation (expressifs) dans le corpus. L'annotation *cool* en tant que marque de confirmation et en tant que marque d'évaluation était donc associée à la mention « 4 Oui ».

Pour résumer, plusieurs fichiers excel ont été établis sur la base des données extraites de l'annotation du corpus, fichiers modifiés et enrichis selon les besoins pour permettre d'effectuer différentes analyses. En plus de ce traitement informatique des données issues du corpus, un travail de prise en note papier avait aussi pu être réalisé tout au long du travail d'analyse, dans le but de pouvoir revenir en détail sur les données.

Ces différentes méthodes de traitement des données ont rendu possible l'extraction des résultats et exemples qui feront l'objet de la prochaine et dernière partie de ce mémoire, dédiée à l'analyse des résultats et à leur discussion. Cette partie sera composée de deux chapitres. Un premier chapitre (chapitre 9) se centrera sur les différents résultats obtenus après l'analyse des données du corpus (observations générales, usages selon certains paramètres de proximité, inventaire et des formes observées, etc.). Le deuxième chapitre (chapitre 10) sera consacré à la discussion de ces résultats mais aussi du mémoire de manière générale, dans le but notamment d'en souligner les limites.

Partie 3

-

Analyse et discussion des résultats

Chapitre 9. Résultats de l'analyse du corpus

En suivant la méthodologie établie dans la partie précédente de ce mémoire (partie 2), une analyse à la fois quantitative et parfois plus qualitative a pu être effectuée sur les données issues de l'annotation de notre corpus. En termes de fréquence des RVC dans le corpus, cette analyse a donné des résultats généraux tous corpus confondus (point 1 du chapitre), mais aussi des résultats plus spécifiques selon le degré de proximité des interactants (point 2). Plusieurs inventaires des formes des RVC annotés dans le corpus ont aussi pu être effectués (point 3). Pour finir ce chapitre (point 4), nous proposerons quelques remarques concernant le contenu lexical et les structures syntaxiques généralement retrouvées par type de RVC dans notre corpus.

1. Observations générales

Ce premier point présentera les résultats de notre analyse quantitative en termes de fréquences des RVC (tous types confondus mais par type fonctionnel). Le total de RVC annotés dans le corpus est de 204, avec 104 marques d'enregistrement, 12 marques de confirmation et 71 expressifs (c'est-à-dire 23 marques de convergence et 48 marques d'évaluation) et 17 annotations avec une double valeur. En termes de nombre de mots, ce total revient à 540 mots tous RVC confondus, avec 209 mots pour les ME, 40 pour les MC, 243 mots pour les expressifs (les MC comptant pour 78 mots et les ME pour 165) et 48 mots pour les doubles valeurs.

Les différents résultats seront présentés dans un premier temps, avant de passer aux remarques que l'on peut en tirer.

1.1. Pourcentages des différents types de RVC

La figure ci-dessous présente les pourcentages des différents types et sous-types de RVC par rapport au total de RVC annotés dans le corpus :

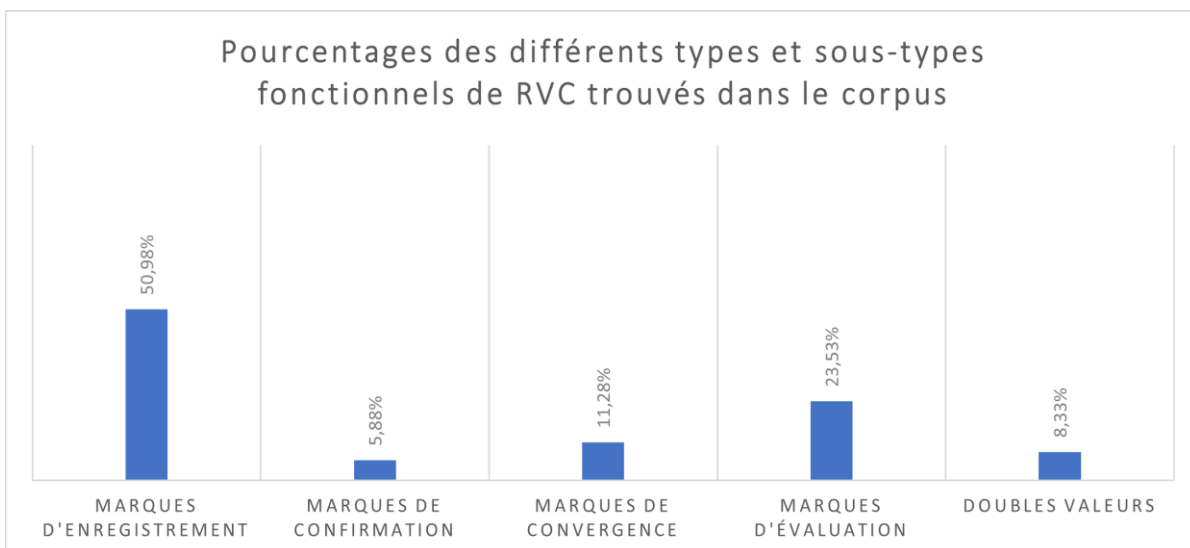


Figure 7 Pourcentages des différents types et sous-types de RVC à l'intérieur des RVC annotés dans le corpus

1.2. Types des doubles valeurs

La figure ci-dessous présente les pourcentages des différentes doubles valeurs retrouvées dans le corpus, tout d'abord vis-à-vis du total des doubles valeurs puis vis-à-vis du total des RVC.

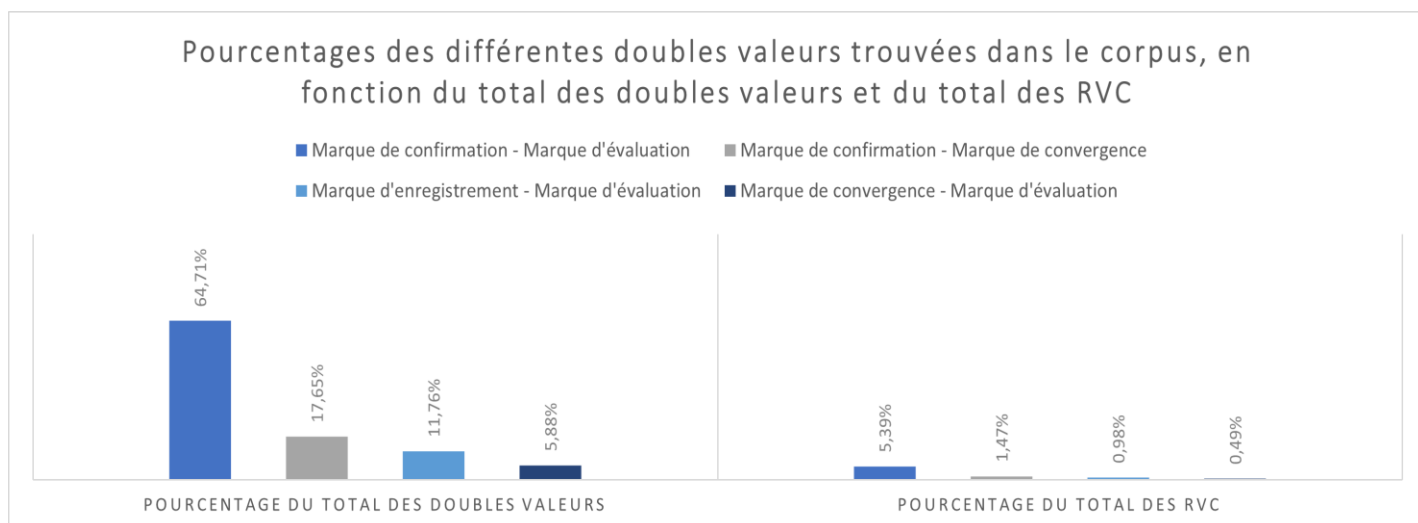


Figure 8 Pourcentages des différentes doubles valeurs retrouvées dans le corpus, vis-à-vis du total des doubles valeurs et vis-à-vis du total des RVC

1.3. Fréquences des RVC dans le corpus

Les figures ci-dessous présentent les fréquences des RVC et de leurs types et sous-types dans le corpus. Les pourcentages présentés indiquent le nombre de mots formant des RVC sur 100 mots.

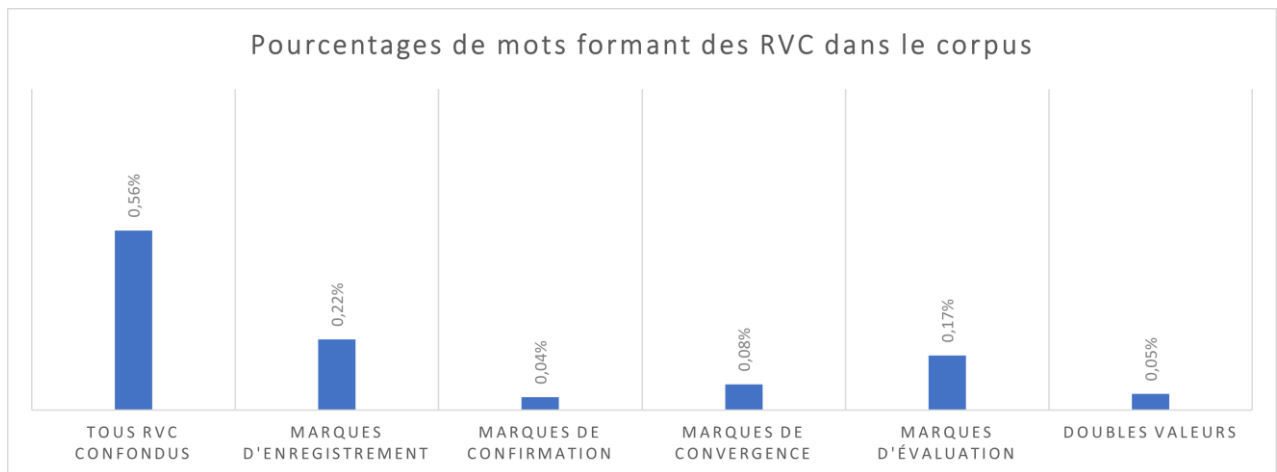


Figure 9 Pourcentages des mots formants des RVC dans le corpus

Dans le cas des interactions avec deux participants uniquement (9 sur les 17 conversations), il a été possible de calculer le nombre de RVC produit par les locuteurs en fonction du nombre de mots de leur interlocuteur.

Les résultats sont les suivants : en moyenne, les locuteurs produisent 0,21 RVC pour 100 mots de leur interlocuteur. Il est à noter cependant que cette moyenne prend uniquement en compte les données de locuteurs produisant effectivement des RVC. Or un bon nombre de locuteurs ne rentrent pas dans ce cas de figure, et ne produisent aucun RVC (sur les 18 locuteurs des interactions à deux, seuls 13 locuteurs produisent au moins un RVC). Il est donc possible de proposer deux autres moyennes, une pour les locuteurs adoptant une position de Récepteur tout au long d'une interaction (seuls à produire des RVC) et une pour les locuteurs jouant un rôle parfois de Récepteur et parfois de Locuteur (interactions dans lesquelles les deux locuteurs produisent des RVC). Ces moyennes sont les suivantes : 0,21 RVC pour 100 mots de l'interlocuteur dans les conversations où les deux locuteurs produisent des RVC, contre 0,22 RVC pour 100 mots dans les conversations où un seul des locuteurs produit des RVC. La différence entre ces moyennes étant quasiment nulle, on peut en conclure que le nombre de RVC produit par un Récepteur n'est apparemment pas influencé par le fait qu'il garde ce statut de Récepteur sur une longue durée, comme on peut supposer que c'est le cas lors de conversations à deux où l'autre interlocuteur ne produit pas de RVC.

Pour conclure, quelques observations générales peuvent être faites d'après ces différentes données :

- Tout d'abord les RVC sont peu fréquents dans le corpus, puisque les mots inclus dans des RVC correspondent à moins d'un pour cent du total de mots dans le corpus (0,56%).
- La répartition des différents types de RVC n'est pas équilibrée, et la majorité des RVC présents dans le corpus sont des marques d'enregistrement (50,98%). Les marques d'évaluation (type expressif) forment le deuxième type fonctionnel le plus fréquent (23,53% des RVC du corpus).
- Si l'on s'intéresse directement au nombre de RVC réalisé en fonction du discours d'un interlocuteur (calcul possible dans les interactions avec deux participants uniquement), il est intéressant de noter qu'un locuteur gardant apparemment une position de Récepteur tout au long de la conversation (seul à utiliser des RVC) ne produit en moyenne pas plus de RVC qu'un locuteur alternant entre une position de Récepteur et une position de Locuteur.

Ces différentes observations générales nous serviront de base pour la suite de l'analyse, qui s'intéressera plus précisément à la variation des RVC. Ainsi nous nous pencherons d'abord sur l'influence possible du score de proximité sur le nombre et le type de RVC rencontrés (point 2), puis sur les variations formelles (contenu lexical) des RVC observées dans le corpus (points 3 et 4).

2. Fréquences selon le degré de proximité

Étant donné notre intérêt pour la variation au sein des régulateurs verbaux de connivence, un calcul de leur fréquence selon le score de proximité des participants a été effectué. Les résultats de ce calcul en fonction du score total de proximité feront l'objet d'un premier point, puis les résultats en fonction de paramètres précis seront présentés. Ces paramètres sont la combinaison des paramètres 08 et 10 (dialogisme/monologisme et développement thématique), ainsi que le paramètre 02 (relation intime ou non).

2.1. Fréquences des RVC selon le score général de proximité

Afin de faciliter les calculs et d'obtenir une meilleure vue d'ensemble, les fréquences ont été calculées sur des tranches de scores et non pas sur chaque score de proximité possible. Les tranches de scores observées pour les scores globaux de proximité sont les suivantes :

taux de proximité faible (de -02 à -05 points), taux assez faible (de -06 à -10 points), taux assez fort (de -11 à -15 points) et taux fort (de -16 à -19 points)⁴⁵.

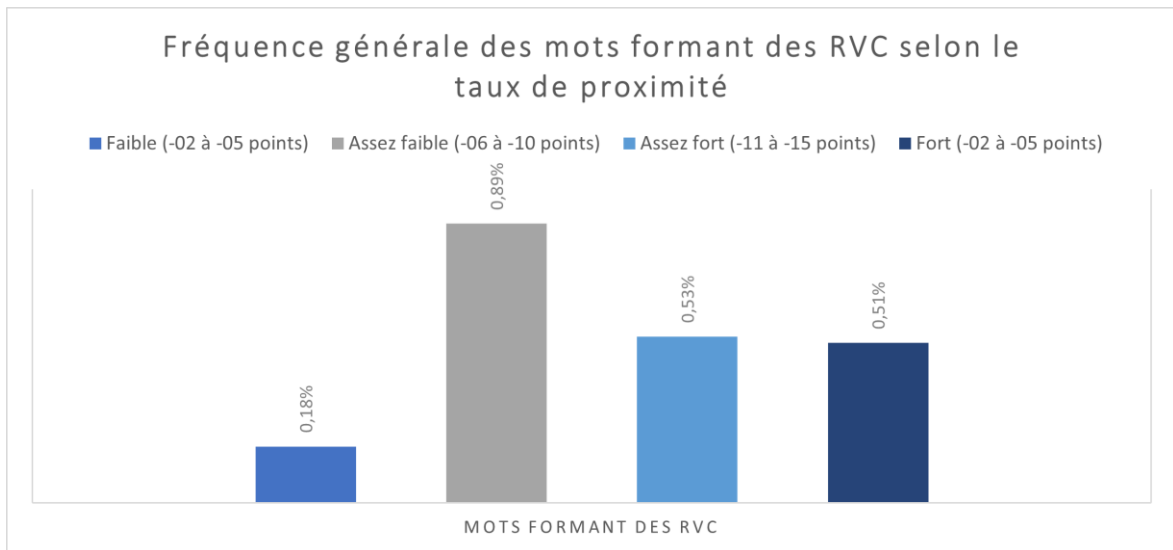


Figure 10 Histogramme des pourcentages de mots formant des RVC selon le taux de proximité dans le corpus

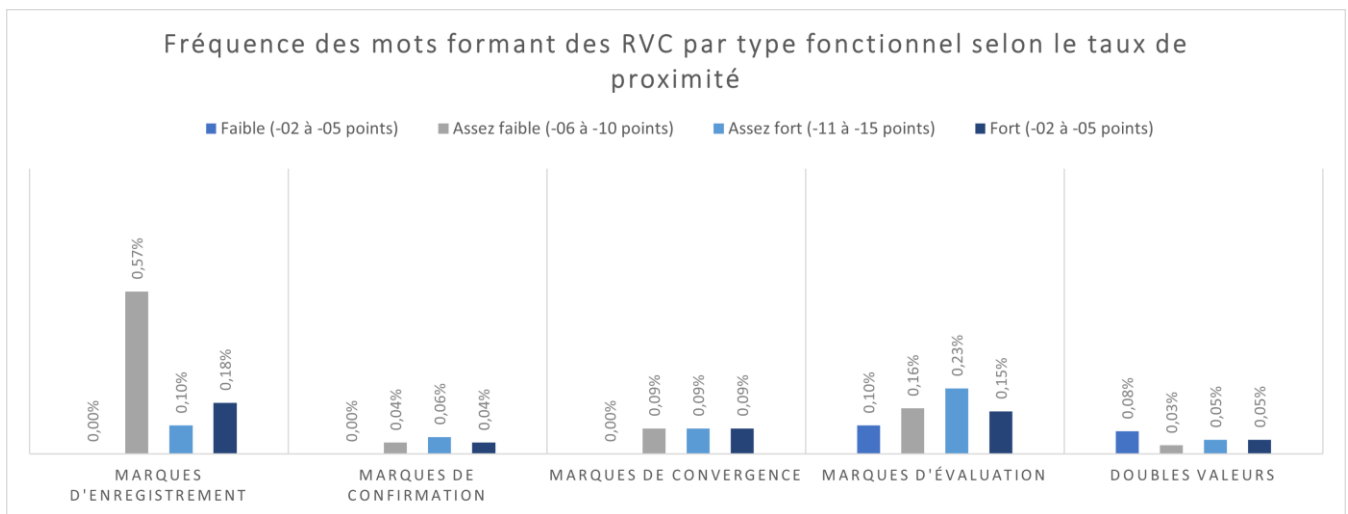


Figure 11 Histogramme des pourcentages de mots formant des RVC par type fonctionnel selon le taux de proximité dans le corpus

Un total de 4 locuteurs (pour 9133 mots dont 4 RVC (soit 16 mots)) ont obtenu un score faible de proximité (entre -02 et -05 points). 75% de ces locuteurs ne produisent aucun RVC (tous les RVC de cette tranche sont produits par le même locuteur).

⁴⁵ -02 et -19 étant les valeurs minimales et maximales des scores dans notre corpus.

Un taux assez faible de proximité (de -06 à -10 points) est atteint par 11 locuteurs, pour un total de 18 232 mots dont 66 RVC (soit 162 mots). Encore une fois, certains locuteurs dans cette tranche de scores ne produisent aucun RVC, mais ce total équivaut cette fois-ci à 36,36% des locuteurs (4 locuteurs sur 11).

16 locuteurs possèdent un score de proximité assez fort (de -11 à -15 points), pour 26 053 mots dont 52 RVC (soit 139 mots). Parmi ces 16 locuteurs, 2 n'utilisent aucun RVC (soit 12,5% des locuteurs).

Pour finir, tous les locuteurs de la tranche de scores allant de -16 à -19 points (taux de proximité fort), c'est-à-dire 11 locuteurs, produisent au moins un RVC. Au total, ces locuteurs prononcent 43 603 mots dont 82 RVC (soit 223 mots).

S'il est difficile de tirer des conclusions valides concernant les RVC produits avec un score de proximité faible (de -02 à -05 points), tant les données sont faibles et peu représentatives, il est néanmoins possible de dégager des dynamiques intéressantes pour les tranches de scores plus élevées (taux assez faible à fort), mais aussi de faire quelques observations quant aux locuteurs ne produisant pas de RVC.

Ainsi il apparaît que les locuteurs possédant un score de proximité assez faible (de -06 à -10 points) sont les plus productifs en termes de fréquence de mots formant des RVC (0,89% des mots). De plus, cette productivité concerne majoritairement les marques d'enregistrement. Ce fait peut être en partie expliqué par la présence d'une locutrice dirigeant un entretien dans cette tranche de scores. En effet, cette locutrice est celle qui produit le plus de RVC dans tous le corpus (7,45% de ses mots forment des RVC), tous ayant fonction de marque d'enregistrement. Étant donné le format d'entretien de la conversation à laquelle elle participe, il semble que les marques d'enregistrement qu'elle produit puissent être apparentées à des « interventions de nature incitatives », selon les termes de Ruquoy (1995).

Les participants avec un taux de proximité assez fort et fort (de -11 à -15 points et de -16 à -19 points) présentent des pourcentages de mots formant des RVC moins importants (0,53% et 0,51% des mots), mais relativement stables, avec une différence de 0,02% uniquement entre ces deux tranches de score. Quant au pourcentage de locuteurs ne produisant aucun RVC, celui-ci décroît pour chaque tranche de scores et passe de 75% des locuteurs (taux de proximité faible) à 36,36% (taux assez faible), 12,5% (taux assez fort) et finalement 0% pour la tranche de scores tendant vers une proximité maximale (taux fort). Il

semble donc qu'une plus grande proximité implique une utilisation de RVC, sinon plus fréquente, étendue et répartie sur l'ensemble des participants.

Ces données nous permettent déjà de revenir sur notre hypothèse d'une inscription des régulateurs dans la proximité (voir chapitre 2. 2). En effet, il apparaît que la fréquence des RVC n'est pas corrélée avec le score de proximité des interactants, une fréquence plus importante étant même observée lors d'une conversation de type entretien (tendant ainsi supposément vers la distance) et dans le discours des locuteurs possédant un score de proximité assez faible. Il semble par conséquent que cette hypothèse nécessite d'ores et déjà d'être rediscutée, ce que nous continuerons à faire dans le point 3. 3 de ce chapitre en observant cette fois-ci le contenu lexical des RVC selon le score de proximité.

2.2. Fréquences des RVC selon les scores de proximité aux paramètres 8 & 10 (dialogisme/monologisme et développement thématique libre/fixe)

Un système de tranche de scores a aussi été utilisé afin d'étudier la variation des RVC selon la combinaison des paramètres 08 et 10 des paramètres situationnels établis par Koch & Esterreicher (2001). Les paramètres 08 et 10 correspondent au continuum entre le dialogisme et le monologisme d'une part, et le continuum entre liberté et fixation thématique d'autre part. Le choix de s'intéresser à ces deux paramètres relève de l'intention de tester l'influence des *multi-unit turns* sur le nombre de RVC retrouvés. En effet, une partie des auteurs ayant travaillé sur les régulateurs considèrent que ceux-ci sont retrouvés principalement, voire uniquement (Goodwin, 1986, Eiswirth, 2020b entre autres), dans de tels tours de parole. Pour rappel, les *multi-unit turns* sont des tours de parole étendus, obtenus lorsqu'un même locuteur garde la parole après plusieurs *turn-constructive units*, ou TCU⁴⁶. Nous émettons l'hypothèse que ces tours sont plus fréquents voire majoritaire dans des interactions monogérées avec une fixation des thèmes, mais peu fréquents dans des conversations co-gérées avec une liberté thématique importante. En suivant cette hypothèse, l'examen des usages des RVC et surtout de leur fréquence selon les scores combinés des paramètres 08 & 10 nous permettra de vérifier une influence des *multi-unit turns* sur l'usage et la fréquence des RVC.

⁴⁶ Voir chapitre 2. 1.

Les tranches de scores pour la combinaison de ces deux paramètres sont les suivantes : taux de proximité faible (04 à 02 points), moyen (-01 à 01 points) et fort (-02 à -04 points).

7 locuteurs possèdent un taux de proximité faible pour ces deux paramètres (04 à 02 points), s'inscrivant ainsi dans le monologisme et la fixation thématique. Ces locuteurs produisent 10 558 mots dont 36 RVC (soit 80 mots). Le taux de proximité moyen est atteint par 16 locuteurs (01 à -01 points), pour un total de 27 641 mots dont 60 RVC (soit 170 mots). La parole de ces locuteurs se situe entre le monologisme et le dialogisme, et entre la fixation et la liberté thématique. Finalement, 19 locuteurs ont obtenu un taux de proximité fort pour ces deux paramètres dans notre corpus (-02 à -04 points), produisant un total de 58 882 mots dont 108 RVC (soit 290 mots). Ces locuteurs s'inscrivent dans le dialogisme et la liberté thématique.

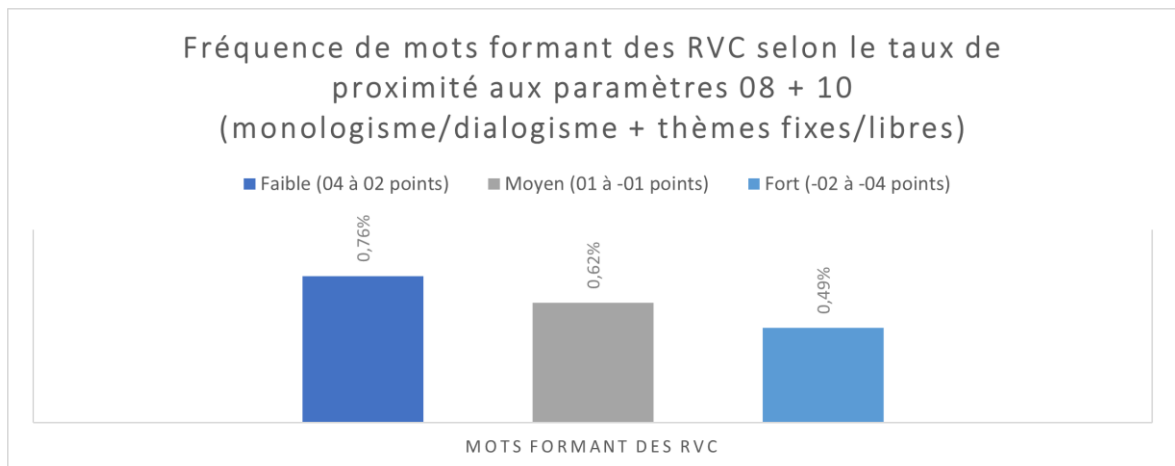


Figure 12 Histogramme des pourcentages de mots formant des RVC selon le taux de proximité aux paramètres 08 + 10

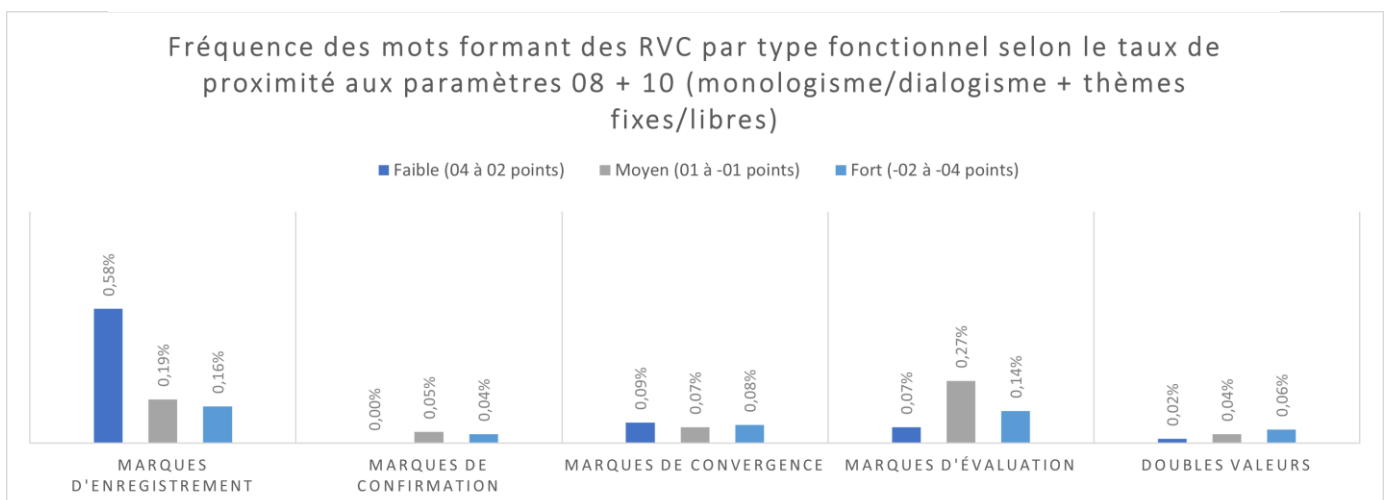


Figure 13 Histogramme des pourcentages de mots formant des RVC par type fonctionnel selon le taux de proximité aux paramètres 08 + 10

Quelques observations peuvent être établies sur la base de ces données : tout d'abord le pourcentage total de RVC par tranche de scores est décroissant, avec un pourcentage moindre plus la proximité est grande (0,76% de mots formant des RVC pour le taux de proximité faible, 0,62% pour le taux moyen et 0,49% pour le taux fort). S'il est difficile de valider l'hypothèse d'une plus grande présence de régulateurs dans les *multi-unit turns* en se basant uniquement sur ces données, il est néanmoins possible de dire que ces données vont dans le sens d'une telle hypothèse. Il serait donc intéressant de vérifier ce fait avec plus de données, en se basant cette fois-ci non pas sur le score de proximité aux paramètres 08 et 10 (qui offre certes une indication quant au taux possible de *multi-unit turns* mais qui reste peu précis), mais bien sur une réelle prise en compte de ce type de tours.

Un autre fait intéressant à noter dans ces données est le pourcentage plus important de marques d'enregistrement dans la tranche de scores tendant le maximum vers la distance (de 02 à 04 points, taux de proximité faible). Il semble donc qu'une interaction tendant vers le monologisme et la fixation thématique implique un usage plus important de RVC d'un côté, mais aussi de l'autre côté une variation moindre des types fonctionnels de RVC, avec une préférence pour les RVC ayant fonction de marques d'enregistrement.

2.3. Fréquences des RVC selon les scores de proximité au paramètre 2 (interlocuteur intime/inconnu)

L'influence potentielle d'un dernier paramètre situationnel sur la production des RVC a pu être testée, en l'occurrence le paramètre 02 (interlocuteur intime/inconnu). L'hypothèse en lien avec cette analyse est l'idée que le récepteur montrera un engagement personnel moindre vis-à-vis d'un inconnu, avec un taux plus important de marques d'enregistrement comparé aux autres types fonctionnels.

Seuls 3 locuteurs ont obtenu un taux de proximité faible (0 à 01⁴⁷ points : interlocuteurs moyennement à peu connus) pour ce deuxième paramètre, pour un total de 7 488 mots dont 7 RVC (soit 13 mots).

⁴⁷ Aucun locuteur n'a obtenu le score maximal de distance (02 points) pour ce paramètre dans notre corpus.

14 locuteurs ont obtenu un taux de proximité moyen (-01 point : interlocuteur bien connu) pour ce deuxième paramètre, pour un total de 24 870 mots dont 46 RVC (soit 126 mots).

Le score maximal de proximité au paramètre 02 est atteint par 25 locuteurs (taux de proximité fort : interlocuteur très bien connu), pour un total de 64 663 mots dont 151 RVC (soit 401 mots).

Les histogrammes présentant les résultats de cette analyse sont proposés ci-dessous :

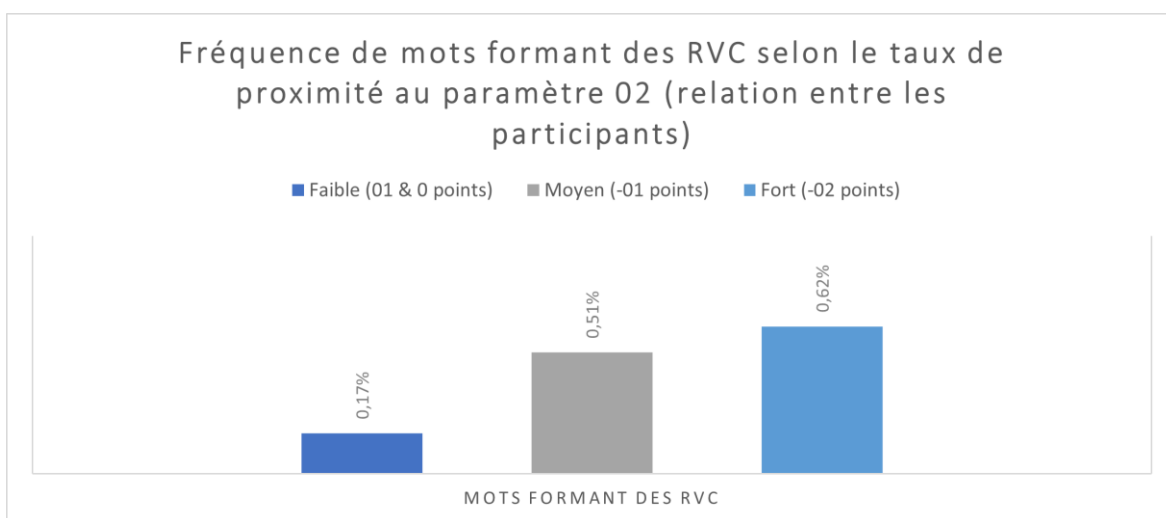


Figure 12 Histogramme des pourcentages de mots formant des RVC selon le taux de proximité au paramètre 02

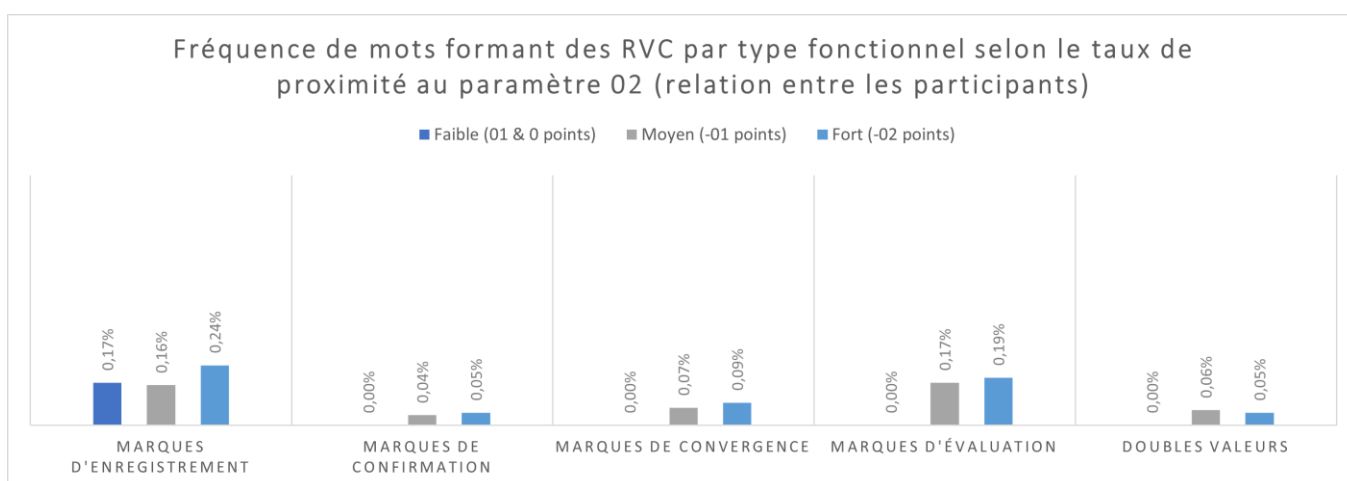


Figure 13 Histogramme des pourcentages de mots formant des RVC par type fonctionnel selon le taux de proximité au paramètre 02

Ces différents résultats montrent qu'il existe effectivement, dans notre corpus du moins, une différence d'usage des RVC face à un interlocuteur intime ou inconnu. En effet,

les locuteurs s'adressant à des interlocuteurs relativement inconnus (taux de proximité faible) produisent très peu de RVC, et tous les RVC qu'ils produisent sont des marques d'enregistrement (0,17% des mots). A l'inverse les locuteurs s'adressant à des locuteurs intimes (taux de proximité moyen et faible) produisent plus de mots formant des RVC (0,51% et 0,62% des mots respectivement) au total, et les types fonctionnels retrouvés sont relativement équilibrés⁴⁸.

Pour conclure, il semble donc que la proximité exerce effectivement une certaine influence sur la variation fonctionnelle des RVC dans notre corpus, en particulier pour un type précis : les marques d'enregistrement. En effet, celles-ci sont plus fréquentes lorsque la distance est à son maximum pour les paramètres 08 et 10, ainsi que pour le paramètre 02 (c'est-à-dire dans le cas d'interactions monologiques présentant une certaine fixation thématique, et/ou en présence d'interlocuteurs peu connus). Il existe cependant une différence au niveau du nombre total de RVC produits entre les paramètres 08 & 10 et le paramètre 02. On retrouve en effet un pourcentage de RVC très faible face à un locuteur peu connu (paramètre 02), mais un pourcentage un peu plus élevé dans une interaction monologique à thème fixe (paramètres 08 et 10). Ces observations ne sont plus valables dans le cas du score général de proximité en revanche, même si une tendance vers une production de RVC généralisée à tous les locuteurs plus la proximité est grande peut être observée.

Ces différentes observations font ressortir un certain impact du score de proximité sur les RVC utilisés par les locuteurs dans notre corpus. Cependant cet impact est loin d'être unifié, et certains effets sont convergents entre paramètres (par exemple la préférence pour les marques d'enregistrement dans le cas d'un taux de proximité faible aux paramètres 08 & 10 et 02), tandis que d'autres paraissent divergents (la fréquence totale des RVC est différente pour ces deux paramètres). Il paraît ainsi nécessaire de continuer à étudier le comportement des RVC en fonction du score de proximité, et ce de façon plus détaillée (sur des temps plus courts, en s'intéressant aux 10 paramètres situationnels et non seulement à deux d'entre eux), afin de mieux comprendre les dynamiques d'utilisation de ces éléments linguistiques.

⁴⁸ Les marques d'enregistrements et marques d'évaluation étant encore un fois légèrement plus fréquentes que les autres, comme c'est déjà le cas pour les résultats généraux sur tout le corpus (figure 9).

Les fréquences des RVC dans le corpus, établies sur la totalité du corpus mais aussi selon le détail de certains paramètres situationnels, nous ont par conséquent permis de tirer quelques tendances générales concernant les RVC et leurs types fonctionnels dans le corpus. Afin d'aller plus loin dans notre analyse et de mieux comprendre ces tendances, il semble cependant primordial de s'intéresser aussi au contenu lexical des RVC. En effet, ce contenu pourrait nous en apprendre plus sur la manière dont ces éléments linguistiques sont utilisés dans les conversations. Cet examen nous permettra aussi de revenir sur l'impact de ce contenu lexical sur les types fonctionnels observés.

3. *Inventaire des formes des RVC dans le corpus*

Ce troisième point dédié aux résultats de l'analyse des données proposera plusieurs inventaires des différentes formes de RVC retrouvées dans le corpus. Un premier inventaire, établi en fonction du nombre d'éléments lexicaux (et particules satellites dans certains cas⁴⁹) dont les RVC sont composés, sera proposé. Un deuxième inventaire reprendra ces formes selon les différents types et sous-types fonctionnels. Enfin un troisième inventaire, établi en fonction du score de proximité total (par tranches de scores), sera présenté. Pour terminer ce troisième point, le contenu lexical des doubles valeurs sera détaillé, ainsi que le pourcentage de formes uniques dans le corpus.

3.1. *Inventaire des formes observées des RVC selon le nombre d'éléments lexicaux (+ particules satellites dans certains cas)*

Afin d'obtenir une meilleure idée de la variation formelle possible des RVC dans notre corpus, nous avons tout d'abord examiné les différentes combinaisons formelles possibles en termes de nombre d'éléments lexicaux + particules satellites (voir note 49). Les tableaux résultant de cet examen sont présentés dans l'annexe 2, p. 131.

Ces différents tableaux nous permettent d'avoir une meilleure idée de la forte variation formelle des RVC, qui peuvent donc être composés de un jusqu'à sept éléments lexicaux + particules satellites. Ils ne nous donnent en revanche aucune indication sur une

⁴⁹ La présence possible d'éléments satellites est signalée par l'indication + satellites entre parenthèses. Ces différentes particules satellites regroupent les vocalisations telles que ah, oh, hum, etc. ; les marqueurs tels que ben, bah, mais, etc. ainsi que les oui et ouais.

influence possible des différents types et sous-types fonctionnels sur cette variation formelle. Pour y remédier, un inventaire des formes des RVC selon notre typologie sera proposée dans la partie 3. 2. ci-dessous.

3.2. Inventaire des formes observées des RVC selon les types et sous-types fonctionnels

Pour commencer sur ce point, il est important de noter que les différents tableaux présentés ci-dessous ne neutralisent pas les doubles valeurs, contrairement aux pourcentages présentés dans le point 2 (neutralisation grâce à un calcul dédié). Une forme doublement annotée est donc prise en compte dans les deux types correspondants dans les tableaux ci-dessous.

Tableau 19 Inventaire des RVC fonctionnant comme marques d'enregistrement dans le corpus

Formes observées pour les RVC ayant fonction de marques d'enregistrement			
d'accord	(77)	ah OK	(1)
OK	(14)	d'accord OK	(1)
ah d'accord	(4)	d'accord ouais ouais	(1)
hum d'accord	(4)	d'acc	(1)
ah ouais d'accord	(2)	ouais d'accord	(1)

Tableau 20 Inventaire des RVC fonctionnant comme marques de confirmation dans le corpus

Formes observées pour les RVC ayant fonction de marques de confirmation					
cool	(4)	bah oui forcément ouais	(1)	hum c'est cool	(1)
c'est clair	(3)	c'est normal	(1)	bien sûr	(1)
ouais c'est vrai	(3)	c'est vrai	(1)	ouais c'est clair	(1)
c'est chaud	(2)	c'est ça	(1)	ouais c'est sûr	(1)
ah c'est chaud	(1)	c'est pas faux	(1)	ouais c'est cool ouais	(1)
bah c'est cool	(1)	hum bien	(1)	ouais je te jure	(1)

Tableau 21 Inventaire des RVC expressifs fonctionnant comme marques de convergence dans le corpus

Formes observées pour les RVC expressifs – marques de convergence					
tu m'étonnes	(5)	ben normal	(1)	mais oui c'est rigolo	(1)
grave	(3)	ben grave ouais	(1)	ouais c'est vrai	(1)
exactement ouais	(1)	ben ouais c'est ça	(1)	ouais c'est clair c'est clair	(1)
ah ouais carrément	(1)	c'est normal	(1)	ouais j'avoue	(1)
bah tu m'étonnes	(1)	super beau hein	(1)	ouais je te jure	(1)
bah j' imagine ouais	(1)	putain ouais	(1)	ouais ça doit être chaud hein	(1)
bah oui forcément	(1)	mais bien sûr ouais ouais ouais	(1)	oui oui c'est normal	(1)

Tableau 22 Inventaire des RVC expressifs fonctionnant comme marques d'évaluation dans le corpus

Formes observées pour les RVC expressifs – marques d'évaluation					
cool	(5)	ah ouais c'est fou ça	(1)	hum ah c'est bien ça	(1)
oh putain	(3)	bah c'est cool	(1)	hum c'est cool	(1)
ah c'est chaud	(2)	c'est malheureux	(1)	hum d'accord	(1)
c'est chaud	(2)	c'est cool ça	(1)	très utile	(1)
l'enfer	(2)	c'est fou ça	(1)	putain	(1)
abusé	(1)	c'est nul ça	(1)	putain atroce	(1)
top	(1)	c'est chaud quand même	(1)	putain c'est chaud	(1)
ah bien	(1)	c'est affreux quoi	(1)	truc de fou	(1)
ah: chouette:	(1)	c'est trop affreux	(1)	oh c'est une horreur	(1)
ah bah c'est bien	(1)	c'est la merde	(1)	oh la vache	(1)
ah ben ça c'est chouette ça	(1)	c'est pas cool	(1)	oh punaise	(1)
ah c'est trop bien	(1)	je te jure	(1)	oh oh oh putain	(1)
ah tout pourri	(1)	d'accord	(1)	ouais c'est bizarre hein	(1)
ah la vache	(1)	OK	(1)	ouais c'est cool ouais	(1)
ah putain	(1)	tant mieux	(1)	ouais c'est pas cool ça	(1)
ah mais c'est trop bien ça	(1)	han putain	(1)	ça doit faire trop bizarre	(1)
ah ouais carrément	(1)	hum bien	(1)	roh putain c'est affreux	(1)
ah ouais c'est chiant ça	(1)	hum: chaud quand même	(1)		

Le tableau ci-dessous présente de plus le détail des doubles valeurs retrouvées dans le corpus. Certaines de ces doubles valeurs seront discutées plus en détails dans le point 4. 1 de ce chapitre.

Tableau 23 Inventaire des RVC ayant reçu une double valeur dans le corpus

Formes observées pour les RVC avec une double valeur et détail des types concernés			
cool	(4)	Marque de confirmation	Expressif – Marque d'évaluation
c'est chaud	(2)	Marque de confirmation	Expressif – Marque d'évaluation
ah c'est chaud	(1)	Marque de confirmation	Expressif – Marque d'évaluation
ah ouais carrément	(1)	Expressif – Marque de confirmation	Expressif – Marque d'évaluation
bah c'est cool	(1)	Marque de confirmation	Expressif – Marque d'évaluation
c'est normal	(1)	Marque de confirmation	Expressif – Marque de convergence
d'accord	(1)	Marque d'enregistrement	Expressif – Marque d'évaluation
hum bien	(1)	Marque de confirmation	Expressif – Marque d'évaluation
hum c'est cool	(1)	Marque de confirmation	Expressif – Marque d'évaluation
hum d'accord	(1)	Marque d'enregistrement	Expressif – Marque d'évaluation
ouais c'est cool ouais	(1)	Marque de confirmation	Expressif – Marque d'évaluation
ouais c'est vrai	(1)	Marque de confirmation	Expressif – Marque de convergence
ouais je te jure	(1)	Marque de confirmation	Expressif – Marque de convergence

Ces différents inventaires reprenant les formes des RVC observés dans le corpus selon leur type ou sous-type fonctionnel nous permettent de tirer quelques conclusions quant à la variation formelle des RVC. Ainsi, il semble que cette variation soit plus importante au sein des expressifs, en particulier des marques d'évaluation. Afin de prouver et de détailler cette observation, un dernier calcul cherchant à établir le pourcentage de formes uniques par type et sous-type fonctionnel dans le corpus a été effectué.

Au total, 84,27% des formes observées sont des formes uniques, apparaissant une seule fois dans le corpus. Ce nombre varie de plus selon le type de RVC :

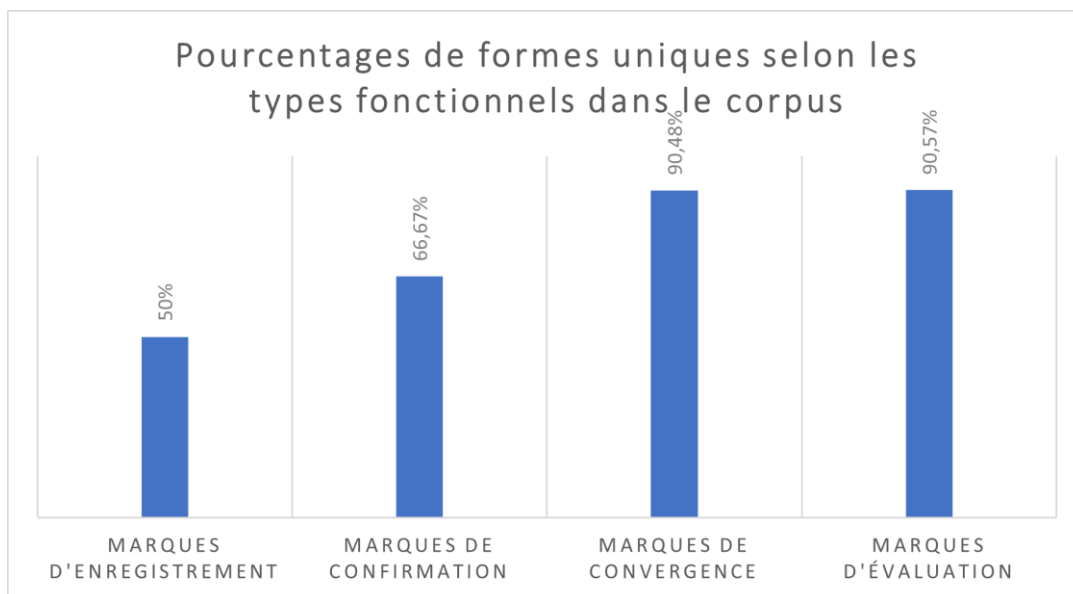


Figure 14 Pourcentages de formes uniques au sein des différents types et sous-types fonctionnels

Ces pourcentages nous permettent de confirmer notre première observation, et de l'étendre à l'ensemble des expressifs. Il existe bel et bien une plus grande variation des formes des RVC au sein des RVC de type expressif dans notre corpus, variation associée à un plus grand nombre de formes uniques. Une raison possible de cette plus forte variation pourrait être l'importance de la subjectivité inhérente aux RVC expressifs, qui exigent de s'adapter à la situation et reposent sur une réelle subjectivité des interactants plutôt que sur une réponse de « routine », adaptée à n'importe quelle situation.

Nous reviendrons sur les détails du contenu lexical des RVC selon leur type fonctionnel dans le point 4 de ce chapitre, en examinant en particulier les unités lexicales sur lesquelles ils sont construits, ainsi que les structures syntaxiques récurrentes. Avant ceci cependant, il conviendra de revenir, comme prévu, sur la variation formelle des RVC en fonction de degré de proximité.

3.3. Inventaire des formes observées des RVC selon le degré de proximité

Nous avons vu précédemment que la fréquence générale des RVC variait en fonction du degré de proximité, avec une fréquence plus importante dans le cas d'un score assez faible (0,89% de mots formant des RVC, prédominance des marques d'enregistrements), et plus faible mais assez stable pour les scores plus élevés (0,53% et 0,51% de mots formant des RVC pour les scores assez forts et forts). Ces fréquences générales, mais aussi les fréquences détaillées par type fonctionnel, ne nous donnent cependant que peu d'informations quant à la manière dont sont utilisés les RVC en fonction de la proximité.

Par conséquent, nous nous intéresserons désormais non pas à la fréquence, mais au contenu lexical des RVC observés dans notre corpus, en fonction du degré de proximité. Les tranches de scores faible (-02 à -05 points), assez faible (-06 à -10 points), assez forte (-11 à -15 points) et forte (-16 à -19 points) sont à nouveau mobilisées pour cette analyse. Quatre tableaux sont ainsi proposés ci-dessous, par tranche de scores. Il est à noter que les cases avec un fond bleu dans ces tableaux indiquent les doubles valeurs.

Tableau 24 Inventaire des RVC produits par les locuteurs avec un score de proximité faible

RVC produits par des locuteurs avec un score de proximité faible (-02 à -05 points)			
M. d'enregistrement	M. de confirmation	M. de convergence	M. d'évaluation
	ah c'est chaud (1)		ah c'est chaud (1)
	c'est chaud (1)		c'est chaud (1)
			c'est chaud quand même (1)
			hum chaud quand même (1)

Comme cela a déjà pu être dit, une seule locutrice produit des RVC au sein des locuteurs avec un faible taux de proximité (entre -02 et -05 points). Il est donc difficile de tirer de réelles conclusions sur le contenu lexical et les types fonctionnels retrouvés, mais il est intéressant de noter que la locutrice utilise systématiquement des formes composées autour de l'unité lexicale *chaud*. De plus, une double valeur marque de confirmation/marque d'évaluation a été assignée à deux des RVC qu'elle produit (double valeur qui, comme nous allons le voir, semble être le signe d'une désémantisation de *chaud*).

Il se trouve que cette locutrice s'adresse à un locuteur avec lequel elle est plutôt intime (score de -01 point pour l'intimité), locuteur qui lui fait un récit de son expérience à l'armée sur un ton assez émotionnel, avec quelques anecdotes sortant de l'ordinaire. Ce contexte paraît expliquer l'utilisation de RVC plutôt expressifs sémantiquement, mais qui sont aussi des RVC de routine. En effet, ces RVC semblent finalement indiquer un engagement relativement faible de cette locutrice dans la conversation (fait exacerbé par l'utilisation systématique de la seule unité lexicale *chaud*). Cet engagement relativement faible peut être mis en lien avec le faible score de proximité de cette locutrice, même s'il est difficile de dire si les deux sont strictement liés en se basant uniquement sur ces données. Il est aussi possible que cette locutrice fasse preuve d'un engagement personnel faible dans la conversation, ou d'une faible expressivité, parce qu'elle ne sait pas quoi dire vis-à-vis des anecdotes parfois dures et choquantes de son interlocuteur.

Tableau 25 Inventaire des RVC produits par les locuteurs avec un score de proximité assez faible

RVC produits par des locuteurs avec un score de proximité assez faible (-06 à -10 points)			
M. d'enregistrement	M. de confirmation	M. de convergence	M. d'évaluation
d'accord (42)	c'est clair (1)	ah ouais carrément (1)	l'enfer (2)
hum d'accord (3)	ouais c'est sûr (1)	ben normal (1)	ah ben ça c'est chouette ça (1)
OK (3)		mais bien sûr ouais ouais ouais (1)	ah ouais carrément (1)
ah d'accord (1)		oui oui c'est normal (1)	c'est malheureux (1)
d'accord ouais ouais (1)		tu m'étonnes (1)	c'est nul ça (1)
ouais d'accord (1)			c'est pas cool (1)
			d'accord (1)
			oh punaise (1)
			ouais c'est pas cool ça (1)

En termes de contenu lexical, il est intéressant de noter que les marques d'enregistrement au sein de cette tranche de scores assez faible sont majoritairement des *d'accord*, accompagnés ou non de particules satellites telles que *hum* ou *ouais*. Ce fait, associé à la forte fréquence des marques d'enregistrement pour cette tranche, va dans le sens d'un usage usuel de *d'accord*. Seules quelques marques d'enregistrement prennent la forme de *OK*, qui paraît plus familier. Cette absence de formes lexicales familières est aussi notable au sein des autres types fonctionnels, y compris expressifs. Ainsi la seule interjection observée est *punaise*, et les autres unités lexicales retrouvées sont assez peu intensives (*normal*, *cool* - cette dernière unité lexicale étant associée à une négation dans toutes ses occurrences), et parfois assez désuètes (*malheureux*, *l'enfer*). Un dernier point notoire est l'utilisation fréquente de l'anaphore *ça* au sein des marques d'évaluation, qui pourrait être produite pour rappeler le discours du Locuteur. Le style général utilisé par les locuteurs ayant un score de proximité assez faible semble par conséquent être plutôt soutenu.

Tout comme la locutrice associée à un score de proximité faible (voir tableau précédent), les locuteurs de cette tranche de scores assez faible réutilisent régulièrement les mêmes unités lexicales, par participant et/ou par conversation. Ainsi les deux *l'enfer* sont produits par deux locutrices différentes dans la même conversation, les RVC *ben normal* et *oui oui c'est normal* sont aussi produits par une seule locutrice, et pour finir tous les RVC dans lesquels l'anaphore *ça* est utilisée sont retrouvés dans la même conversation. Il apparaît ainsi encore une fois, selon nous, que l'usage de RVC avec un score de proximité assez

faible est un usage de routine. On pourrait par conséquent considérer que la production de RVC, associée à un score de proximité faible à assez faible, relève plutôt de la politesse, de la routine (avec de ce fait un certain aspect inconscient), que d'une utilisation réellement subjective, personnalisée des RVC.

Tableau 26 Inventaire des RVC produits par les locuteurs avec un score de proximité assez fort

RVC produits par des locuteurs avec un score de proximité assez fort (-11 à -15 points)			
M. d'enregistrement	M. de confirmation	M. de convergence	M. d'évaluation
d'accord (9)	ouais c'est vrai (2)	tu m'étonnes (3)	abusé (1)
OK (6)	bah forcément ouais (1)	bah tu m'étonnes (1)	ah bien (1)
ah d'accord (1)	c'est chaud (1)	exactement ouais (1)	ah c'est chaud (1)
hum d'accord (1)	hum bien (1)	putain ouais (1)	ah chouette (1)
	ouais c'est clair (1)	ouais j'avoue (1)	ah la vache (1)
	ouais c'est cool ouais (1)	bah oui forcément (1)	ah ouais c'est chiant ça (1)
Suite des marques d'évaluation			
ah tout pourri (1)	hum bien (1)	oh putain (1)	très utile (1)
c'est chaud (1)	hum d'accord (1)	OK (1)	truc de fou (1)
c'est fou ça (1)	oh c'est une horreur (1)	ouais c'est cool ouais (1)	top (1)
c'est trop affreux (1)	oh la vache (1)	putain atroce (1)	
ça doit faire trop bizarre (1)	oh oh oh putain (1)	putain c'est chaud (1)	

Un fait notoire au premier abord, est que la variation des formes lexicales est beaucoup plus importante au sein de cette tranche de scores de proximité (score assez fort, entre -11 et -15 points), en particulier au sein des marques d'évaluation (type expressif). Les marques d'évaluation sont par ailleurs le type fonctionnel le plus utilisé dans cette tranche de scores assez forte (0,23% de mots formant des RVC dans cette tranche de score pour ce type, tandis que les autres types fonctionnels ne dépassent pas les 0,10%, voir figure 11). Il apparaît de plus que les unités lexicales retrouvées (*putain, chaud, chiant, top, etc.*) renvoient à un style moins soutenu que celles trouvées dans la tranche de scores de proximité assez faible. Enfin, certaines formes lexicales dans cette tranche de scores sont plutôt intensives (par exemple *fou, horreur, affreux, ou encore forcément*), et certaines sont même renforcées par des intensifieurs (*tout pourri, trop affreux, trop bizarre, très utile*), ou par l'ajout d'une interjection (*putain atroce, putain c'est chaud*).

On peut ainsi émettre l'hypothèse que les locuteurs se situant dans cette tranche de scores de proximité (score assez fort) ont une utilisation plus personnelle des RVC, c'est-à-dire une utilisation moins liée à la politesse. En effet, les formes lexicales retrouvées semblent être adéquates pour mettre en avant la personnalité des différents locuteurs, avec une plus grande subjectivité et une variation plus importante. L'usage de RVC paraît ainsi moins usuel, et basé au contraire sur de vraies dynamiques interpersonnelles où la personnalité des différents interactants est prise en compte.

Tableau 27 Inventaire des RVC produits par les locuteurs avec un score de proximité fort

RVC produits par des locuteurs avec un score de proximité fort (-16 à -19 points)			
M. d'enregistrement	M. de confirmation	M. de convergence	M. d'évaluation
d'accord (27)	cool (4)	grave (3)	cool (5)
OK (5)	c'est clair (2)	bah j'imagine ouais (1)	oh putain (2)
ah d'accord (2)	bah c'est cool (1)	ben grave ouais (1)	ah bah c'est bien (1)
ah ouais d'accord (2)	bien sûr (1)	ben ouais c'est ça (1)	ah c'est trop bien (1)
	c'est ça (1)	c'est normal (1)	ah mais c'est trop bien ça (1)
	c'est normal (1)	mais oui c'est rigolo (1)	ah ouais c'est fou ça (1)
	c'est pas faux (1)	ouais c'est clair c'est clair (1)	ah putain (1)
	c'est vrai (1)	ouais c'est vrai (1)	bah c'est cool (1)
	hum c'est cool (1)	ouais ça doit être chaud hein (1)	c'est affreux quoi (1)
	ouais c'est vrai (1)	ouais je te jure (1)	c'est cool ça (1)
	ouais je te jure (1)	super beau hein (1)	c'est la merde (1)
		tu m'étonnes (1)	han putain (1)
Suite des marques d'évaluation			
hum ah c'est bien ça (1)	je te jure (1)	putain (1)	tant mieux (1)
hum c'est cool (1)	ouais c'est bizarre hein (1)	roh putain c'est affreux (1)	

La fréquence des marques d'enregistrement revient à la hausse au sein de la tranche de scores indiquant un taux de proximité fort (voir figure 11), même si les marques d'évaluation gardent aussi une place importante (0,18% et 0,15% respectivement, pour une fréquence globale de 0,51% mots formant des RVC dans le discours des locuteurs avec un taux fort de proximité). Il semble que la variation lexicale soit moins importante au sein de

cette tranche de scores, comparé à la tranche de scores assez forte, avec des unités lexicales récurrentes telles que *chaud*, *cool* ou *bien* (souvent associées à une double valeur marque de confirmation/marque d'évaluation) ou encore *grave*, *clair* et *putain*. L'intensité paraît cependant toujours présente au sein des expressifs, avec plusieurs interjections (*putain*), des intensifieurs (*super beau hein*, *ah c'est trop bien*, *ah mais c'est trop bien ça*) ou encore la présence du marqueur discursif *grave*. Contrairement aux locuteurs des autres tranches de scores, les locuteurs ayant obtenu un score de proximité fort produisent quelques RVC formés autour de la première personne du présent (*j'imagine* et *je te jure*). Pour terminer, il est possible d'observer que cette tranche de scores présente de nombreux cas de doubles valeurs (3 marques de confirmation/marques de convergence et 3 marques de confirmation/marques d'évaluation).

A partir de ces différentes observations, il est possible d'émettre plusieurs hypothèses. Tout d'abord, il est probable que la présence de nombreux segments avec une double valeur soit liée à un besoin moins prononcé de s'exprimer clairement, sans ambiguïté, pour les locuteurs avec un score de proximité fort. En effet, ces doubles valeurs correspondent à des cas où la prosodie du RVC ne permettait pas d'en déterminer une seule fonction précise.

Une deuxième hypothèse est celle d'un retour vers une utilisation plus rituelle des RVC, possiblement liée à un besoin moins important de montrer son originalité (le score de proximité étant déjà très élevé). Les Récepteurs peuvent cependant utiliser des RVC expressifs très subjectifs (présence du *je*, intensité, etc.) dans certains cas. Il est ainsi probable que les locuteurs s'inscrivant dans une situation de forte proximité communicative éprouvent d'un côté un besoin moindre de faire part d'originalité lorsqu'ils sont en position de Récepteurs, mais aussi un besoin d'émettre des réactions plus personnelles, parfois intenses.

Ces différentes observations et hypothèses, qui ne pourront être testées qu'avec plus de données, viennent une nouvelle fois nuancer notre première hypothèse d'une inscription des régulateurs au sein des éléments linguistiques typiquement liés à la proximité. Nous avons observé (section 2. 1. de ce chapitre) que ceux-ci sont au contraire plus fréquents dans la tranche de proximité assez faible, en particulier ceux ayant fonction de marque d'enregistrement. L'analyse des formes lexicales des RVC nous permet de préciser ces observations. En effet, nous faisons l'hypothèse, d'après nos données, que plus le score de proximité est faible (-02 à -10 points dans notre corpus), plus l'usage des RVC est rituel, lié

à la politesse. Au contraire, un score de proximité plus fort (-11 à -19 points dans notre corpus) paraît lié à une variation lexicale plus importante des RVC, avec un style moins soutenu et surtout une subjectivité plus importante. Curieusement, ce fait paraît d'autant plus important au sein de la tranche de scores assez forte, possiblement du fait d'un besoin plus important d'originalité et de connivence afin d'augmenter le taux de proximité.

Pour conclure, les différentes données extraites de notre corpus nous permettent de nuancer l'hypothèse émise initialement dans notre état de l'art (p. 21 de ce mémoire) d'un lien fort entre la présence de régulateurs et la proximité communicative. De fait, les régulateurs (ou du moins, les RVC) ne sont pas uniquement utilisés dans une situation de forte proximité communicative, mais aussi dans des situations tendant relativement vers la distance (score de proximité faible et assez faible). Il est de plus possible d'observer une différence d'utilisation des RVC selon le taux de proximité, tant en termes de types fonctionnels que de contenu lexical. Il apparaît en effet que les marques d'enregistrement, ainsi qu'un registre général plus soutenu, une utilisation plus rituelle, prévalent dans les tranches de scores plus faibles (taux faible à assez faible). En revanche, les types fonctionnels expressifs, en particulier les marques d'évaluation, sont plus retrouvés dans les tranches de score plus élevées, et permettent aux locuteurs de produire des RVC plus variés et surtout plus personnels.

Il semble ainsi que de manière générale, malgré la faible proportion de mots formant des RVC par rapport au total de mots dans ce corpus, il existe effectivement une certaine variation à la fois formelle et fonctionnelle des RVC. Cette variation s'exprime dans les taux généraux de différents types au sein des RVC (plus de marques d'enregistrement et de marques d'évaluation globalement), dans l'observation des différents scores généraux et particuliers de proximité, ainsi que dans les différentes formes des RVC. En effet, il semble qu'il existe une forte variation formelle au sein des RVC, avec de nombreuses combinaisons de particules satellites et d'éléments lexicaux possibles (des combinaisons impliquant jusqu'à sept unités ayant été retrouvées). Cette variation paraît d'autant plus importante au sein des RVC expressifs, avec un inventaire de formes plus conséquents et un très fort pourcentage de formes uniques (plus de 90% pour les deux sous-types). De plus, la proximité semble jouer un certain rôle quant au contenu lexical des RVC dans notre corpus, avec une différence d'usage entre les tranches de scores plus faibles et les plus fortes.

Une explication plus poussée et une meilleure compréhension de la variation des formes des RVC nécessite cependant une analyse plus détaillée, s'intéressant aussi aux

unités lexicales centrales et aux structures syntaxiques généralement retrouvés. Une telle analyse sera proposée dans la section suivante (point 4), ainsi que quelques observations sur les doubles valeurs retrouvées dans le corpus.

4. Analyse détaillée de la variation formelle des RVC selon les types et sous-types fonctionnels

Ce dernier point dédié à l'analyse du corpus nous permettra de revenir sur des aspects plus spécifiques des RVC. Nous nous intéresserons ainsi, par type fonctionnel, aux unités lexicales centrales⁵⁰ majoritairement retrouvées dans le corpus, ainsi qu'aux structures syntaxiques les plus fréquentes. Les réflexions que nous tirerons de cet examen nous permettront de revenir sur le lien entre forme et fonction, et l'importance qu'il revêt pour les *régulateurs verbaux de connivence*. Les tableaux reprenant les différentes unités lexicales et structures syntaxiques du corpus sont disponibles en annexe 3, p. 133.

4.1. Marques d'enregistrement

Au niveau des marques d'enregistrement, il est frappant d'observer qu'elles sont uniquement basées sur les unités lexicales *d'accord* et *OK*, ou bien sur une combinaison des deux. Il semble donc que l'usage de *d'accord* et *OK* soit un usage rituel, et que seuls ces deux unités lexicales puissent remplir la fonction de marque d'enregistrement. Cette notion de « rituel » est renforcée par la forte fréquence des marques d'enregistrement (voir point 2 de ce chapitre).

L'usage de *d'accord* et *OK* uniquement pourrait de plus s'expliquer par la forte désémantisation de ces deux unités lexicales, qui sont souvent citées comme marqueurs discursifs⁵¹ dans la littérature. En effet, cette forte désémantisation semble tout à fait adaptée aux marques d'enregistrement, qui servent minimalement d'« accusé de réception », et

⁵⁰ C'est-à-dire les unités lexicales sur lesquelles le RVC est construit, par exemple *OK* pour *OK*, mais *chaud* pour *c'est chaud quand même*.

⁵¹ Lire par exemple le numéro 154 de la revue *Langue Française* (2007) pour des articles dédiés aux marqueurs discursifs, avec quelques références à *OK* et *d'accord*

n'offrent aucune indication sur l'interprétation et l'opinion du Récepteur sur le discours entendu (engagement personnel minimal).

4.2. Marques de confirmation

Pour ce qui est des marques de confirmation, il est important de noter que deux des unités lexicales les plus retrouvées (*cool* et *chaud*, trouvées dans 26,96% et 11,54% des RVC de ce type) présentent systématiquement une double valeur avec les marques d'évaluation (type expressif). L'exemple (10) ci-dessous est un tel cas de double valeur.

(10)

Exemple 10 *c'est chaud* avec une double valeur marque de confirmation - marque d'évaluation (expressif)

(extrait de *armee_mer_15*, TCOF)

- | | | |
|---|----------------|--|
| 1 | -04_TCOF_01_02 | ils ont juste laissé notre section (0.6) |
| 2 | | au cas où (2.2) |
| 3 | | donc euh bah les u- musulmans ils en ont profité |
| 4 | | quand ils ont vu qu'on était plus beaucoup quoi |
| 5 | | pour garder les lieux (2.2) |
| 6 | -02_TCOF_01_01 | hum (0.5) c'est chaud |
| 7 | -04_TCOF_01_02 | donc voilà hein |

Les doubles valeurs marque de confirmation – marque d'évaluation constituent un cas très particulier. Il s'agit en effet de cas où l'unité lexicale en elle-même (*cool* et *chaud* étant les deux unités majoritairement concernées) contient une connotation évaluative et affective, et serait donc attendue comme marque d'évaluation. Cependant, il arrive que ces RVC soient prononcés sur un ton factuel, qui ne renvoie pas le caractère expressif des marques d'évaluation. Ils semblent par conséquent correspondre à des marques de confirmation dans de tels contextes : le Récepteur indique qu'il a bien reçu l'information et la ratifie en allant dans le sens du Locuteur, avec un faible degré d'engagement personnel dans la conversation. Dans la totalité du corpus, seuls 36,36% des *cool* et 57,14% des *chaud* possèdent uniquement une valeur de marque d'évaluation, au lieu d'une double valeur avec les marques de confirmation. Ces deux unités lexicales sont de plus parmi celles les plus fréquemment retrouvées dans le corpus, derrière *d'accord* et *OK* : sur les 43 unités lexicales centrales retrouvées dans le corpus, *cool* est la troisième plus fréquente et *chaud* la quatrième

plus fréquente⁵². Ces deux faits vont dans le sens, selon nous, d'une désémantisation et d'une sorte de « routinisation » de ces deux unités lexicales, avec une perte d'expressivité.

Après ces doubles valeurs, les unités lexicales centrales les plus fréquentes au sein des marques de confirmation sont les adjectifs *vrai* (15,38% des cas), *ex-aequo* avec *clair*, et *sûr* (7,69% des cas). Ces adjectifs ne possèdent pas de connotation affective, et semblent parfaitement adaptés aux marques de confirmation en ce qu'ils permettent de ratifier la parole du locuteur, de la confirmer, sans sentiment particulier. Ils sont utilisés dans la structure *c'est* + *ADJ* dans la majorité des cas.

4.3. Marques de convergence

Au sein des marques de convergence, les unités lexicales centrales les plus fréquentes sont *étonner* (uniquement retrouvée dans « tu m'étonnes », 22,22% des cas), *grave* (14,81% des cas) et *normal* (11,11% des cas). Il semble qu'elles permettent, tout comme les adjectifs *vrai*, *clair* et *sûr* pour les marques de confirmation, d'aller dans le sens du Locuteur et de le confirmer dans ces propos, mais cette fois-ci de façon plus abstraite et surtout plus subjective. Les structures syntaxiques les plus productives sont les adverbes seuls, les marqueurs discursifs (*grave*), les conjugaisons et les structures basées sur les adjectifs (*ADJ* seul, *c'est* + *ADJ* ou intensifieur + *ADJ*).

Les particules satellites sont particulièrement intéressantes à observer au sein des marques de convergence. En effet, deux tiers des RVC de ce type sont formés avec au moins une particule satellite. Parmi celles-ci, les plus fréquentes sont *ben/bah* et *oui/ouais* (observées avant ou après l'unité lexicale principale dans ce dernier cas). On retrouve aussi des *mais* et *hein*, par exemple. Il semble que ces différentes particules satellites permettent d'appuyer la parole du locuteur, en renforcement de la sémantique de l'unité lexicale centrale.

Ces particules satellites paraissent de plus jouer un rôle pour transformer une construction lexicale généralement associée au type fonctionnel marque d'évaluation, en une marque de convergence. Les deux exemples ci-dessous (11, 12) présentent de tels cas :

⁵² En termes de pourcentages, *cool* est ainsi retrouvé dans 5,39% des RVC du corpus et *chaud* dans 3,92% des RVC du corpus (ex-aequo avec *putain*).

(11)

Exemple 11 *super beau hein* en tant que marque de convergence (expressif) (extrait de Apéritif entre ami(e)s
- Glasgow, CLAPI)

- 1 -16_CLAPI_01_01 ouais et donc on est on a loué une voiture on a fait euh tsk deux jours là bas on a dormi euh dans un bi n bi avec la vue sur euh sur quoi d'ailleurs ((rires))
- 2 -16_CLAPI_01_02 sur la mer l'océan
- 3 -16_CLAPI_01_01 (2.6) la mer (0.3) ouais c'était trop bien
- 4 -16_CLAPI_01_02 **super beau hein**
- 5 -16_CLAPI_01_01 c'était trop trop beau

(12)

Exemple 12 *mais oui c'est rigolo* en tant que marque de convergence (expressif) (extrait de Apéritif entre ami(e)s - Glasgow, CLAPI)

- 1 -16_CLAPI_01_01 non ça on avait pas fait exprès mais au lieu de tu vois au lieu de réagir genre oh on est désolé on était mort de rire
- 2 genre on se foutait de leur gueule alors que pas du tout c'est juste que tu vois c'était le
- 3 -16_CLAPI_01_02 **mais oui c'est rigolo**
- 4 -16_CLAPI_01_01 l'accumulation et au moment de payer *NNAAMMEE* il me dit tu as de la monnaie et je dis non

Ces deux exemples sont tirés de la même conversation, et sont produits à un moment où la locutrice -16_CLAPI_01_01 (désormais L1) partage un récit de son séjour en Écosse. Il se trouve que son interlocutrice, -16_CLAPI_01_02 (désormais L2), a aussi visité les mêmes endroits en Écosse auparavant. L2 partage ainsi l'expérience de L1, ce qui semble expliquer que son utilisation d'une forme évaluative (*super beau*), associée à *hein* pour montrer son accord, montre finalement sa convergence sur l'expérience de L1, au lieu d'être une simple évaluation. L'exemple (12) ne correspond pas à un partage d'expérience concret, cependant L2 se positionne bien dans cet exemple comme en accord avec l'expérience décrite par L1. Elle le montre en utilisant une forme évaluative correspondant à la description de L1 (qui accompagne son récit de rires un peu auparavant) : *c'est rigolo*, associé à *mais oui* pour souligner son accord. Il est de plus intéressant de noter que L1 produit un *tu vois* dans son discours juste avant le régulateur de L2 dans cet exemple (12).

Les particules satellites paraissent ainsi jouer un rôle particulièrement important au sein des marques de convergence, en appuyant précisément sur celle-ci, en soulignant l'accord du Récepteur avec le Locuteur.

4.4. Marques d'évaluation

Pour finir, les unités lexicales les plus retrouvées au sein des marques d'évaluation (désormais EME) sont *cool* (17,74% des cas) et *chaud* (11,29%), encore une fois⁵³, mais aussi *putain* (11,29%), *bien* (9,68%), ou encore *fou* (4,84%). Il est frappant de noter que ces unités lexicales sont toutes soit des interjections, soit des adjectifs avec une connotation appréciative ou dépréciative et possédant un fort aspect subjectif. Elles sont donc tout à fait adaptées pour permettre une évaluation personnelle et expressive des propos du Locuteur par le Récepteur.

Les structures syntaxiques observées sont très variées au sein des marques d'évaluation⁵⁴, avec un total de 19 structures différentes. Les structures basées sur des adjectifs restent particulièrement productives, puisqu'elles sont présentes dans 58,06% des EME. Les interjections, presque uniquement retrouvées au sein de ce type fonctionnel, forment aussi une part importante des EME. Les structures contenant une interjection correspondent ainsi à 20,97% des EME, et l'on retrouve également quelques cas particulièrement expressifs formés par l'addition d'une interjection et d'un adjectif. Un dernier point notoire en ce qui concerne les structures syntaxiques des EME est la présence régulière du déictique *ça*, en renforcement dans des structures contenant déjà *c'est* + *ADJ*. Ces structures avec *ça* représentent 16,13% des EME. Les intensifieurs sont aussi plus fréquents au sein de ce type fonctionnel, et sont inclus dans 8,06% des EME.

Ces différentes observations structurelles semblent fortement liées à la fois à la fonction d'évaluation des EME (majorité de structures basées sur des adjectifs, présence du déictique *ça*) ainsi qu'à leur aspect expressif et fortement subjectif (haut degré d'engagement personnel du Récepteur : présence d'interjections et d'intensifieurs).

⁵³ En tant que marques d'évaluations ces lexèmes n'ont cependant pas systématiquement une double valeur, comme on a pu le voir.

⁵⁴ Fait qui se reflète dans le nombre très important de formes uniques, de même que pour les marques de convergence (voir point 3 de ce chapitre).

4.5. D'autres détails sur les doubles valeurs

Nous avons déjà pu discuter du cas des doubles valeurs marques de confirmation/marques d'évaluation dans la section 4.2 ci-dessus. Il ne s'agit cependant pas des seules combinaisons de valeurs annotées dans notre corpus (voir tableau 23, p. 97). Il conviendra ainsi dans cette section de discuter des autres cas de doubles valeurs, en particulier de la combinaison marque de confirmation/marque de convergence.

Pour rappel, les doubles valeurs correspondent à des segments auxquels il était trop difficile d'assigner un type fonctionnel précis. Deux types fonctionnels différents étaient alors assignés à de tels segments (17 au total dans notre corpus). Dans la majorité des cas, le flou ayant causé l'association à deux types fonctionnels différents était lié à un ton ambigu, trop difficile à discerner. Dans certains cas la portée du RVC pouvait aussi être difficile à distinguer (c'est notamment le cas de l'unique double valeur marque de convergence/marque d'évaluation, voir exemple (13)). Il est cependant légitime de se demander si certaines doubles valeurs ne pourraient pas être liées à un problème dans la typologie que nous avons développée⁵⁵. Nous avons pu discuter le cas des doubles valeurs marques de confirmation/marques d'évaluation dans la section 4. 2, doubles valeurs qui paraissent plutôt liées à un changement linguistique qu'à un réel problème dans la typologie. Mais qu'en est-il des autres cas de doubles valeurs ?

Un seul segment présente une double valeur marque de convergence/marque d'évaluation :

(13)

⁵⁵ Il est d'autant plus important de se questionner à ce sujet étant donné le manque de consensus lors de l'accord inter-annotateurs (voir chapitre 6).

Exemple 13 *ah ouais carrément* en tant que marque de convergence/marque d'évaluation (expressifs) (extrait de coree_ghu_14, TCOF)

- 1 -09_TCOF_02_02 c'était super drôle aussi de payer parce que je payais en cash
 3 -10_TCOF_02_01 ((rires))
 4 -09_TCOF_02_02 donc euh si tu veux [c'était]
 5 -10_TCOF_02_01 [les sous] tu as dû [te sentir riche] ((rires))
 6 -09_TCOF_02_02 [on payait en mil- en] millions de wons
 euh vu la [différence]
 7 -10_TCOF_02_01 [**ah ouais carrément**] ((rires))
 8 -09_TCOF_02_02 et le truc c'est que là-bas i-

Cette double valeur est liée à une difficulté d'évaluer la portée du RVC, du fait de sa superposition avec le discours du Locuteur (-09_TCOF_02_02). En effet, *ah ouais carrément* pourrait aussi bien être une marque de convergence, à mettre en lien avec le segment *vu la différence* de -09_TCOF_02_02 (sa fonction serait alors de montrer la convergence, l'accord de la Réceptrice sur cette différence), ou bien être une marque d'évaluation portant sur toute la situation décrite par -09_TCOF_02_02, traitant cette situation comme extraordinaire. Le rire de la Réceptrice (-10_TCOF_02_01) n'aide pas à faire la distinction entre les deux, puisqu'il impacte la façon dont elle prononce ce RVC, le rendant peu intelligible. Le problème paraît ainsi effectivement provenir de la situation d'énonciation dans ce cas.

Deux RVC sont assignés à la fois une valeur de marque d'enregistrement et une valeur de marque d'évaluation. Ces deux RVC sont basés sur l'unité lexicale *d'accord*. L'exemple (14) ci-dessous illustre un de ces cas :

(14)

Exemple 14 *hum d'accord* en tant que marque d'enregistrement/marque d'évaluation (extrait de film_tre_15, TCOF)

- 1 -15_TCOF_04_02 je sais pas du coup j'ai passé le test ils me disaient C1 donc euh
 2 -15_TCOF_04_01 ah ouais mais euh normalement enfin non le niveau euh
 3 pour le pour les instits c'est B2
 4 -15_TCOF_04_02 **hum d'accord**
 5 -15_TCOF_04_01 mais euh non mais

La double valeur est, pour ces deux cas, liée à une prononciation ambiguë de *d'accord*, avec un ton qui n'est pas le ton factuel communément retrouvé pour les marques d'enregistrement. Le ton employé renvoie en effet plutôt à une utilisation évaluative de *d'accord* (ton impressionné, ou bien ton ambigu vexé/moqueur dans le cas de l'exemple

(14)). Ainsi, dans l'exemple (14), les deux locutrices discutent des niveaux de langue en Anglais (B2 et C1 en particulier). -15_TCOF_04_02 (désormais L2) explique que le niveau C1 n'est pas forcément inatteignable, puisqu'une autre personne et elle-même l'ont obtenu. -15_TCOF_04_01 (désormais L1) lui répond cependant que le niveau nécessaire pour être professeur des écoles est le B2 uniquement. L2 répond à ces informations par le RVC *hum d'accord*, et L1 continue ensuite sur sa lancée (ligne 5, *mais euh*). Le type fonctionnel de ce *hum d'accord* est cependant ambigu. En effet, il est possible d'interpréter le ton de L2 comme un ton moqueur et/ou vexé : le *hum* ressemble presque à un petit rire, et le *d'accord* est court, prononcé de manière sèche. Le fait que L1 revienne ensuite sur sa lancée avec un *non mais* participe aussi à cette interprétation. Il reste néanmoins très difficile de savoir s'il s'agit bien d'un cas de marque d'évaluation, ou bien d'un cas de marque d'enregistrement seulement avec ces données.

Pour finir, trois RVC possèdent une double valeur marque de confirmation/marque de convergence. Il s'agit des cas qui pourraient le plus être dus à un problème dans la typologie. En effet, ces deux types fonctionnels sont assez similaires, la majeure différence se situant dans la qualité expressive, plus engagée, des marques de convergence. A l'exception de cette qualité expressive pour les marques de convergence, ces types fonctionnels permettent en effet tous deux au Récepteur d'aller dans le sens du Locuteur, de sous-entendre un certain accord avec lui.

Pour vérifier le possible impact de la typologie sur ces doubles valeurs, nous nous sommes par conséquent intéressés aux pourcentages de RVC avec cette double valeur au sein de chacun de ces deux types fonctionnels. Ainsi, les doubles valeurs MC/EMC constituent 11,54% des marques de confirmation (20% si l'on exclue les doubles valeurs MC/EME), et 11,11% des EMC. Sans être très élevés, ces pourcentages restent suffisamment importants pour garder en tête la possibilité d'un problème de définition voire de catégorisation dans notre typologie, en particulier pour la catégorie des marques de confirmation. Le faible nombre de marques de confirmation sans double valeur (12 sur les 26 trouvés dans le corpus, c'est-à-dire moins de la moitié – 46,15%) dans notre corpus appuie de plus la nécessité de préciser voire redéfinir la catégorie des marques de confirmation en particulier à l'avenir.

Avant de conclure sur cette section de notre analyse et sur ce chapitre 9 en général, il est essentiel de revenir sur un point qui avait déjà pu être abordé dans notre état de l'art, c'est-à-dire la difficulté que pose la multifonctionnalité des régulateurs. La présence de

doubles valeurs et de cas indécidables dans notre corpus est en effet au moins en partie liée à la nature même des régulateurs, qui sont des éléments linguistiques et/ou comportementaux inscrits dans une interaction, une situation d'énonciation donnée. Des cas où le/la chercheur.se ne peut reconstituer le sens des segments analysés paraissent donc inévitables, même s'ils peuvent être réduits par différents procédés (voir chapitre 6). La multifonctionnalité des segments analysés, qui peuvent être des régulateurs mais aussi servir d'autres fonctions pragmatiques dans certains cas, ou bien avoir un sens référentiel dans d'autres, est le premier facteur de difficulté. À ce facteur s'ajoutent les problèmes liés à la nature des enregistrements audio (volume insuffisant ou au contraire trop important pour bien distinguer certains segments dans certains cas, superposition de parole, etc.). De plus, il est aussi nécessaire de rappeler dans notre cas que l'analyse a uniquement été effectuée sur des fichiers audio, et non vidéo (y compris, dans un esprit d'unité, pour le corpus CLAPI, qui proposait de tels fichiers vidéo). Certains indices multimodaux (regards, gestes, etc.) qui auraient pu nous aider dans notre interprétation des régulateurs verbaux de connivence n'ont par conséquent pas pu être pris en compte.

Pour conclure, il semble donc que les unités lexicales les plus fréquemment retrouvées par type et sous-type fonctionnel soient bel et bien adaptées aux fonctions en question. La variété lexicale selon les types de RVC paraît aussi faire écho à leurs fonctions. Ainsi les RVC expressifs, en particulier, présentent une forte variété lexicale qui peut être expliquée par la subjectivité de l'affect exprimé et le fort engagement du Récepteur, qui s'adapte au discours donné du Locuteur. Certaines unités lexicales ont cependant pu être annotées avec une double valeur dans notre corpus. Ces doubles valeurs paraissent liées à un phénomène de changement linguistique (désémantisation) dans certains cas (doubles valeurs marques de confirmation/marques d'évaluation), mais aussi dans d'autres cas à un problème d'intelligibilité du segment (ton ambigu). Les doubles valeurs marques de confirmation/marques de convergence, quant à elle, pourraient être liées à un problème de catégorisation dans notre typologie.

Chapitre 10. Discussion

Avant de conclure ce mémoire, il conviendra de discuter dans ce dernier chapitre des limites de notre travail, en revenant en particulier sur ce qui n'a pas pu être fait.

Les régulateurs sont un ensemble de comportements linguistiques très complexes à analyser, notamment parce qu'ils jouent un rôle au cœur des interactions. Une définition et un travail exhaustif sur ces comportements proposeraient donc des éléments définitoires et analytiques sur les notions suivantes : la multimodalité, la prosodie, les tours de paroles, les marqueurs discursifs et phraséologimes, etc. Un corpus conséquent, et plus particulièrement un travail massif d'annotation du corpus seraient nécessaires pour accomplir un tel projet. Les moyens et le temps disponibles pour ce travail de mémoire étant insuffisants pour accomplir une tâche aussi importante, plusieurs notions ont dû être expliquées rapidement, voire n'ont malheureusement pas pu être expliquées et analysées.

Concernant le travail que nous avons pu accomplir, il est donc nécessaire de rappeler qu'il s'agit uniquement d'un travail exploratoire, dédié à ouvrir la voie à d'autres travaux de typologisation et de prise en compte de la variation au sein des régulateurs. Pour ce travail exploratoire, seul un sous-ensemble spécifique de régulateurs a été choisi, et seuls les régulateurs rentrant dans ce sous-ensemble des *régulateurs verbaux de connivence* ont été analysés. Ce fait revêt toute son importance dans la partie quantitative de l'analyse. En effet, les données statistiques obtenues ne concernent que ce sous-ensemble de régulateurs. Il est donc possible de faire des observations sur la façon dont les différents types fonctionnels envisagés sont présents dans les interactions, mais il est très difficile de tirer des réelles conclusions sur le comportement des locuteurs de notre corpus en tant que Récepteurs. De fait, l'analyse proposée ne prend pas du tout en compte la variation et le style individuel des locuteurs. Certains d'entre eux, qui ne produisent aucun RVC ou très peu, sont en réalité très productifs en tant que Récepteurs, mais recourent plutôt à des régulateurs vocaux tels que *hum* ou *ouais*. L'influence du style propre à chaque locuteur est aussi notée dans plusieurs travaux au sujet des intensifieurs (Eiswirth 2020b, par exemple). Il aurait donc été pertinent, avec plus de temps et de moyens pour le faire, de s'intéresser à cette variation individuelle afin de mieux comprendre la façon dont les locuteurs utilisent les RVC en contexte (quels types de RVC utilisent-ils dans quel contexte ? ont-ils des choix lexicaux spécifiques au sein des différents types ? Recourent-ils plutôt aux RVC ou à des régulateurs vocaux ? etc.).

Une autre limitation que nous pouvons apporter à notre travail concerne la typologie fonctionnelle déployée. En effet, celle-ci n'a malheureusement pu être testée qu'une seule fois dans le cadre d'un accord inter-annotateurs, et présente donc, dans la version présentée ici, des aspects qui n'ont pas été revus par des personnes autres que l'auteur de ce mémoire. L'analyse proposée dans ce mémoire a cependant permis d'effectuer un premier test de cette typologie, test qui s'est avéré mitigé, comme nous avons pu le voir en traitant les doubles valeurs. Il apparaît en effet que la catégorie des marques de confirmation serait à revoir dans de futurs travaux (re-définition ou abandon de la catégorie).

Concernant l'analyse de la proximité dans le corpus, il est aussi primordial de rappeler que, si le choix a été fait d'établir des scores par conversation pour ce mémoire, le degré de proximité et de distance varie en réalité au sein même d'une conversation. Prenons un exemple rapide avec le paramètre 03 (émotionnalité forte / faible) pour illustrer ce fait : lors d'une seule et même conversation un locuteur peut raconter dans un premier temps une anecdote factuelle, avec un faible taux d'émotion, puis passer à une anecdote très personnelle ou bien à un débat engagé, avec une implication émotionnelle beaucoup plus importante. Une analyse plus fine aurait peut-être permis d'obtenir des résultats différents, plus sûrs que les résultats obtenus dans ce mémoire. De plus, les degrés de proximité établis pourraient être revus de façon plus approfondie, avec des échelles spécifiques à chaque paramètre et des critères mieux définis, plus stables que ceux que nous avons utilisés dans le présent mémoire.

Vis-à-vis de l'examen des données, les analyses qui auraient pu être effectuées en plus de celles proposées sont nombreuses. Ainsi, l'analyse formelle aurait pu être menée jusqu'au bout (en observant le comportement des particules satellites au sein de chaque type fonctionnel, par exemple), et une étude du contexte environnant aurait pu être mise au point. Les régulateurs étant, on l'a vu, ancrés dans leur situation d'énonciation, l'observation des tours de parole précédant et suivant les RVC aurait probablement permis d'obtenir des informations sur le rôle possible de phénomènes tels que les hésitations, la présence de phatiques, etc. sur l'apparition de RVC. L'étude de la prosodie des tours de parole environnants aurait aussi été pertinente en ce sens.

Pour résumer, le mémoire que nous avons proposé ici possède avant tout une visée exploratoire. De ce fait, il ne rend pas complètement compte de la complexité des RVC ou même des régulateurs en général. De plus, seuls les régulateurs constituant notre objet d'étude, les *régulateurs verbaux de connivence*, ont été annotés dans ce mémoire. Ce fait

réduit les comparaisons et les analyses possibles, telle que l'analyse du style individuel des locuteurs dans leur utilisation de régulateurs, ou encore de l'influence du contexte environnant sur la présence et sur le type de régulateurs retrouvés. Cela étant, il nous semble que le mémoire que nous proposons rempli son rôle de travail exploratoire, comme nous le reverrons une dernière fois dans la conclusion.

Conclusion

Les régulateurs constituent un ensemble de comportements linguistiques permettant à un locuteur se positionnant comme auditeur, ou Récepteur, de signifier à son interlocuteur détenant la parole son écoute, et son engagement dans la conversation. Si ces comportements sont étudiés internationalement depuis plus de quarante ans, ils sont loin d'avoir été traités de façon exhaustive dans la littérature française. En effet, peu de travaux francophones se concentrent sur des régulateurs particuliers et/ou s'intéressent aux différentes micro-fonctions que peuvent avoir les régulateurs.

Par le biais d'une approche onomasiologique, guidée par les données, nous avons ainsi souhaité aller dans le sens d'une plus grande prise en compte des micro-fonctions spécifiques des régulateurs en français hexagonal, et de leur variation par extension. Pour ce faire, nous avons défini et nous sommes concentrés sur un sous-ensemble spécifique de régulateurs : les *régulateurs verbaux de connivence*, qui possèdent un contenu lexical et font partie des régulateurs simples. L'écoute répétée de notre corpus, constitué de conversations orales tendant le plus possible vers le vernaculaire, nous a permis d'établir une typologie fonctionnelle de ces RVC valable pour le français hexagonal uniquement, que nous avons aussi pu lier à la littérature existante. Afin d'aller plus loin que la présentation simple de cette typologie, nous avons aussi proposé des analyses cherchant à examiner la façon dont les différentes fonctions dégagées étaient représentées dans le corpus. Nous nous sommes aussi intéressés au lien possible entre l'aspect conceptionnel des conversations (tendance vers la proximité ou la distance) et la variation des régulateurs verbaux de connivence, avec une hypothèse initiale de lien fort entre régulateurs et proximité. Les données extraites de notre corpus nous ont cependant emmenés à reconsidérer et à nuancer cette hypothèse, du fait d'une différence d'utilisation des régulateurs verbaux de connivence selon le degré de proximité (différence entre le taux faible et le taux fort). Le contenu lexical et les structures syntaxiques des régulateurs verbaux de connivence ont aussi pu être discutés, selon leur type fonctionnel, afin de faire ressortir de possibles liens entre ces formes et les fonctions envisagées dans notre typologie.

Si, comme nous l'avons vu, ce mémoire n'a donc pas pu rendre compte de toute la complexité des régulateurs et analyser plus d'éléments linguistiques ou contextuels pouvant influencer leur variation fonctionnelle, il a tout de même rempli son rôle de travail

exploratoire. En effet, le travail de typologisation proposé ici sur ce sous-ensemble spécifique des *régulateurs verbaux de connivence* ouvrira, nous l'espérons, la voie à d'autres travaux de typologisation portant sur des régulateurs spécifiques, et à une plus grande prise en compte de la variation au sein des régulateurs.

De tels travaux de typologisation pourraient ainsi porter non seulement sur les micro-fonctions des régulateurs, mais aussi sur leurs différentes macro-fonctions pragmatiques et interactionnelles, qui, d'après les données que nous avons recueillies lors de notre travail d'analyse, mériteraient aussi d'être catégorisées dans le détail. De fait, les régulateurs verbaux de connivence dans notre corpus, et ce peu importe leur type fonctionnel précis, présentaient plusieurs macro-fonctions qui nous paraissent relativement peu détaillées dans la littérature, avec entre autres une fonction de relance du Locuteur, et une fonction de synthèse du segment de conversation en cours. Par conséquent, même si un tel résultat paraît très difficile à atteindre du fait de la quantité de travail à fournir, il semble qu'une typologie complète des régulateurs prendrait en compte non seulement leurs différentes micro-fonctions selon la langue donnée, mais aussi leurs différentes macro-fonctions pragmatiques et interactionnelles possibles.

Bibliographie

Arstein, R. & Poesio, M. (2008). Inter-coder agreement for computational linguistics. *Computational linguistics*, 34 (4), 555-596. Consulté sur : <https://doi.org/10.1162/coli.07-034-R2>

Bertrand, R., Ader, M., Blache, P., Ferré, G., Espesser, R. et Rauzy, S. (2009). Représentation, édition et exploitation des données multimodales : le cas des backchannels du corpus CID. *Cahiers de linguistique*, 33 (2), 183-212. Consulté sur : <https://hal.archives-ouvertes.fr/hal-00380698> (Pagination de PDF)

Degand, L. et Simon, A. C. (2011). L'analyse en unités discursives de base : pourquoi et comment ? *Langue Française*, 170, 45-59. Consulté sur : <https://www.cairn.info/revue-langue-francaise-2011-2-page-45.htm>

Eiswirth, M. (2020a). *It's all about the interaction: listener responses as a discourse-organisational variable*. (Thèse de Doctorat, The University of Edinburgh). Consulté sur : <http://dx.doi.org/10.7488/era/64>

Eiswirth, M. (2020b). Increasing interactional accountability in the quantitative analysis of sociolinguistics variation. *Journal of pragmatics*, 170, 172-188. Consulté sur : <https://doi.org/10.1016/j.pragma.2020.08.018>

Gadet, F. (2007). *La variation sociale en français* (nouvelle édition revue et augmentée). Ophrys.

Gadet, F. (2017). L'oralité ordinaire à l'épreuve de la mise à l'écrit : ce que montre la proximité. *Langages*, 208, 113-129. Consulté sur : <https://www.cairn.info/revue-langages-2017-4-page-113.htm>

de Gaulmyn, M.-M. (1987). Les régulateurs verbaux : le contrôle des récepteurs. Dans J. Cosnier & C. Kerbrat-Orecchioni (dir.), *Décrire la conversation* (p. 203-224). Presses Universitaires de Lyon.

Gharbi, A. (2020). Les formules expressives de la conversation : description linguistique et équivalence en Arabe. *Lublin Studies in Modern Languages and Literature*, 44, 161-173. Consulté sur : <https://doi.org/10.17951/lsmll.2020.44.1.161-173>

Goodwin, C. (1986). Between and within: alternative sequential treatment of continuers and assessments. *Human Studies*, 9, 205-217. Consulté sur : <https://doi.org/10.1007/BF00148127>

Guerin, E. (2017). Eléments pour une approche communicationnelle de la variation. Dans H. Tyne, M. Bilger, C. Paul & E. Guerin (dir.), *La variation en question(s) : Hommage à Françoise Gadet* (p. 57-73). Peter Lang.

Halté, P. (2013). *Les marques modales dans les chats : étude sémiotique et pragmatique des interjections et des émoticônes dans un corpus de conversations synchrones en ligne*. (Thèse de doctorat, Université de Lorraine, Nancy, Université du Luxembourg). Consulté sur : <http://www.theses.fr/2013LORR0308>

Kerbarth-Orecchioni, C. (2001). Oui, Non, Si : un trio célèbre et méconnu. *Marges Linguistiques*, 2, 95-119. Consulté sur : http://www.revue-texto.net/Parutions/Marges/Marges_sommaire.html

Kleiber, G. (2006). Sémiotique de l'interjection. *Langages*, 161, 10-23. Consulté sur : <https://www.cairn.info/revue-langages-2006-1-page-10.htm>

Koch, P. & Oesterreicher, W. (2001). Gesprochene Sprache und geschriebene Sprache – Langage parlé et langage écrit. Dans G. Holtus, M. Metzeltin & C. Schmitt (eds.), *Lexikon der Romanistischen Linguistik. Band I, 2 : Methodology* (p. 584-627). Max Niemeyer Verlag.

Mahrer, R. (2019). Parler, écrire : “continuum communicatif” et rupture matérielle. *Pratiques*, 183-184. Consulté sur : <https://doi.org/10.4000/pratiques.6842>

Maynard, S. (1990). Conversation management in contrast: Listener response in Japanese and American English. *Journal of Pragmatics*, 14, 397-412. Consulté sur : [https://doi.org/10.1016/0378-2166\(90\)90097-W](https://doi.org/10.1016/0378-2166(90)90097-W)

Modicom, P.-Y. (2015). L'opposition oralité – scripturalité dans l'analyse de DISCOURS/TEXTES : une introduction au programme de Koch & Oesterreicher. *Atelier du groupe ÉLIS*. Paris : Université Paris-Sorbonne, déc. 2015. Consulté sur : <https://halshs.archives-ouvertes.fr/halshs-01242845v2>

De Nuchèze, V. & Coletta, J.-M. (Eds.) (2002). *Guide terminologique pour l'analyse des discours : lexique des approches pragmatiques du langage*. Peter Lang.

O'Keeffe, A. & Adolphs, S. Response tokens in British and Irish discourse: corpus, context and variational pragmatics. Dans P. Schneider & A. Barron (eds.), *Variational pragmatics: a focus on regional varieties in pluricentric languages* (p. 69-98). John Benjamins Publishing Company. Consulté sur : <https://doi.org/10.1075/pbns.178.05ok> (pagination du PDF)

Persson, R. (2014). *Ressources linguistiques pour la gestion de l'intersubjectivité dans la parole en interaction : Analyses conversationnelles et phonétiques*. (Thèse de Doctorat, Lund University). Consulté sur : [https://portal.research.lu.se/portal/en/publications/ressources-linguistiques-pour-la-gestion-de-lintersubjectivite-dans-la-parole-en-interaction--analyses-conversationnelles-et-phonetiques\(0986c104-9c34-4152-af32-63fe391b7f17\).html](https://portal.research.lu.se/portal/en/publications/ressources-linguistiques-pour-la-gestion-de-lintersubjectivite-dans-la-parole-en-interaction--analyses-conversationnelles-et-phonetiques(0986c104-9c34-4152-af32-63fe391b7f17).html)

Ruquoy, D. (1995). Situation d'entretien et stratégie de l'interviewer. Dans L. Albarello, F. Digneffe, J.-P. Hiernaux, C. Maroy, D. Ruquoy & P. de Saint-Georges (aut.), *Pratiques et méthodes de recherche en sciences sociales* (p. 59-82). Armand Colin.

Sacks, H., Schegloff, E. & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50 (4), 696-735. Consulté sur : <https://doi.org/10.1007/BF00148127>

Schegloff, E. (1982). Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. Dans D. Tannen (ed.), *Analyzing Discourse: Text and Talk* (p. 71-93). Georgetown University Press.

Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. John Benjamins.

Traverso, V. (2016). *Décrire le français parlé en interaction*. Editions Ophrys.

Tutin, A. (2019). Phrases préfabriquées des interactions : quelques observations sur le corpus CLAPI. *Cahiers de Lexicologie*, 114, 63-91.

Logiciels informatiques :

ELAN (Version 6.0) [Logiciel]. (2020). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Obtenu sur <https://archive.mpi.nl/tla/elan>

Mathet, Y., Widlöcher, A., & Métivier J.-P. (2015). The Unified and Holistic Method Gamma (γ) for Inter-Annotator Agreement Measure and Alignment. *Computational linguistics*, 41 (3), 437-79. Consulté sur : https://doi.org/10.1162/COLI_a_00227.

GammaSoftware (Version 2.0) [Logiciel]. Obtenu sur <https://gamma.greyc.fr/>.

Corpus :

Gadet, F. (dir.) (2017). *Les parlers jeunes dans l'Ile-de-France multiculturelle*. Ophrys.

(2019). *MPF* [Corpus]. ORTOLANG (Open Resources and TOols for LANGuage) - www.ortolang.fr, v3, <https://hdl.handle.net/11403/mpf/v3>.

Analyse et traitement informatique de la langue française - UMR 7118 (ATILF)
(2020). *TCOF : Traitement de Corpus Oraux en Français* [Corpus]. ORTOLANG (Open Resources and TOols for LANGuage) - www.ortolang.fr, v2.1, <https://hdl.handle.net/11403/tcof/v2.1>.

CLAPI : Corpus de Langue Parlée en Interaction [Corpus]. <http://clapi.icar.cnrs.fr>

Table des exemples

Exemple 1 <i>d'accord</i> et <i>hum</i> en tant que régulateurs (extrait de coree_ghu_14, TCOF)	7
Exemple 2 Extrait de la conversation politique_ham_15 du corpus TCOF	16
Exemple 3 <i>OK</i> comme marque d'enregistrement (type <i>continuer</i>) (extrait de voyage_hab_14, TCOF)	44
Exemple 4 <i>d'accord</i> comme marque d'enregistrement (type <i>assessment</i>) (extrait de voyage_hab_14, TCOF)	44
Exemple 5 <i>c'est clair</i> comme marque de confirmation (extrait de apéritif entre ami(e)s – pois, CLAPI)	45
Exemple 6 <i>tu m'étonnes</i> comme marque de convergence (expressif) (extrait de voyage_hab_14, TCOF)	47
Exemple 7 <i>oh putain</i> et <i>oh c'est une horreur</i> comme marques d'évaluation (expressifs) (extrait de tourisme_arn_15, TCOF)	49
Exemple 8 <i>OK</i> comme marque d'évaluation (expressif) (extrait de tourisme_arn_15, TCOF)	51
Exemple 9 Extrait d'un passage avec une très forte implication émotionnelle des participants	68
Exemple 10 <i>c'est chaud</i> avec une double valeur marque de confirmation - marque d'évaluation (expressif) (extrait de armee_mer_15, TCOF)	106
Exemple 11 <i>super beau hein</i> en tant que marque de convergence (expressif) (extrait de Apéritif entre ami(e)s - Glasgow, CLAPI).....	108
Exemple 12 <i>mais oui c'est rigolo</i> en tant que marque de convergence (expressif) (extrait de Apéritif entre ami(e)s - Glasgow, CLAPI).....	108
Exemple 13 <i>ah ouais carrément</i> en tant que marque de convergence/marque d'évaluation (expressifs) (extrait de coree_ghu_14, TCOF)	111
Exemple 14 <i>hum d'accord</i> en tant que marque d'enregistrement/marque d'évaluation (extrait de film_tre_15, TCOF).....	111

Table des illustrations et tableaux

Figure 1 Vue d'ensemble des approches interactionnistes du langage permettant l'appréhension et l'étude des régulateurs.....	13
Figure 2 Schéma récapitulatif des règles d'organisation de la parole dans une conversation (Sacks, Schegloff & Jefferson, 1974).....	17
Figure 3 Schéma de l'utilisation d'un régulateur dans une conversation.....	18
Figure 4 Schéma des étapes du processus d'analyse des données du corpus, en vue de l'élaboration de la typologie	59
Figure 5 Acteurs de la grille d'annotation selon leur relation hiérarchique (extrait du fichier ELAN traitant le corpus apéritif entre ami(es) - pois de CLAPI)	75
Figure 6 Annotation finalisée d'un RVC expressif de type marque d'évaluation (extrait du fichier ELAN traitant le corpus apéritif entre ami(e)s – glasgow de CLAPI)	77
Figure 7 Pourcentages des différents types et sous-types de RVC à l'intérieur des RVC annotés dans le corpus	84
Figure 8 Pourcentages des différentes doubles valeurs retrouvées dans le corpus, vis-à-vis du total des doubles valeurs et vis-à-vis du total des RVC.....	84
Figure 9 Pourcentages des mots formants des RVC dans le corpus.....	85
Figure 10 Histogramme des pourcentages de mots formant des RVC selon le taux de proximité dans le corpus	87
Figure 11 Histogramme des pourcentages de mots formant des RVC par type fonctionnel selon le taux de proximité dans le corpus	87
Figure 14 Histogramme des pourcentages de mots formant des RVC selon le taux de proximité au paramètre 02	92
Figure 15 Histogramme des pourcentages de mots formant des RVC par type fonctionnel selon le taux de proximité au paramètre 02	92
Figure 16 Pourcentages de formes uniques au sein des différents types et sous-types fonctionnels	98
*	
Tableau 1 Les paramètres situationnels de Koch & Oesterreicher (2001).....	20
Tableau 2 Résumé de quelques classifications fonctionnelles des régulateurs trouvées dans la littérature.....	27

Tableau 3 Classifications fonctionnelles des régulateurs par Maynard (1989) et Bertrand et al. (2009)	35
Tableau 4 Résumé de quelques classifications fonctionnelles des régulateurs trouvées dans la littérature (régulateurs complexes exclus).....	38
Tableau 5 Tableau synthétique des fonctions majoritairement retrouvées dans la littérature	39
Tableau 6 Présentation synthétique de la typologie des régulateurs verbaux de connivence	42
Tableau 7 Contenu du corpus final utilisé dans le cadre du présent mémoire	57
Tableau 8 Les paramètres situationnels de Koch & Cesterreicher, version de 2008 (p. 201, cité par Modicom, 2015, p. 4)	65
Tableau 9 Détails de la grille de score pour le paramètre 01 (parole privée/publique).....	66
Tableau 10 Détails de la grille de score pour le paramètre 02 (relation de familiarité entre les participants/les participants ne se connaissent pas)	67
Tableau 11 Détails de la grille de score pour le paramètre 03 ((faible) implication émotionnelle) .	68
Tableau 12 Détails de la grille de score pour le paramètre 04 (prévalence/effacement de la situation d'énonciation)	69
Tableau 13 Détails de la grille de score pour le paramètre 05 (proximité/éloignement des référents)	69
Tableau 14 Détails de la grille de score pour le paramètre 06 (proximité/distance spatio-temporelle)	70
Tableau 15 Détails de la grille de score pour le paramètre 07 (coopération/pas de coopération entre les participants).....	70
Tableau 16 Détails de la grille de score pour le paramètre 08 (dialogisme/monologisme)	71
Tableau 17 Détails de la grille de score pour le paramètre 09 (discours spontané/réfléchi).....	72
Tableau 18 Détails de la grille de score pour le paramètre 10 (développement thématique libre/fixe)	73
Tableau 19 Inventaire des RVC fonctionnant comme marques d'enregistrement dans le corpus....	95
Tableau 20 Inventaire des RVC fonctionnant comme marques de confirmation dans le corpus....	95
Tableau 21 Inventaire des RVC expressifs fonctionnant comme marques de convergence dans le corpus	96
Tableau 22 Inventaire des RVC expressifs fonctionnant comme marques d'évaluation dans le corpus	96

Tableau 23 Inventaire des RVC ayant reçu une double valeur dans le corpus	97
Tableau 24 Inventaire des RVC produits par les locuteurs avec un score de proximité faible	99
Tableau 25 Inventaire des RVC produits par les locuteurs avec un score de proximité assez faible	100
Tableau 26 Inventaire des RVC produits par les locuteurs avec un score de proximité assez fort	101
Tableau 27 Inventaire des RVC produits par les locuteurs avec un score de proximité fort	102
Tableau 28 Tableau récapitulatif reprenant le nom codé des participants et leurs noms originaux (participants + conversations).....	129
Tableau 29 Inventaire des RVC composés d'un seul élément lexical	131
Tableau 30 Inventaire des RVC composés de deux éléments lexicaux (+ satellites).....	131
Tableau 31 Inventaire des RVC composés de trois éléments lexicaux (+ satellites)	131
Tableau 32 Inventaire des RVC composés de quatre éléments lexicaux (+ satellites)	132
Tableau 33 Inventaire des RVC composés de cinq éléments lexicaux (+ satellites)	132
Tableau 34 Inventaire des RVC composés de six éléments lexicaux (+ satellites).....	132
Tableau 35 Inventaire des RVC composés de sept éléments lexicaux (+ satellites).....	132
Tableau 36 Inventaire des unités lexicales observées au sein des marques d'enregistrement dans le corpus	133
Tableau 37 Inventaire de unités lexicales observées au sein des marques de confirmation dans le corpus	133
Tableau 38 Unités lexicales observées au sein des marques de convergence (expressifs) dans le corpus	133
Tableau 39 Unités lexicales observées au sein des marques d'évaluation (expressifs) dans le corpus	134
Tableau 40 Inventaire des structures syntaxiques (particules satellites non prises en compte) observées dans le corpus pour les marques d'enregistrement	134
Tableau 41 Inventaire des structures syntaxiques (particules satellites non prises en compte) observées dans le corpus pour les marques de confirmation	134
Tableau 42 Inventaire des structures syntaxiques (particules satellites non prises en compte) observées dans le corpus pour les marques de convergence (expressifs).....	134

Tableau 43 Inventaire des structures syntaxiques (particules satellites non prises en compte) observées dans le corpus pour les marques d'évaluation (expressifs).....	135
---	-----

Table des annexes

Annexe 1 Noms codés des participants associés à leur nom et conversations originales	129
Annexe 2 Inventaire des formes observées des RVC selon le nombre d'éléments lexicaux + particules satellites.....	131
Annexe 3 Inventaires des unités lexicales et structures syntaxiques retrouvées dans le corpus (par type fonctionnel).....	133

Annexe 1

Noms codés des participants associés à leur nom et conversations originales

Tableau 28 Tableau récapitulatif reprenant le nom codé des participants et leurs noms originaux (participants + conversations)

Nom codé	Nom original conversation	Nom original participant
-16_CLAPI_01_01	Apéritif entre ami(e)s - glasgow	JUD
-16_CLAPI_01_02	Apéritif entre ami(e)s - glasgow	PAT
-15_CLAPI_02_01	Apéritif entre ami(e)s - pois	JUL
-15_CLAPI_02_02	Apéritif entre ami(e)s - pois	ROM
-17_CLAPI_02_03	Apéritif entre ami(e)s - pois	ANN
-16_CLAPI_03_02	Apéritif entre ami(e)s - rupture	ALB
-17_CLAPI_03_01	Apéritif entre ami(e)s - rupture	JUS
-17_CLAPI_03_03	Apéritif entre ami(e)s - rupture	ARN
-14_MPF_01_03	Aristide2b_ANON	Aristide
-19_MPF_01_01	Aristide2b_ANON	Daniel
-19_MPF_01_02	Aristide2b_ANON	Margot
-07_MPF_02_01	Aristide5b_ANON	Louis
-07_MPF_02_02	Aristide5b_ANON	Aristide
-11_MPF_02_03	Aristide5b_ANON	Joël
-02_TCOF_01_01	armee_mer_15	L1
-04_TCOF_01_02	armee_mer_15	L2
-08_TCOF_02_03	coree_ghu_14	L3
-09_TCOF_02_02	coree_ghu_14	L2
-10_TCOF_02_01	coree_ghu_14	L1
-15_TCOF_03_01	cours_mau_15	L1
-15_TCOF_03_03	cours_mau_15	L3
-16_TCOF_03_02	cours_mau_15	L2
-15_TCOF_04_01	film_tre_15	L1
-15_TCOF_04_02	film_tre_15	L2

-02_TCOF_05_02	how_pro_14	L2
-07_TCOF_05_01	how_pro_14	L1
-05_TOCF_06_02	nonno_crz_14	L2
-07_TOCF_06_01	nonno_crz_14	L1
-08_TOCF_06_03	nonno_crz_14	L3
-09_TCOF_07_01	paternel_gos_14	L1
-10_TCOF_07_02	paternel_gos_14	L2
-15_TCOF_08_01	politique_ham_15	L1
-15_TCOF_08_02	politique_ham_15	L2
-16_MPF_03_01	Roberto2c_ANON	Stéphane
-16_MPF_03_02	Roberto2c_ANON	Farid
-13_TCOF_09_01	showgirl_vos_15	L1
-13_TCOF_09_02	showgirl_vos_15	L2
-10_TCOF_10_02	tourisme_arn_15	L2
-13_TCOF_10_01	tourisme_arn_15	L1
-13_TCOF_10_03	tourisme_arn_15	L3
-11_TCOF_11_01	voyage_hab_14	L1
-12_TCOF_11_02	voyage_hab_14	L2

Annexe 2

Inventaire des formes observées des RVC selon le nombre d'éléments lexicaux + particules satellites

Tableau 29 Inventaire des RVC composés d'un seul élément lexical

RVC composés d'un seul élément lexical (26)			
OK	(15)	abusé	(1)
cool	(5)	putain	(1)
grave	(3)	top	(1)

Tableau 30 Inventaire des RVC composés de deux éléments lexicaux (+ satellites)

RVC composés de 2 éléments lexicaux + vocaux (97)					
d'accord	(77)	ah: chouette:	(1)	hum bien	(1)
oh putain	(3)	ben normal	(1)	oh punaise	(1)
l'enfer	(2)	bien sûr	(1)	putain atroce	(1)
ah bien	(1)	d'acc	(1)	putain ouais	(1)
ah OK	(1)	exactement ouais	(1)	tant mieux	(1)
ah putain	(1)	han putain	(1)	très utile	(1)

Tableau 31 Inventaire des RVC composés de trois éléments lexicaux (+ satellites)

RVC composés de 3 éléments lexicaux + vocaux (34)					
tu m'étonnes	(5)	ah tout pourri	(1)	d'accord OK	(1)
ah d'accord	(4)	bah ouais forcément	(1)	je te jure	(1)
hum d'accord	(4)	ben grave ouais	(1)	oh la vache	(1)
c'est clair	(3)	c'est ça	(1)	ouais d'accord	(1)
c'est chaud	(2)	c'est malheureux	(1)	ouais j'avoue	(1)
ah la vache	(1)	c'est normal	(1)	super beau hein	(1)
ah ouais carrément	(1)	c'est vrai	(1)	truc de fou	(1)

Tableau 32 Inventaire des RVC composés de quatre éléments lexicaux (+ satellites)

RVC composés de 4 éléments lexicaux + vocaux (27)					
ouais c'est vrai	(3)	c'est cool ça	(1)	hum c'est cool	(1)
ah c'est chaud	(2)	c'est fou ça	(1)	hum: chaud quand même	(1)
ah ouais d'accord	(2)	c'est la merde	(1)	oh oh oh putain	(1)
bah c'est cool	(1)	c'est nul ça	(1)	ouais c'est clair	(1)
bah j'imagine ouais	(1)	c'est pas cool	(1)	ouais c'est sûr	(1)
bah oui forcément ouais	(1)	c'est pas faux	(1)	ouais je te jure	(1)
bah tu m'étonnes	(1)	c'est trop affreux	(1)		
c'est affreux quoi	(1)	d'accord ouais ouais	(1)		

Tableau 33 Inventaire des RVC composés de cinq éléments lexicaux (+ satellites)

RVC composés de 5 éléments lexicaux + vocaux (11)					
ah bah c'est bien	(1)	c'est chaud quand même	(1)	ouais c'est cool ouais	(1)
ah c'est trop bien	(1)	mais oui c'est rigolo	(1)	oui oui c'est normal	(1)
ben ouais c'est ça	(1)	oh c'est une horreur	(1)	roh putain c'est affreux	(1)
ça doit faire trop bizarre	(1)	ouais c'est bizarre hein	(1)		

Tableau 34 Inventaire des RVC composés de six éléments lexicaux (+ satellites)

RVC composés de 6 éléments lexicaux + vocaux (6)					
ah ouais c'est chiant ça	(1)	mais bien sûr ouais ouais ouais	(1)		
ah ouais c'est fou ça	(1)	ouais ça doit être chaud hein	(1)		
hum ah c'est bien ça	(1)	ouais c'est pas cool ça	(1)		

Tableau 35 Inventaire des RVC composés de sept éléments lexicaux (+ satellites)

RVC composés de 7 éléments lexicaux + vocaux (3)	
ah ben ça c'est chouette ça	(1)
ah mais c'est trop bien ça	(1)
ouais c'est clair c'est clair	(1)

Annexe 3

Inventaires des unités lexicales et structures syntaxiques retrouvées dans le corpus (par type fonctionnel)

Les résultats présentés ci-dessous n'ont pas été transformés en pourcentages, et indiquent directement le nombre d'occurrences trouvées dans le corpus. Pour rappel le total de RVC trouvés dans le corpus est de 204, 221 en comptant les doubles valeurs une fois par type, comme c'est le cas dans les données présentées ci-dessous. L'abréviation MD désigne les marqueurs discursifs.

Tableau 36 Inventaire des unités lexicales observées au sein des marques d'enregistrement dans le corpus

Unités lexicales utilisées dans les marques d'enregistrement	
d'accord	90
OK	15
d'accord + OK	1
Total	106

Tableau 37 Inventaire de unités lexicales observées au sein des marques de confirmation dans le corpus

Unités lexicales utilisées dans les marques de confirmation			
cool	7	jurer	1
vrai	4	faux	1
clair	4	bien	1
chaud	3	normal	1
sûr	2	anaphore (ça)	1
forcément	1	Total	26

Tableau 38 Unités lexicales observées au sein des marques de convergence (expressifs) dans le corpus

Unités lexicales utilisées dans les marques de convergence (expressifs)					
étonner	6	chaud	1	putain	1
grave	4	clair	1	rigolo	1
normal	3	exactement	1	sûr	1
anaphore (ça)	1	forcément	1	vrai	1
avouer	1	imaginer	1	carrément	1
beau	1	jurer	1	Total	27

Tableau 39 Unités lexicales observées au sein des marques d'évaluation (expressifs) dans le corpus

Unités lexicales utilisées dans les marques d'évaluation (expressifs)							
cool	11	chouette	2	horreur	1	punaise	1
chaud	7	d'accord	2	jurer	1	putain + affreux	1
putain	7	enfer	2	malheureux	1	putain + atroce	1
bien	6	la vache	2	merde	1	tant mieux	1
fou	3	abusé	1	nul	1	top	1
affreux	2	carrément	1	OK	1	utile	1
bizarre	2	chiant	1	pourri	1	Total	62

Tableau 40 Inventaire des structures syntaxiques (particules satellites non prises en compte) observées dans le corpus pour les marques d'enregistrement

Structures syntaxiques utilisées au sein des marques d'enregistrement	
MD	104
MD raccourci	1
MD + MD	1
Total	106
Total de structures différentes	3

Tableau 41 Inventaire des structures syntaxiques (particules satellites non prises en compte) observées dans le corpus pour les marques de confirmation

Structures syntaxiques utilisées au sein des marques de confirmation			
MD	4	c'est + NEG + Adj	1
Adv	1	Adj	1
conj	1	c'est + ça	1
INT + Adj	1	Total	26
c'est + Adj	16	Total de structures différentes	8

Tableau 42 Inventaire des structures syntaxiques (particules satellites non prises en compte) observées dans le corpus pour les marques de convergence (expressifs)

Structures syntaxiques utilisées au sein des marques de convergence (expressifs)			
MD	4	c'est + Adj x2	1
Adv	3	c'est + ça	1
conj	9	interjection	1
Adj	1	phraséologisme	1
INT + Adj	2	Total	27

c'est + Adj	4	Total de structures différentes	10
-------------	---	---------------------------------	----

Tableau 43 Inventaire des structures syntaxiques (particules satellites non prises en compte) observées dans le corpus pour les marques d'évaluation (expressifs)

Structures syntaxiques utilisées au sein des marques d'évaluation (expressifs)					
MD	8	c'est + NEG + Adj	1	interjection + c'est + Adj	2
Adv	1	c'est + Adj + ça	6	phraséologisme	1
conj	1	c'est + INT + Adj + ça	1	Pro + Nom	2
Adj	6	c'est + NEG + Adj + ça	2	c'est + Pro + Nom	2
INT + Adj	2	ça + c'est + Adj + ça	1	Nom + de + Nom	1
c'est + Adj	12	interjection	10	Total	62
c'est + INT + Adj	2	interjection + Adj	1	Total de structures différentes	19

Table des matières

Remerciements	3
Sommaire	5
Introduction	7
PARTIE 1 - ÉTAT DE L'ART	9
CHAPITRE 1. APPROCHES THEORIQUES ET METHODOLOGIQUES.....	10
CHAPITRE 2. QUELQUES FONDEMENTS NECESSAIRES A L'APPREHENSION DU FONCTIONNEMENT DES REGULATEURS : L'ORGANISATION DE LA PAROLE DANS LES CONVERSATIONS ET LA DIMENSION CONCEPTIONNELLE DU LANGAGE.....	14
CHAPITRE 3. APPROCHE ET DEFINITION DES <i>REGULATEURS VERBAUX DE CONNIVENCE</i>	23
PARTIE 2 - METHODOLOGIE.....	53
CHAPITRE 5. CHOIX ET ELABORATION DU CORPUS	54
CHAPITRE 6. ÉLABORATION DE LA TYPOLOGIE.....	59
CHAPITRE 7. ANNOTATION DU CORPUS	64
CHAPITRE 8. TRAITEMENT DES DONNEES OBTENUES	79
PARTIE 3 - ANALYSE ET DISCUSSION DES RESULTATS	82
CHAPITRE 9. RESULTATS DE L'ANALYSE DU CORPUS.....	83
CHAPITRE 10. DISCUSSION	114
Conclusion.....	117
Bibliographie.....	119
Table des exemples	123
Table des illustrations et tableaux	124
Table des matières	136

MOTS-CLÉS : *Backchannel*, Régulateurs verbaux de connivence, interactions, typologie fonctionnelle, variation

RÉSUMÉ

Le contact et l'engagement des locuteurs est primordial au sein des interactions, et ce y compris lorsque seul un des locuteurs s'exprime. Les *backchannel*, ou régulateurs en Français, font partie des comportements d'écoute permettant de manifester cet engagement sans reprendre la parole. S'il existe différents types de régulateurs, leurs micro-fonctions ont peu été étudiées en détail. Ce mémoire tente de remédier à ce fait, en proposant une typologie fonctionnelle d'un sous-ensemble de régulateurs : les régulateurs verbaux de connivence, soit des régulateurs simples possédant un contenu lexical. La typologie et les résultats obtenus sont basés sur l'étude d'un corpus oral de Français hexagonal. En plus de la typologie proposée, ce mémoire présente des résultats concernant la variation des micro-fonctions envisagées selon le continuum conceptionnel de la proximité et de la distance (Koch & Œsterreicher, 2001), et s'interroge aussi sur le lien entre la forme et la fonction de ces régulateurs.

KEYWORDS : *Backchannel*, Régulateurs verbaux de connivence, interaction, functional classification, variation

ABSTRACT

Speakers need to be constantly engaged in an interaction for it to truly be *interactive*. For this reason, speakers show *backchannel* behaviours just as they are only listening, to exhibit this continued attention. Different types of backchannels (*régulateurs* in French) exist, however few studies worked on their specific micro-functions. To rectify this fact, this thesis offers a functional classification of a subset of backchannels: the *régulateurs verbaux de connivence*, simple backchannels with a lexical content. This classification is based on the study of an oral corpus of hexagonal French conversations. This thesis also provides results as to the variation of the micro-functions outlined according to the conceptional content of the conversation, between proximity and distance (Koch & Œsterreicher, 2001). A special interest is also taken in discussing the link between form and function for this subset of backchannel behaviour.