

# **Cerner les marqueurs de la spontanéité énonciative : choix méthodologiques et apport de données expérimentales**

Identifying markers of enunciative spontaneity in spoken French:  
methodological orientations and experimental data

Isabelle Stabarin<sup>1</sup>

**Abstract:** What is enunciative spontaneity and how does it manifest itself in language? After having laid some theoretical and methodological milestones, we report on the experimental aspects of our research on markers of spontaneity in spoken French. We illustrate how a corpus recorded according to a new protocol allows us to compare semantically equivalent predications, uttered with different degrees of spontaneity. The results obtained lead to a first inventory of the formal aspects of spontaneity and show that linguistic complexity is one of its main markers.

**Key words:** enunciative spontaneity, definitional matrix analysis (AMD), linguistic complexity, experimental corpus, solicited reformulations.

## **1. Introduction**

Les travaux qui se développent autour de l'oral spontané – ou français parlé spontané – sont de plus en plus nombreux, dans pratiquement tous les domaines des sciences du langage. Cependant, si ces deux appellations sont aujourd'hui largement plébiscitées par les chercheurs, elles se réfèrent à des corpus hétérogènes et pris dans des contextes très variés. Il nous semble donc opportun de se pencher sur ces étiquettes. Nos interrogations concernent tout d'abord l'adjectif spontané : désormais, son emploi semble aller tellement de soi qu'aucune étude, à notre connaissance, n'en a approfondi la définition. Mais qu'est-ce au juste que la spontanéité en langue ? Et surtout : quels sont ses aspects formels ?

C'est sur ces interrogations que se fonde notre recherche. D'abord, pour essayer de combler en partie ce que nous percevons comme un manque à un niveau théorique. Ensuite, parce que nous

---

<sup>1</sup> Università degli Studi di Trieste; [isabelle.stabarin@deams.units.it](mailto:isabelle.stabarin@deams.units.it).

croyons qu'étudier la langue à travers le prisme de la spontanéité peut faire émerger des éléments nouveaux pour la description du français et de son fonctionnement dans la communication.

Cet article présente le parcours relatif à cette recherche sur les marqueurs linguistiques de la spontanéité en français parlé<sup>2</sup>, dans ses aspects expérimentaux. Nous y exposons le cadre théorique adopté, soit l'Analyse matricielle définitoire, ainsi que les outils méthodologiques qu'elle comprend, et leurs potentialités dans ce contexte. Sur cette base, nous formulons l'hypothèse selon laquelle la complexité syntaxique est un marqueur de spontanéité. Nous illustrons aussi, plus particulièrement, comment un corpus *ad hoc* a été recueilli, face aux limites constatées dans l'exploration d'autres corpus. Ce nouveau corpus, enregistré selon un protocole inédit, permet de comparer des prédications sémantiquement équivalentes, énoncées à des degrés différents de spontanéité par un locuteur donné et de vérifier notre hypothèse de recherche. Nous proposons enfin un échantillon des résultats obtenus menant à un premier recensement des aspects formels de la spontanéité énonciative.

## **2. La spontanéité énonciative : perspectives définitionnelles**

Nous appréhendons la spontanéité énonciative selon deux perspectives, par ailleurs complémentaires.

La première perspective est *qualitative* : la spontanéité caractérise une énonciation non préparée, ou « non préméditée » et un énoncé oral « conçu et perçu dans le fil de son énonciation » (Luzzati 2007 : 25). Di Cristo (2016) souligne le caractère « impromptu » des énoncés produits avec spontanéité, Traverso (2016) montre bien, au moyen des corpus d'oral conversationnel qu'elle analyse, la part d'improvisation, l'instantanéité, l'immédiateté qui les définissent. « L'oral spontané » s'oppose dans cette première acception à « l'oral préparé ». L'oral préparé, souvent un écrit oralisé, peut être fictionnel (ex. le dialogue d'une série) ou au contraire ancré dans la vie réelle (ex. : un discours officiel), ces deux types d'énonciation pouvant aussi s'entrecouper<sup>3</sup>. Parmi les différents types de discours oraux, c'est la conversation à bâtons rompus, où l'interaction est dense, qui constitue l'exemple prototypique de l'oral spontané : à chaque tour de parole les interlocuteurs doivent réagir pour continuer à communiquer.

La seconde perspective est *quantitative* : le locuteur qui interagit dans une énonciation *très* spontanée laisse libre cours à son élan communicationnel sans se soucier de la forme de son

<sup>2</sup> Thèse soutenue à Sorbonne université en novembre 2019 (directeurs successifs A. H. Ibrahim et C. Martinot).

<sup>3</sup> C'est le cas si un acteur interrompt sa récitation pour faire une considération personnelle et change de plan d'énonciation, ou si un homme politique quitte un discours bien rodé pour répondre à une remarque ou à une question inattendue.

discours. Il peut au contraire réfréner son élan communicationnel/énonciatif pour exercer un contrôle sur sa propre énonciation et sur les énoncés déjà produits, modalité qui requiert un effort cognitif plus important. C'est l'aspect *scalaire* de la spontanéité que recouvre ce deuxième aspect, puisque des degrés différents de spontanéité peuvent s'observer en fonction de la posture plus ou moins vigilante du locuteur. Cette variation potentielle de la spontanéité génère des variations linguistiques intrapersonnelles, parfois même au sein de la même énonciation. Elle peut impliquer une pluralité de registres de langue, mais ce n'est pas systématique (Blanche-Benveniste 1997, Depoux & Stabarin 2017, Stabarin 2017, Depoux 2017).

Dans la première perspective, les marqueurs de spontanéité sont les signes de l'immédiateté énonciative, tels que les disfluences du discours en élaboration. Il s'agit de ruptures sur l'axe syntagmatiques comme les répétitions, troncations, fillers et faux-départs (Bouraoui 2008, Bazillon 2011, Dutrey 2014). Dans la seconde perspective, les marqueurs de spontanéité témoignent du degré de spontanéité *vs* contrôle. Et ils font partie du niveau morphosyntaxique de la langue. C'est cette perspective que nous privilégions pour notre recherche des marqueurs de spontanéité.

Ayant caractérisé la spontanéité comme un élan communicationnel /énonciatif, qui génère un discours non prémédité, nous nous sommes demandé ce qui peut freiner cet élan et le modifier. Nous avons pu constater, au fur et à mesure que ce travail progressait, que les *facteurs de contrôle* sont nombreux et très variés. Principalement de nature linguistique, environnementale, sociale ou psychologique, ces facteurs se combinent et ne sont de la sorte pas toujours clairement identifiables. Nous retenons alors de manière plus générale le concept de *marché linguistique*, développé par Pierre Bourdieu, en l'interprétant comme un grand facteur de contrôle qui peut en recouvrir d'autres : « Il y a marché linguistique toutes les fois que quelqu'un produit un discours à l'intention de récepteurs capables de l'évaluer, de l'apprécier et de lui donner un prix » (Bourdieu 1982 : 122).

### **3. Cadres et hypothèse de recherche**

#### **3.1. L'analyse matricielle définitoire (AMD)**

Le cadre théorique et méthodologique qui soutient en grande partie notre travail est l'Analyse matricielle définitoire (AMD) développée par Amr Helmy Ibrahim à partir de 1994. Ce cadre s'est révélé comme étant tout à fait approprié à notre recherche. Il nous a permis aussi bien de formuler notre hypothèse de recherche que de la vérifier à l'aide des outils méthodologiques de ce même cadre.

### 3.1.1. Transformations / Reformulations

L'AMD s'inspire des théories de Zellig Sabbetaï Harris et du développement qui en a été fait par Maurice Gross, concepteur du « Lexique-grammaire ». Les recherches qui émanent de ce courant s'appuient sur les transformations linguistiques qui, dans le respect des équivalences sémantiques, permettent de faire ressortir les structures de base des énoncés ainsi analysés, éventuellement par des procédures comparatives successives<sup>4</sup>.

Ainsi, par exemple, l'énoncé *Votre position est indéfendable* peut être reformulé par : *La position qui est la vôtre est indéfendable / La position qui est la vôtre ne peut pas être défendue / Il n'y a pas de défense possible de votre position / Votre position est impossible à défendre / On ne peut pas se défendre quand on est dans une position comme la vôtre*, etc.

Ce procédé a l'avantage de montrer le fonctionnement des éléments de la langue en rapport direct avec le sens qui leur est propre en fonction des positions occupées dans la phrase, du traitement morphosyntaxique qu'on leur réserve et de l'environnement lexical.

### 3.1.2. Prédication

Pour l'AMD, la notion de prédication est essentielle. Comme on le constate dans Ibrahim (éd.) (2009), elle peut recouvrir des conceptions assez différentes selon les cadres théoriques d'analyse. La question est complexe, et nous retiendrons principalement ici que prédiquer équivaut à dire quelque chose dont le sens est achevé et accessible, c'est-à-dire perçu comme complet : « une prédication se distingue par sa complétude » (Ibrahim 2015 : 344), c'est « une forme qui se parachève au regard d'un contenu » (*ibid.* : 343).

Ce qui nous semble particulièrement fructueux dans l'AMD est le fait qu'elle amène à se demander quels sont les éléments prédicatifs d'un énoncé, ou quelle est la force prédicative de ces éléments en cas de prédication distribuée (Ibrahim 2015 : 355). Le verbe, par exemple, n'occupe pas forcément une fonction centrale dans la phrase, et peut avoir juste une fonction de support, puisque les noms ou même des éléments d'autre nature peuvent être prédicatifs.

<sup>4</sup> Appelés « classes d'équivalence », ces cadres « qui permettent des changements de forme sans changer l'information [d'un énoncé de départ] » sont de 4 types : paraphrase, transformation, restructuration et équivalent métalinguistique (Ibrahim 2015 : 59). Dans notre travail, pour des exigences de simplification autant que de pertinence par rapport au corpus analysé, nous ne distinguons pas les typologies de classes d'équivalence et nous nous en tenons à la notion de *reformulation paraphrastique* (Ibrahim 2015 : 296 ; Martinot 1994, 2018).

### 3.1.3. Complexité et matrices analytiques définitives (MAD)

Héritage du courant harrissien, pour lequel « la complexité notoire de la grammaire, dont la majeure partie est créée par les réductions, n'est pas due à la complexité de l'information » (Harris 2007 : 45), les notions de simplicité et complexité linguistiques deviennent vite centrales pour l'AMD. Dans sa description des constructions d'une langue, Ibrahim distingue tout d'abord deux types de simplicité/complexité :

[...] la simplicité en langue n'est pas la même selon qu'il s'agit d'une simplicité d'usage qui a recours à une forme préconstruite, qui peut être le résultat d'une formation complexe, ou d'une simplicité intrinsèque, une forme à la fois indécomposable et facile d'accès. (Ibrahim 2013: 20)

Cette distinction présente un intérêt certain pour la compréhension de structures plus propres à la langue parlée spontanée et l'appréhension d'un système qui semble lui être spécifique : bien que cela puisse sembler paradoxal, la complexité grammaticale n'est pas dans la longueur et l'extension, mais bien dans la condensation des énoncés. Nous partageons l'explication qu'en donne Ibrahim :

En plus de [la] complexité explicite faite de choix [parmi un éventail de combinaisons possibles], il y a dans la langue une deuxième forme de complexité dont le locuteur n'est le plus souvent pas conscient. Il s'agit d'une complexité pour ainsi dire préconstruite, sous-jacente aux formes plus ou moins fortement grammaticalisées de la langue et qui a généralement tous les dehors de la simplicité. Du fait de la fréquence de son usage ou du fait que cet usage n'implique pas de choisir entre plusieurs formes possibles dans le même contexte, elle est perçue comme simple. Sa complexité réside dans le parcours qui l'a rendue disponible, dans sa généalogie. La description de la langue montre que l'accès à cette complexité sous-jacente, même quand il est dans les faits très rapide, reste opaque. Les formes qui relèvent de cette complexité ont une apparence élémentaire voire première et indécomposable. Elles sont en fait décomposables. **Ainsi par exemple l'expression: Assis ! a-t-elle une généalogie fort complexe de type: Je (te + vous) (demande + ordonne) de (t'asseoir + vous asseoir + rester assis)!<sup>5</sup>** (Ibrahim 2013: 23)

L'exemple reporté en gras montre comment un dépliage de la langue met au jour la complexité d'un énoncé – apparemment simple – qui est en fait réduit/condensé. C'est ce principe qui est au centre de l'Analyse matricielle définitive et dont nous nous inspirons pour notre recherche, en nous appuyant sur l'équivalence réduction / complexité.

<sup>5</sup> C'est nous qui soulignons.

Les matrices analytiques définitoires (MAD) sont des modèles d'explicitation qui déplient les structures grammaticales et décomposent le lexique en termes élémentaires. Elles donnent aussi des informations de type énonciatif qui découlent de l'énoncé. Les séquences ainsi obtenues permettent un accès au sens sans ambiguïté. Nous reportons l'exemple d'une matrice générée selon ce principe (Ibrahim 2001 : 110), à partir de l'énoncé radiophonique *dans le genre compil de ténor, ça vous changera des pavarottismes habituels*<sup>6</sup> :

[MAD : (une personne + quelqu'un) dit à un moment non précisé que dans l'ensemble (d'éléments assez nombreux + d'extraits assez célèbres) pour former un genre dénommé compilation de ténor et abrégé en compil de ténor, quelqu'un changera, passant de son état défavorable dû au fait que les imitations en série du modèle que Pavarotti donne en exemple à suivre dans le pavarottisme, font de son nom propre dans les pavarottismes un nom commun et un exemple à ne pas suivre, à un état favorable dû à une compil différente des compils des pavarottismes.] (Ibrahim 2001 : 110)

Le développement que nous ferons de nos matrices ne suit pas fidèlement le modèle proposé par Ibrahim (2001, 2015), mais l'adapte à nos objectifs de recherche, aux spécificités de nos corpus et aux différents niveaux de l'analyse, explicitant notamment l'aspect pragmatique. En particulier, nous ne les développons pas exhaustivement et le lexique choisi n'est pas toujours indécomposable. Aussi présentons-nous nos matrices sous l'appellation [MAD<sub>adaptée</sub>], qui se réfère directement au principe des MAD tout en signalant qu'elles sont élaborées différemment.

### **3.2. Hypothèse de recherche : la complexité linguistique comme marqueur de spontanéité ?**

Lors de repérages au sein de corpus dont les modalités d'enregistrement garantissent le caractère très spontané des énoncés (PCF<sup>7</sup> en particulier), ou en observant les interactions de notre entourage, nous relevons des constantes, comme le caractère réduit de certaines constructions prédicatives.

C'est le cas du *si* assertif : sa valeur prédicative répond implicitement à une phrase formulée négativement et peut, selon l'intonation, le contexte mais aussi les mimiques qui l'accompagnent, avoir des valeurs pragmatiques très différentes : ex. *Tu ne veux pas de café ? Si !* La réponse correspond aussi bien à [MAD<sub>adaptée1</sub> : bien sûr que je veux du café je me demande bien comment tu peux imaginer que je vais m'en passer !] qu'à [MAD<sub>adaptée2</sub> : j'en veux bien un si ça ne

<sup>6</sup> Été 2000 sur France Musique à propos de Fritz Wunderlich.

<sup>7</sup> Corpus Phonologie du français contemporain.

te dérange pas trop de le faire], ou encore à [MAD<sub>adaptée3</sub> : je réponds à l'affirmatif tout en exprimant mon hésitation quant au choix que je dois faire car je suis encore en train d'y réfléchir]. D'ailleurs, chaque valeur de *si* entrainera des répliques différentes et difficilement interchangeables, comme *Ne t'énerve pas je t'en fais un tout de suite* ou *tu bois pas trop de café ?*, pour le premier cas, *J'en bois aussi volontiers un*, pour le deuxième cas, ou *Tu as peur de ne pas dormir ?*, pour le troisième. *Si* est donc bien une forme condensée, qui, employée seule à l'oral et supportée par des intonations adéquates, peut avoir plusieurs valeurs sémantiques et pragmatiques, et est absolument compatible avec l'élan conversationnel et le caractère « implicite » des structures propres à l'oral spontané.

Autre forme remarquable par sa complexité, le pronom sujet *ça*, particulièrement présent à l'oral. Dans la prédication *c'est qu'ça bouffe un doberman*<sup>8</sup>, il permet de combiner différentes valeurs sémantiques [loi générale ou habitude x abondance ou intensité x conséquence(s) implicite(s) que l'on déduit à partir de notre connaissance du monde]. Autre exemple, la construction prédicative *ça craint*<sup>9</sup> se référant à un quartier mal fréquenté, peut être dépliée en [MAD<sub>adaptée</sub> : dans ce quartier les gens ont habituellement de bonnes raisons d'avoir peur pour leur sécurité], où la combinaison des mêmes valeurs sémantiques nous semble se confirmer.

Les formes réduites soutiennent donc effectivement une économie linguistique (segmentale) qui permet des interactions rapides et spontanées entre locuteurs et renforce expressivité et complicité. Toutefois pour la recherche de marqueurs dont on n'a pas l'intuition préalable, ce type de recensement montre très vite ses limites. Sur la base de ces observations, nous émettons l'hypothèse qu'un locuteur qui s'exprime spontanément au cours d'une interaction animée a tendance à produire des énoncés syntaxiquement plus complexes que s'il surveille son énonciation. Cependant, pour vérifier une telle hypothèse, il nous faut établir si une variation de la spontanéité fait effectivement varier la complexité linguistique des énoncés produits, chez un même locuteur.

#### 4. Constitution d'un corpus *ad hoc* « spontanéité »

C'est dans le contexte de l'expression réitérée d'une même intention sémantique et dans le cadre minimal d'une prédication que la variation intrapersonnelle nous semble le plus fructueusement observable. Cependant, un tel corpus de variantes nous a d'abord semblé difficile à recueillir, même en envisageant d'enregistrer un

<sup>8</sup> Tiré d'un sketch de Coluche, intitulé *L'autostoppeur*.

<sup>9</sup> Tiré du corpus PFC : *on n'est pas allés dans les quartiers où ça craignait* (72715 21abl1 Dijon).

locuteur dans des situations très différentes – ou du moins pas assez riche pour y asseoir notre analyse et envisager un recensement systématique des marqueurs présents dans le corpus.

#### 4.1. Hypothèse expérimentale et résultats exploratoires

Remarquant toutefois que dans les conversations courantes l'autoreformulation est récurrente, qu'elle soit spontanée ou qu'elle soit sollicitée pour différentes raisons (bruit ambiant, audition défaillante ou distraction de l'interlocuteur, message ambigu voire incompréhensible en réception, ou prononcé trop rapidement, etc.), nous émettons une seconde hypothèse, de type expérimental cette fois : si l'on signifie à un locuteur qui s'exprime spontanément – au cours d'une conversation informelle où il est à l'aise – que l'on n'a pas compris ce qu'il a dit, celui-ci fera un effort pour rendre accessible ce qu'il veut dire, en prêtant une attention accrue à la forme de son discours, soit en répétant l'énoncé pour en améliorer la forme phonique, soit en le reformulant tout en lui conservant son sens premier.

Une telle séquence de reformulation offrirait l'opportunité d'étudier le passage d'une énonciation spontanée à une énonciation plus vigilante, au niveau des équivalences sémantiques prédicatives. La sollicitation aurait alors aussi bien une fonction de réducteur de la spontanéité qu'une fonction d'indicateur de la variation, permettant de bien circonscrire le phénomène à étudier : l'énoncé à l'amont de la sollicitation peut être considéré comme un échantillon de spontanéité maximale ou en tout cas supérieure à l'énoncé de même sens en son aval. Les deux énoncés peuvent être en outre aisément comparés.

Ci-après, un exemple tiré d'enregistrements exploratoires, qui illustre le processus envisagé :

- (1) La conversation est enregistrée dans un contexte familial, où LOC2 a préparé un gâteau

LOC1 t'as pris du gâteau ? (a)  
 LOC2 ... [SOLLICITATION MUETTE] (b)  
 LOC1 tu as pris du gâteau ? (c)  
 LOC2 hmm ? [2e SOLLICITATION] (d)  
 LOC1 tu as pris du gâteau ? je te demande si t'as goûté mon gâteau si tu as pris un morceau de gâteau (e) (corpus *Ad hoc*-SP)

L'interaction (1) comporte deux sollicitations, la première (b) est silencieuse, la seconde (d) correspond à une expression non voisée à intonation montante : « hmm ? », mais toutes deux fonctionnent effectivement comme des invitations à redire. Si nous analysons maintenant les énoncés produits pendant cette interaction, nous constatons que la première formulation (a) comporte bien, par rapport

aux reformulations qui suivent (c et e), des marques de réduction ou condensation, qui vont même au-delà de la réduction syntaxique attendue :

- une élision du pronom qui brouille la perception des frontières de mots : *t'as pris du gâteau* (a) vs la forme non élidée *tu as pris* (c) qui les reconstitue ;

- une interrogation directe avec une structure argumentale de type  $[N_0 V N_1]$  : *tu as pris du gâteau ?* vs une interrogation indirecte introduite par le verbe performatif demander (e) : *je te demande si...* ;

- un nom déterminé par l'article partitif : *du gâteau* (a) vs un déterminant nominal : *un morceau de gâteau* (e), et vs un déterminant possessif : *mon gâteau* (e) correspondant dans ce contexte d'énonciation au [gâteau que j'ai fait + que j'ai apporté] ;

- un terme sémantiquement plus générique *pris* (a) par rapport à une unité lexicale plus précise et restrictive : *goûté* (e).

On peut affirmer à partir de cet exemple (1) que les éléments, voire les phénomènes mis au jour en amont de la requête de répétition/reformulation sont bien les marqueurs d'une spontanéité maximale. L'élan est au contraire freiné par la sollicitation et la reformulation est alors le produit d'une énonciation plus vigilante. Ce premier constat nous a encouragée à continuer l'expérience, de manière à recueillir un corpus spécifique, constitué d'autoreformulations sollicitées par un enquêteur qui feindrait une mauvaise compréhension.

Nous avons donc continué à enregistrer des interactions selon ces modalités et elles ont confirmé la validité de ce système pour obtenir des séquences permettant d'observer la variation intrapersonnelle et la scalarité de la spontanéité énonciative. Les reformulations de la part du locuteur soucieux de mieux se faire comprendre vont bien dans le sens d'une explicitation des liens grammaticaux ou d'un « dépliage définitoire » au niveau lexical, mais aussi méta-énonciatif et phonétique.

L'extrait suivant (2) montre bien aussi selon quelles modalités s'effectue ce dépliage :

(2) Mail d'anniversaire

F1 non ben Lauredane\* j'ai Lauredane\* j'ai reçu un p(e)tit mail là pour mon anniversaire

vs

F2 oui oui là je hm parce que Lauredane ne m'a pas rappelée là aujourd'hui j'ai juste eu un mail ↗ euh dans l(e)quel elle me souhaitait un bon anniversaire ↗  
(corpus *Ad hoc*-SP)

Remarquons tout de suite qu'ici les formulations F1 et F2 ont le même sens, mais F1 comprend 14 mots alors que F2 en compte 24 outre une fragmentation de l'information avec deux requêtes de

feedback ↗, le tout témoignant d'une expansion de la forme en F2 *vs* d'une réduction en F1.

Dans la formulation très spontanée F1, *petit (j'ai reçu un petit mail)* est employé comme un minimiseur ayant le sens de *seulement*, ce que nous confirme l'emploi de *juste* dans la reformulation F2 (*j'ai juste eu un mail*). La formulation avec *petit* est ambiguë (un mail court ?) alors que la reformulation lève cette ambiguïté.

La préposition *pour* contribue à la réduction de la prédication seconde en F1 (*pour mon anniversaire*) alors qu'en F2 nous constatons le dépliage d'une proposition relative, avec un pronom relatif composé (*euh un mail dans lequel elle me souhaitait un bon anniversaire*). Le lexème *souhaiter* est employé en F2 mais pas en F1 où il est implicite et interprétable selon notre connaissance du monde, nos connaissances partagées ou notre vécu. Enfin, le filler *euh* (F2) indique une réflexion avant la production d'une construction aux liens analytiques explicités qui requiert plus de mobilisation à un niveau cognitif et plus de concentration.

#### 4.2. Protocole de recueil du corpus et considérations

Les résultats obtenus par le stratagème de reformulation sollicitée se sont révélés inespérés, dans la mesure où près de 95% des sollicitations ont effectivement donné lieu à des reformulations, avec un taux très bas d'énoncés répétés à l'identique. Pour la constitution de ce corpus spécifique, que nous désignons comme « corpus *Ad hoc*-SP»<sup>10</sup>, nous avons enregistré de novembre 2016 à janvier 2018 des interactions se déroulant dans notre quotidien. Qu'il s'agisse de simples prises de nouvelles entre proches ou de séances de travail, ces dialogues entrent dans les interactions habituelles du locuteur LOC1, soit nous-même. LOC1 fait office d'enquêteur et il recueille et transcrit le corpus.

Le corpus *Ad hoc*-SP se compose de conversations entre deux locuteurs (LOC1 et un de ses interlocuteurs) qui se connaissent bien et qui, pour la plupart, communiquent de vive voix – en présence ou téléphoniquement – depuis plusieurs années. Les tours de parole sont souvent denses, et les sujets abordés extrêmement variés : achats d'un camping-car, de semelles ou de fruits de mer, travaux de rénovation, choix d'une montre pour un cadeau commun, discussion sur la réforme de l'orthographe, commentaires sur le temps qu'il fait, inquiétudes face à la crise, nouvelles des enfants, gestion du silence dans les bibliothèques, etc.

En tout neuf locuteurs (enquêteur LOC1 compris) sont impliqués dans cette expérience, avec un nombre d'interventions très variable pour chacun. Toutes les interactions avec autoreformulation sollicitée sont des dialogues qui impliquent l'enquêteur LOC1 et un de ses huit

<sup>10</sup> SP pour « spontanéité ».

interlocuteurs. L'appellation LOC1 désigne toujours l'enquêteur, alors que les huit autres locuteurs sont désignés indistinctement par LOC2, ce qui contribue aussi à préserver leur identité<sup>11</sup>.

Au moment de l'interaction, les locuteurs ne savent pas qu'ils sont enregistrés<sup>12</sup>, ce qui constitue une entorse aux « bonnes pratiques » (Baude *et al.* 2006). C'est cependant à cette condition que nous avons pu mettre en œuvre un protocole adéquat à l'étude des marqueurs de spontanéité. Notre corpus a été anonymisé lors de sa transcription de manière à empêcher l'identification des personnes enregistrées.

Pour ce qui est des facteurs sociaux qui peuvent influencer le style des locuteurs, signalons que tous les locuteurs du corpus *Ad hoc*-SP sont âgés de plus de 50 ans. Si une bonne moitié d'entre eux sont titulaires d'un diplôme universitaire, le plus souvent un doctorat, les durées de scolarisation sont dans l'ensemble extrêmement variées. Pour préserver leur anonymat, nous ne décrivons pas davantage chacun d'eux.

Les interactions ont été conduites via Skype audio ou, plus rarement, par téléphone portable (smartphone). Elles ont été enregistrées, avec une qualité satisfaisante, au moyen des logiciels *MP3 Skype recorder* ou *Call recorder*, téléchargés en ligne dans leur version libre. Le corpus *Ad hoc*-SP regroupe 42 conversations, totalisant plus de 70 autoreformulations sollicitées, qui, en vertu des découpages opérés, correspondent à une durée audio d'environ 2 heures, pour une retranscription de 15000 mots environ. Au départ, le corpus recueilli était beaucoup plus important (presque 30 heures de conversation).

Le protocole appliqué pour la constitution de ce corpus prévoit, comme nous l'avons illustré, une ou plusieurs sollicitations adressées au locuteur-informateur. Ces sollicitations se font en général au moyen des énoncés suivants, éventuellement combinés entre eux : « pardon ? » ; « j'ai pas compris » ; « j'ai rien compris » ; « j'ai pas entendu » ; « quoi ? » ; « attends, qu'est-ce que tu as dit ? », etc.

Si l'intention initiale était de solliciter notre interlocuteur à deux reprises, voire trois, à partir d'un même énoncé, dans la réalisation effective des enquêtes les sollicitations ont été plutôt isolées. Cette pratique réitérée aurait pu en effet paraître suspecte aux locuteurs-témoins ou les déstabiliser au point de compromettre l'interaction. Or les interactions avec ces locuteurs se déroulent généralement hors des lois du « marché linguistique », au sein d'une énonciation spontanée où l'on se soucie davantage de ce que l'on dit que de *comment* on le dit. Le stratagème de sollicitation a pour but de faire varier la spontanéité en

<sup>11</sup> Les différents énoncés d'un même locuteur sont davantage caractérisants s'ils peuvent être regroupés.

<sup>12</sup> L'autorisation de ces locuteurs anonymes a été demandée par la suite, avant la diffusion de la thèse en vue de sa soutenance, en leur soumettant les transcriptions qui les concernaient. Nous saisissons l'occasion pour les remercier et exprimer notre gratitude à propos de leurs réactions bienveillantes malgré la surprise ou parfois le mécontentement initial qu'on peut bien imaginer.

ramenant provisoirement l'attention de l'interlocuteur sur la forme de ce qu'il énonce. L'analyse des reformulations obtenues, caractérisées par une langue plus analytique, nous confirme que ce procédé donne les résultats attendus.

La question de savoir quand doit intervenir la sollicitation a continué de se poser tout au long de l'enquête. Le moment de la sollicitation s'est fait un peu au hasard au début, puis davantage par intuition ou à la suite d'une prédication que l'enquêteur perçoit comme complexe ou ambiguë. Parfois la sollicitation se fait par réaction à une voix qui faiblit, à un débit qui s'accélère, ou si effectivement la compréhension n'est pas optimale, pour des raisons techniques ou acoustiques. Il n'y a donc pas un vrai choix de la part de l'enquêteur quant au moment où solliciter. Cela se justifie aussi par le fait que celui-ci est pleinement impliqué dans la conversation et que les tours de parole sont denses et rapides. C'est donc seulement dans un second temps, après avoir retranscrit puis analysé l'interaction, que l'intérêt des séquences devient manifeste, et non pendant leur enregistrement. Nous nous sommes efforcée cependant, en tant qu'enquêteur, après avoir retranscrit plusieurs sollicitations qui n'étaient que partielles<sup>13</sup>, de n'intervenir qu'au terme d'une prédication complète.

## 5. Résultats

Au fur et à mesure des enregistrements, nous avons transcrit des séquences d'une dizaine de tours de parole environ autour des sollicitations. Les longueurs dépendent toutefois de la nécessité d'obtenir des unités d'analyse cohérentes et compréhensibles. Un lecteur potentiel doit en effet pouvoir resituer les énoncés reformulés et leur version première dans leur contexte, pour les comprendre. À cet effet, un titre et un bref résumé de la situation chapeautent les extraits de notre corpus, facilitant l'entrée en matière. De même, pour ne pas amputer le sens, les unités de découpage pour lesquelles nous avons opté respectent autant que possible la cohésion discursive : une reprise par pronominalisation, par exemple, peut être séparée de son antécédent nominal par un long énoncé constitué de plusieurs tours de parole que nous avons donc gardé. Après recueil et transcription partielle des corpus, nous avons isolé et analysé les énoncés produits juste en amont et en aval de la sollicitation, soit respectivement les formulations premières (F1), plus spontanées, et les autoreformulations (F2). Un titre a été attribué à chaque interaction, auquel s'ajoute une récapitulation – plus ou moins ample – du contexte, comme en témoigne cet extrait du corpus (3) :

<sup>13</sup> Soit un type de sollicitation portant sur une partie seulement de la prédication qui vient d'être énoncée, par exemple : « *tu as fait **quoi** ?* ».

- (3) Excursions scolaires (LOC1 fait des considérations sur les frais que doivent supporter les parents lors de voyages scolaires impliquant leurs enfants : « mais c'est vrai que j'veux dire pas tout l'monde peut sortir comme ça 190 euros plus euh plus c'qui va à côté pour euh xxx pour une balade scolaire » ; dans l'extrait qui suit, LOC2 approuve en évaluant à 300 euros la somme dont il est finalement question au vu des frais qui s'ajouteront inmanquablement durant le voyage)<sup>14</sup> :

F1 LOC2 ah ouais c'est au minimum ça [*chevauchement*] (F1a)

oui j'dis « ça fait minimum 300 euros » quoi (F1b)

c'est pas c'est pas xxx (*chevauchement*)

LOC1 xxx pardon ? j'ai pas compris [SOLLICITATION]

F2 LOC2 j'dis ça fait minimum 300 euros ↗ (*intonation légèrement montante pour demander un feed-back de compréhension*) c'est pas évident ! (corpus *Ad hoc*-SP)

Types de marqueurs: → modalisation sans/avec verbe introducteur de discours + lexème peu spécifique *vs* précision lexicale + marqueur *quoi* *vs* autre procédé de modalisation

Dans l'exemple (3), nous pouvons distinguer trois formulations différentes de sens équivalent. La première formulation (F1a), *ah ouais c'est au minimum ça*, énoncée dans l'élan de la conversation, est très spontanée ; la deuxième (F1b), *oui j'dis ça fait minimum 300 euros quoi*, est reformulée par LOC2 de sa propre initiative, dans le flux du même tour de parole ; la troisième (F2), *j'dis ça fait minimum 300 euros ↗ c'est pas évident !*, est sollicitée : elle est produite après que l'enquêteur a signifié qu'il n'avait pas compris, obligeant LOC2 à accorder une attention accrue à la forme de son énoncé. Dans les différents extraits que nous présenterons, nous confronterons toujours F1 et F2, mais il n'y aura pas forcément, comme ici, deux formulations F1.

En analysant cette interaction (3), nous relevons que l'énoncé (F1a) *c'est au minimum ça* est aussitôt reformulé en (F1b) par *oui j'dis ça fait minimum 300 euros quoi*. Cette deuxième prédication est introduite par un verbe de discours ; elle est également ponctuée par le marqueur *quoi* dont la fonction ici est à la fois récapitulative et phatique ; *quoi* n'est en revanche pas maintenu en F2. On remarque que *c'est* en (F1a), très polyvalent, est remplacé par une construction qui exprime un peu plus précisément l'idée de coût (*ça fait*) sans toutefois qu'il y ait recours à ce niveau à un lexique spécifique ; de même le pronom *ça* (F1a) est explicité au moyen d'une quantification (*300 euros*) en (F1b).

<sup>14</sup> Les conventions adoptées respectent l'orthographe conventionnelle. Cependant afin de rendre compte d'indices prosodiques importants pour la confrontation entre la formulation spontanée et sa reformulation, des ajustements sont opérés : en particulier, les schwas ou autres phonèmes non produits sont signalés par une apostrophe en début de mot (« 'tends ») et en fin de mot unisyllabique (t'as lu) et mis entre parenthèses à l'intérieur d'un mot (p(e)tit) ; pour les autres mots, les schwas – normalement non prononcés par des locuteurs non méridionaux, en particulier à la fin des mots – ne sont pas signalés.

En F2 on relève à la fin *c'est pas évident*, qui semble avoir été amorcé en F1 (*c'est pas c'est pas*) ; cet énoncé explicite en F2 la prédication qui précède et en fixe le sens, en conformité avec ce qu'avait commencé à dire LOC1 en introduisant le sujet abordé, c'est-à-dire que les frais des excursions scolaires peuvent représenter un problème pour un certain nombre de familles.

En F2 les deux prédications sont séparées par une demande de feedback (↗) exprimée par une intonation légèrement montante et requérant une confirmation de compréhension ou d'adhésion aux propos : *j'dis ça fait minimum 300 euros ↗ c'est pas évident !* Au vu des définitions et classifications proposées par Martinot (1994 ; 2018), nous considérons que la deuxième prédication *c'est pas évident* est une reformulation sémantique de la première, qui ne reprend que le sens de F1 et aucun autre élément segmental. On pourrait objecter que l'énoncé *c'est pas évident* ne reformule pas l'énoncé précédent, mais fournit une évaluation du contenu référentiel qui précède. Cependant la prédication porte bien dans les deux cas sur le poids que représentent ces dépenses : déjà en F1, au moyen de l'intonation notamment et de l'emploi de *minimum*, dans un énoncé identique, LOC2 soulignait l'excès des contraintes financières imposées aux parents d'élèves. L'information *300 euros* pourrait être vue, parce que c'est une donnée chiffrée, comme une donnée objective et être considérée comme l'information principale ; or, ce qui compte ici au niveau prédicatif, c'est que le locuteur dénonce une dépense excessive. Il pourrait aussi bien dire, sur le même ton, *ça fait minimum 20 euros* que *ça fait minimum 500 euros* et dire encore la même chose, soit *cela revient trop cher*, indépendamment du montant annoncé qui ne constitue pas l'information principale.

Il est ainsi utile pour analyser les énoncés de notre corpus de nous interroger sur ce qui est prédiqué, c'est-à-dire ce que dit d'essentiel le locuteur. C'est pourquoi la notion de prédication s'avère être fondamentale pour déterminer la visée du message. De même, un dépliage de type matriciel (Ibrahim 2013, 2015) que nous adaptons à nos objectifs de recherche, à la typologie de notre corpus, et aux différents niveaux de l'analyse nous permet de mieux comprendre le fonctionnement de l'énoncé, voire de l'interaction.

Ainsi, à propos de l'exemple suivant (4), nous montrons comment le développement de matrices explicite des éléments réduits qui semblent incohérents une fois transcrits (ici, *parce que* et *non*) :

(4) Problèmes techniques Skype

LOC2 allo ?

LOC1 oui oui oui j'entends

LOC2 tu m'vois ? tu m'entends ? parce que moi j'ai plus rien là (a)

LOC1 non non j't'entends (b)

(corpus *Ad hoc*-SP : un extrait d'oral spontané, sans sollicitation)

[MAD<sub>adaptée</sub> (a) : je te demande si tu me vois sur ton écran et si tu m'entends à présent **et je te pose la question parce que** de mon côté je n'ai plus rien, ni son ni image]

Alors que *parce que* est peu cohérent au premier abord dans la transcription<sup>15</sup>, car il n'y a pas de lien de causalité direct entre les questions *tu m'vois ? tu m'entends ?* et l'explication qui suit, le dépliage permet de comprendre que le lien de causalité concerne en fait l'aspect énonciatif.

[MAD<sub>adaptée</sub> (b): contrairement à ce que tu présumes à cause du fait que toi tu ne vois rien et n'entends pas, je n'ai pas les mêmes conditions techniques que toi et je t'entends]

*Non* semble incohérent par rapport à l'assertion *j't'entends* mais il ne l'est pas si on se réfère à la matrice.

Ainsi, les MAD<sub>adaptée</sub> aident à appréhender, par reconstruction, des éléments du français parlé qui ont l'air incohérents ou insignifiants mais dont la valeur, aussi bien prédicative que pragmatique, est essentielle.

Ce sont le plus souvent des phénomènes variés de réduction et implicitation que nous relevons dans les reformulations sollicitées, mais il y en a d'autres, qui touchent différents niveaux de la langue. Nous avons initialement répertorié les marqueurs sous sept grandes catégories : *traitement de la forme sonore et de l'information, traitement syntaxique, traitement lexical, valeurs temporelles et aspectuelles, modalisation*. Cependant, l'analyse des extraits met en évidence une concomitance de marqueurs, qui nous semble être un facteur important des synergies dont se prévalent les énoncés produits spontanément.

Nous l'illustrons avec l'exemple suivant (5) :

- (5) Retard de train (LOC1 appelle chez elle pour indiquer qu'elle arrivera plus tard que prévu, suite à un retard de train)

LOC2 bon écoute, fais-moi savoir quelque chose

LOC2 de quand tu pars à quand t'arrives c'est quoi ? une heure ? (F1)

LOC1 non attends j'ai pas compris attends excuse-moi j'ai pas compris [SOLLICITATION]

LOC2 quand le train part ^ (F2-) (*intonation montante pour requérir un premier feed-back de compréhension*)

LOC1 oui

LOC2 combien de temps il met pour arriver à Pordenone ? (F2)

LOC1 euh une heure exacte de Gorizia pour arriver à Pordenone

LOC2 bon alors je verrai, tu m'fais savoir quand tu pars

<sup>15</sup> Nous précisons *dans la transcription*, car dans le fil de la conversation et supportée par l'intonation, la prédication est bien reçue.

Types de marqueurs : → implicitation de l'actualisateur du nom + mot polyfonctionnel (modalisateur /connecteur logique/ phatique) + omission du *ne* de négation

En (5), la formulation spontanée, en amont, implicite l'information partagée, à savoir le fait que le moyen de transport que prend LOC1 est le train, information formalisée en F2 : un changement de sujet pour les verbes *partir* et *arriver* se vérifie en reformulation, soit **tu pars** en F1 vs **le train part** en F2 et *quand t'arrives* (F1) vs *combien de temps il [le train] met pour arriver à Pordenone ?* (F2), où en F2 la ville de destination est nommée. Il en résulte en aval une formulation plus claire et explicite, qui tient aussi à l'emploi d'un mot interrogatif spécifiquement temporel : *c'est quoi ? une heure ?* (F1) vs **combien de temps il met pour arriver ?** (F2). Soit en F1 un mot interrogatif générique (*quoi*) qui n'acquiert son acception temporelle que par sa proximité avec le segment suivant qui en complète le sens (*une heure ?*).

- (6) Respect d'une échéance (Il est ici question de délais et prorogations à accorder à l'auteur d'un article à éditer)

F1 LOC1 en fait c'est plus facile **pour les échéances** avec les gens que tu connais pas du tout

LOC2 comment ? [SOLLICITATION]

F2 LOC1 c'est plus facile je pense euh /**d'faire respecter les échéances** avec des gens que tu n'connais pas du tout (corpus *Ad hoc*-SP)

Types de marqueurs : → nom sans actualisateur *vs* avec actualisateur + modalisation sans verbe introducteur du discours *vs* avec verbe introducteur du discours + négation avec *pas* et omission de *ne* *vs* réalisation de *ne...pas*

Dans l'exemple (6), alors qu'en F1 le nom *échéances* n'est pas actualisé, on constate l'introduction en F2 de l'actualisateur (ici un verbe support, *respecter*) qui précise le sens visé (**respecter les échéances**) et le complète par un verbe opérateur : *faire*. Notons ici que la sollicitation porte sur un énoncé produit par l'enquêteur LOC1, elle n'est donc pas feinte.

La formulation première est effectivement ambiguë du fait de son incomplétude : même si le contexte aurait pu permettre de reconstruire l'actualisation manquante, *échéances* n'est pas à lui seul prédicatif, ses actualisateurs potentiels ne sont pas tous de simples verbes supports, certains (opérateurs) peuvent modifier radicalement le sens du prédicat. On peut fixer une échéance, la respecter, la

proroger, etc. La force prédicative est alors distribuée. Le raccourci était donc risqué car la construction est incomplète.

En (6), une modalisation est introduite dans F2, sous la forme d'un verbe d'opinion en incise : *c'est plus facile je pense euh / d'faire respecter les échéances avec les gens que tu connais pas du tout*. En F1, l'absence d'un verbe introducteur du discours nous fait penser, au premier abord, que la modalisation n'est pas présente dans la formulation la plus spontanée. Or, *en fait* (F1), signifiant ici 'tout compte fait', 'en fin de compte', 'à y bien penser', 'après réflexion', est bien aussi un modalisateur qui va dans le même sens que *je pense* qui apparaît en F2. Mais il a aussi une fonction de phatique en début de phrase. Nous observons donc une différence dans la construction des deux énoncés (F1 vs F2), mais la posture énonciative est la même dans les deux formulations. Il est très vraisemblable qu'une deuxième sollicitation aurait pu faire produire à LOC1 une construction plus analytique (plus développée grammaticalement), s'appuyant sur un verbe introducteur suivi d'une conjonctive, de type : [*je pense/je dis que c'est plus facile de faire respecter les échéances avec les gens que tu ne connais pas du tout*].

Notons également que dans la formulation F2, le phénomène décrit ci-dessus s'accompagne d'une intégration morphosyntaxique, avec l'emploi d'un *ne* de négation, qui garantit plus de clarté sémantique grâce à sa redondance par rapport à *pas du tout* (*avec les gens que tu connais pas du tout vs avec des gens que tu n'connais pas du tout*). Nous remarquons enfin, sans toutefois proposer d'analyse pour le moment, le passage d'un déterminant défini (*les gens*, en F1) à un déterminant indéfini (*des gens*, en F2).

## 6. Conclusions et perspectives

*Ad hoc*-SP rassemble des équivalences sémantiques produites à deux, voire trois, degrés différents de spontanéité. Grâce à la constitution de ce corpus, des marqueurs de spontanéité ont pu être relevés de manière systématique et différents aspects de la réduction se sont précisés, confirmant notre hypothèse de recherche. Les réductions sont parfois d'une telle ampleur dans les énoncés produits très spontanément (F1) qu'on peut s'étonner qu'un locuteur puisse quand même se faire comprendre<sup>16</sup>. Remarquablement, ces choix du

<sup>16</sup> Un de nos relecteurs anonymes – que nous remercions pour ses indications – suggère de prendre en compte les principes de Grice sur l'information et la gestion de l'implicite (théorie des implicatures conversationnelles) : « Étant donné le contexte partagé qu'un locuteur a avec son interlocuteur, le premier apporte sa contribution à la conversation en observant les maximes conversationnelles. Selon la maxime de la quantité, le locuteur fournira la quantité nécessaire d'informations. Les séquences réduites peuvent être utilisées simplement parce que le contexte partagé entre les deux interlocuteurs est tel qu'aucune explication approfondie n'est nécessaire. Ainsi, une

locuteur se révèlent pourtant être une prise de risque calculée qui, loin d'entraver la communication, au contraire, la dynamise. Les formes très réduites contribuent à l'élan du locuteur, quitte à ce que, par un auto-ajustement, celui-ci explicite certains énoncés immédiatement après.

La réduction touche des aspects de la langue que nous n'avons pas pris en compte en travaillant sur d'autres corpus. Les sollicitations ont permis de mettre en évidence – au sein du corpus retranscrit – des réductions que nous n'avons pu identifier comme telles que grâce à leur confrontation avec une forme plus explicite en aval, plus complète, plus précise. Nous retrouvons notamment dans les reformulations (F2) des phénomènes très variés de dépliage grammatical et les marques d'une attention accrue à la précision lexicale. Cela indique, inversement, que c'est bien la réduction grammaticale et un lexique plus général, ou essentiel (et de ce fait parfois ambigu) qui caractérisent la formulation la plus spontanée des énoncés (F1). La variation se manifeste par des effacements/ellipses en F1 et les restitutions correspondantes en F2.

Il nous semble important de souligner que, si les énoncés F2 explicitent et précisent ce qui avait été dit auparavant en F1, ils n'ajoutent rien ou pratiquement rien à l'aspect pragmatico-sémantique par rapport aux formulations plus spontanées : tout avait déjà été dit dans la forme condensée, dont le potentiel d'expression est de ce point de vue remarquable.

L'intonation joue un rôle primordial pour la complétude des énoncés, à plus forte raison pour ce qui est des formulations très

---

séquence réduite n'indique pas nécessairement un degré plus élevé de spontanéité, mais plutôt le fait que le locuteur apporte sa contribution conversationnelle en se basant sur certaines hypothèses quant à ce que son interlocuteur sait et, en obéissant à la maxime conversationnelle de la quantité, il produit simplement un énoncé qu'il juge suffisamment informatif. Lorsque l'interlocuteur demande une clarification, le locuteur se rend compte que son énoncé était sous-informatif. Dans ce cas, il ne devient pas moins spontané, mais il peut simplement reconsidérer la quantité correcte d'informations que son énoncé devrait contenir. Ce processus peut être considéré comme un ajustement aux besoins informationnels de l'interlocuteur » (notre traduction de l'anglais, s'appuyant sur *Deepl.com*).

Nous concordons tout à fait sur le fonctionnement indiqué par Grice (1975) de l'ajustement informationnel face à un locuteur en fonction des informations partagées, le corpus que nous avons recueilli l'illustre d'ailleurs parfaitement. En revanche, cela ne permet aucunement d'exclure qu'il y a une variation de la spontanéité à la base de ces ajustements. La spontanéité maximale, selon les définitions que nous en avons déjà données, correspond à un niveau zéro ou en tout cas un niveau très bas d'autocontrôle sur l'énonciation en cours et sur les énoncés déjà produits. En outre, nos locuteurs-informateurs (LOC2) avaient déjà produit des prédications complètes et parfaitement compréhensibles par LOC1 en F1; ils les reformulent sur une demande spécifique qui est motivée uniquement par le protocole de recherche. La spontanéité est moindre lors de l'énonciation de F2 parce que le locuteur se place dans une posture de contrôle métalinguistique par rapport à ce qu'il énonce, dans un souci de mieux se faire comprendre. Son énoncé F2 n'est par ailleurs guère plus riche en informations, même si certaines sont explicitées. Il comporte surtout des dépliages syntaxiques et des précisions à un niveau lexical.

spontanées, plus réduites, où les liens logiques et les différentes traces de la modalisation ne sont pas formellement exprimés. En effet, une intonation spécifique permet d'impliciter une partie des informations et des aspects pragmatiques.

Les différents phénomènes de réduction agissent en synergie : nous avons relevé une importante concomitance de marqueurs. Cette concomitance fait qu'en reformulation (F2) plusieurs aspects de l'énoncé sont modifiés par rapport à F1. Les équivalences sémantiques peuvent donc apparaître, d'un point de vue formel, très différentes entre elles. Le procédé mis en œuvre pour recueillir le corpus *Ad hoc-SP* donne la possibilité d'observer de manière privilégiée la variation intrapersonnelle et de constater que les phénomènes en jeu dans la production des variantes sont nombreux et complexes. C'est cela qui nous semble apporter de nouveaux résultats.

L'analyse d'*Ad hoc-SP* a déjà mené au recensement d'une cinquantaine de types de marqueurs qui demande à être complété au fur et à mesure d'une relecture attentive des extraits. Ces données, en partie reprises dans cet article, se distribuent en sept grandes catégories (forme sonore, information, syntaxe, lexicale, liens logiques, valeurs temporelles et aspectuelles, modalisation) toutes concernées par des phénomènes de réduction/ condensation.

Pour avancer significativement dans cette recherche sur la spontanéité énonciative, il nous paraît nécessaire de gérer les enregistrements et les transcriptions avec un logiciel spécifique, tel que CLAN ou ELAN, ce que nous n'avons pas pu faire jusqu'à présent. Par ailleurs, la concomitance des marqueurs doit faire l'objet d'un traitement spécifique et rigoureux, au moyen d'un outil informatique adapté, comme un tableur, pour mettre au jour la fréquence de chaque phénomène et leur association. Enfin, il serait fructueux d'élargir le corpus selon le protocole expérimenté, notamment avec des locuteurs plus jeunes.

## Références bibliographiques

- Baude, O., Blanche-Benveniste, C., Calas, M.-F., Cappeau, P. & Cordereix, P. (2006), *Corpus oraux, guide des bonnes pratiques*, CNRS Éditions, Presses Universitaires Orléans ; <https://hal.archives-ouvertes.fr/hal-00357706>.
- Bazillon, T. (2011), *Transcription et traitement manuel de la parole spontanée pour sa reconnaissance automatique*, Université du Maine ; <https://hal.archives-ouvertes.fr/tel-00598427/>
- Blanche-Benveniste, C. (1997), « La notion de variation syntaxique dans la langue parlée », *Langue française*, 115 (*La variation en syntaxe*), p. 19-29.
- Bouraoui, J.-L. M. (2008), *Analyse, modélisation, et détection automatique des disfluences dans le dialogue oral spontané contraint : le cas du Contrôle Aérien*, Université Paul Sabatier-Toulouse III ; <https://hal.archives-ouvertes.fr/tel-00354772/>.

- Bourdieu, P. (1982), « Le marché linguistique », in Bourdieu, P. (éd), *Questions de sociologie*, Les éditions de Minuit, Paris, p. 121-137.
- Corpus *Phonologie du français contemporain* : <http://www.projet-pfc.n>.
- Depoux, P. (2017), *Les redondances prédicatives en français parlé*, L'Harmattan, Paris.
- Depoux, P. & Stabarin, I. (2018), *La variation intrapersonnelle en français parlé : approches et statuts*, CRL, Paris.
- Di Cristo, A. (2016), *Les musiques du français parlé : essais sur l'accentuation, la métrique, le rythme, le phrasé prosodique et l'intonation du français contemporain*, De Gruyter, Berlin/Boston.
- Durand, J., Laks, B. & Lyche, C. (2009), « Le projet PFC: une source de données primaires structurées », in Durand, J., Laks, B. & Lyche, C. (éds), *Phonologie, variation et accents du français*, Hermès, Paris, p. 19-61
- Dutrey, C. (2014), *Analyse et détection automatique de disfluences dans la parole spontanée conversationnelle. Informatique et langage [cs.CL]*, Thèse de doctorat, Université Paris Sud.
- Grice, H. P. (1975), "Logic and conversation", in Cole, P. & Morgan, J. L. (eds), *Syntax and Semantics*, volume 3: *Speech Acts*, Academic Press, New York, p. 41-58.
- Harris, Z. S. (2007), *La langue et l'information* (trad. en français de *Language and Information* (1988) par Amr H. Ibrahim et Claire Martinot), CRL, Paris.
- Ibrahim, A. H. (2001), « Argumentation interne et enchaînements dans les matrices définitives », *Langages*, 142, p. 92-126.
- Ibrahim, A. H. (éd.) (2009), *Prédicats, prédication et structures prédicatives*, CRL, Paris.
- Ibrahim, A. H. (2013), « Une mesure unifiée de la complexité linguistique : l'analyse matricielle définitive », *Nouvelles perspectives en sciences sociales : revue internationale de systémique complexe et d'études relationnelles*, 9/1, p. 17-80.
- Ibrahim, A. H. (2015), *L'analyse matricielle définitive : un modèle pour la description et la comparaison des langues*, CRL, Paris.
- Luzzati, D. (2007), « Le dialogue oral spontané : quels objets pour quels corpora ? », *Revue d'Interaction Homme-Machine*, 8/2, p. 23-36.
- Martinot, C. (1994), *La reformulation dans des productions orales de définitions et explications*, Thèse de doctorat (NR), Université de Paris 8.
- Martinot, C. (2018), « Construction de l'information dans la langue parlée », in Richard, E. (éd.). *Des organisations dynamiques de l'oral*, Peter Lang, Bern, p. 237-253.
- Stabarin, I. (2017), « La complexité comme marqueur de spontanéité ? », in Martinot, C. & Ghoul, D. (éds), *Universalité et grammaire : paradoxe insoluble ou solution matricielle ?*, CRL, Paris, p. 40-51.
- Traverso, V. (2016), *Décrire le français parlé en interaction*, Ophrys, Paris.