



Usages du gérondif et du participe présent en français parlé et écrit : étude comparée basée sur corpus

*A corpus driven comparative study on the uses of present participle and
gérondif in spoken and written French*

Marie-Pierre Escoubas-Benveniste



Édition électronique

URL : <https://journals.openedition.org/tipa/908>

DOI : 10.4000/tipa.908

ISSN : 2264-7082

Éditeur

Laboratoire Parole et Langage

Référence électronique

Marie-Pierre Escoubas-Benveniste, « Usages du gérondif et du participe présent en français parlé et écrit : étude comparée basée sur corpus », *TIPA. Travaux interdisciplinaires sur la parole et le langage* [En ligne], 29 | 2013, mis en ligne le 17 décembre 2013, consulté le 03 novembre 2023. URL : <http://journals.openedition.org/tipa/908> ; DOI : <https://doi.org/10.4000/tipa.908>

Ce document a été généré automatiquement le 3 novembre 2023.



Le texte seul est utilisable sous licence CC BY-NC-ND 4.0. Les autres éléments (illustrations, fichiers annexes importés) sont « Tous droits réservés », sauf mention contraire.

Usages du gérondif et du participe présent en français parlé et écrit : étude comparée basée sur corpus ¹

A corpus driven comparative study on the uses of present participle and gérondif in spoken and written French

Marie-Pierre Escoubas-Benveniste

(0) « Ce livre est une espèce d'ouverture *en disant* eh bien ça peut être ça aussi la philosophie. »
Michel Onfray, Questions du public, Université populaire de Caen, 29/07/2011

1. Introduction

- 1 Le gérondif, dont on a affirmé qu'il était l'« une des formes de complément circonstanciel parmi les plus vivaces et les plus utilisées du français moderne, à l'oral comme à l'écrit » (Halmøy 2003 : 3) n'occupe pourtant pas, semble-t-il, une place centrale dans la langue parlée. En effet, une première analyse de ses emplois attestés dans deux modalités discursives – oral spontané vs écrit de presse – (Floquet *et al.* 2012), indique que cette forme verbale est proportionnellement beaucoup moins présente dans la modalité parlée que dans la modalité écrite.
- 2 Or, le gérondif (i.e. *en sortant*) ne constitue qu'une face de la médaille, l'autre étant le participe présent (i.e. *sortant*), deux formes verbales non finies du français constituant un système complexe et subtil où chaque élément peut aussi bien être contrôlé au sein même de la proposition par le verbe principal (ex. 1) ou par un nom (ex. 2), qu'introduire une prédication secondaire subordonnée à la prédication principale (ex. 1' et 2').

(1) « Le défi des enfants bilingues : grandir et vivre **en parlant** plusieurs langues » de B. Abdelilah-Bauer (2006)

(1') **En parlant** comme ça, tu vas te faire des ennemis.

(2) Le site permet aux parents **parlant** plusieurs langues d'échanger conseils et expériences pour élever leurs enfants.

(2') **Parlant plusieurs langues**, elle se débrouille toujours à l'étranger.

- 3 L'emploi de deux formes non finies morphologiquement distinctes (*gérondif vs participe présent*) pour exprimer ces relations entre les procès constitue une sorte d'exception française au regard d'autres langues romanes et en particulier de l'italien, par exemple, qui n'en utilise qu'une, le *gerundio* (1 et 2'. **parlando più lingue** ; 1'. **parlando così**) ou qui a recours à une relative (2. *che parlano*).² Ainsi, la question « ontologique » de la morphologie du gérondif et du participe présent a-t-elle été au centre de nombreuses recherches, qui ont abouti, rappelons-le rapidement, à deux positions opposées, l'une réduisant le gérondif à un participe avec préposition, l'autre considérant gérondif et participe comme deux modes distincts à part entière.
- 4 Si l'on examine les usages qu'en font les sujets parlants dans leur parole spontanée, il semble que dans certains cas le gérondif manifeste un fonctionnement qui « n'est pas sans rappeler celui d'une construction absolue (ce qui est d'ordinaire l'apanage du participe présent) » (*Ibid.* : 2147), c'est-à-dire qu'il recouvre des emplois que la norme de l'écrit aurait tendance à réserver au participe présent. On assisterait donc dans le français parlé à une sorte d'effacement de la frontière entre ces deux modes verbaux non finis :

(3)

21abm1_LG - BM : Enfin, ouais c'est, enfin moi en tout cas ça me fait rigolo, parce que c'est. C'est p/, c'est p/, c'est plus du tout, j'ai, je me, je suis dans le même ville, dans les mêmes lieux, mais avec euh, tu vois, une vie différente, tu vois, **en ayant** trois ans un peu, tu vois, où comme ça tu as coupé les ponts avec plein de gens. Et tu reviens, et en fait vu que mon a/, activité professionnelle là, tu sais, qui est, qui est relativement en contact avec pas mal de gens, tu te retrouves euh. [Source PFC] (*Ibid.* : 2148)

- 5 Quant au participe présent, Rossi-Gensane (2011 : 88) affirme que « l'oral a abondamment recours aux verbes, mais beaucoup moins aux participes, [...] Les participes présents (par exemple du type **Mangeant** peu, *il ne grossissait pas*) y sont en général peu représentés ».³ Selon l'auteur donc, les participes introduisant une prédication secondaire, à l'instar de 2', seraient rares à l'oral.
- 6 Pour amorcer un tableau complet de l'usage actuel des deux modes verbaux, dans une deuxième étude (Escoubas-Benveniste *et al.* 2012), nous avons inclus le participe présent à l'analyse différentielle des emplois modaux selon la variable diamésique et étendu l'analyse à un deuxième corpus d'oral spontané. Les résultats montrent qu'à la différence de l'italien, où l'usage quantitatif du *gerundio* ne varie pas de manière significative du point de vue diamésique, c'est-à-dire que sa présence à l'oral et à l'écrit est pratiquement identique (Giordano et Voghera, 2002 : 293), le français présente une

distribution différenciée et inversée des formes verbales en *-ant*. D'une part, gérondif et participe présent sont proportionnellement peu nombreux à l'oral par opposition à l'écrit ; d'autre part, leur usage selon la modalité orale spontanée vs écrite est inversé : à l'écrit les participes présents prédominent largement sur les gérondifs, à l'oral au contraire, c'est le gérondif qui domine (Escoubas-Benveniste *et al.* 2012 : 479).

- 7 Au-delà des considérations purement quantitatives, notre enquête empirique a mis en évidence, que ces tendances se confirment même lorsqu'on ne considère que les formes « libres » – du gérondif et du participe présent, à l'exclusion des formes grammaticalisées, pouvant présenter divers degrés de figement (i.e. *en supposant que, en supposant évidemment que, étant donné que*). Ces dernières, qui sont produites comme des unités lexicales figées de type adverbial, prépositionnel ou conjonctionnel plutôt que des constructions verbales libres, sont en outre beaucoup plus représentées dans le mode participe (42,5 %) que dans le mode gérondif (10,7 %) (*Ibid.* : 480). Cela oblige à relativiser encore la réelle disponibilité du participe comme mode verbal du parlé et pourrait s'expliquer par la tendance que manifeste le gérondif en français parlé à recouvrir certains emplois qui seraient dévolus au participe présent, selon les usages de l'écrit (Floquet *et al.* 2012).

2. Objectifs et méthode

2.1. Objectifs de la recherche

- 8 Notre premier objectif de recherche est de vérifier la reproductibilité des résultats obtenus dans nos précédentes études empiriques en réduisant encore les éventuels biais introduits par la différence de volume entre les corpus de l'oral et de l'écrit analysés jusqu'ici. Notre deuxième objectif est d'enrichir la description des faits linguistiques en y ajoutant la dimension diatopique des parlers. Ainsi avons-nous ajouté aux corpus du français parlé de France des échantillons du français parlé de Suisse romande et du Québec.
- 9 Nous reconduisons donc l'analyse quantitative et qualitative mise au point précédemment (Escoubas-Benveniste *et al.* 2012) sur trois nouveaux corpus oraux sélectionnés en fonction de leurs objectifs scientifiques de manière à enrichir quantitativement l'ensemble des données de l'oral spontané tout en y intégrant des échantillons de français parlé en Suisse et au Québec. Nous examinerons la validité de nos hypothèses concernant le statut des deux formes verbales en *-ant* dans le français parlé et dans une variété du français écrit. Nous évaluerons donc a) la rareté des formes verbales libres en *-ant* dans le français parlé par opposition au français écrit journalistique ; b) la distribution différentielle du gérondif et du participe présent dans les deux échantillons de données ; c) la disponibilité de ces modes verbaux dans le français parlé, compte tenu des phénomènes de grammaticalisation. Enfin nous analyserons certains aspects morphosyntaxiques, syntaxiques et lexico-sémantiques des emplois attestés à l'oral et de leurs variantes, ainsi que certains usages du gérondif qui semblent spécifiques de l'oral spontané.

2.2. Méthode de constitution des données

- 10 Les formes graphiques terminées par *-ant* sont fréquentes en français et ne sont pas nécessairement des verbes (i.e. *tant, pourtant, autant, instant...*). Lorsqu'elles le sont, elles se caractérisent par leur ambiguïté morphosyntaxique et par la discontinuité des formes graphiques. Les mots *consultant* et *attirant* peuvent en effet respectivement représenter des non-verbes (nom ou adjectif verbal) et des verbes (participe présent ou base verbale de gérondif). Dans le cas du gérondif, la forme verbale complète ne se limite pas au mot graphique puisqu'elle requiert la présence, non nécessairement contiguë de l'élément *en* – qui a son tour est ambigu, puisqu'il peut représenter le pronom clitique matérialisant un argument du participe – et éventuellement, la présence d'éléments syntaxiques entre ces deux constituants. La recherche automatisée des formes participiales et gérondivales suppose donc la mise en place d'une procédure d'extraction chronologiquement ordonnée qui tienne compte des problèmes méthodologiques que posent ces propriétés.
- 11 Le repérage des données utilisées pour l'analyse linguistique – formes verbales et contextes d'apparition – est le résultat de deux démarches distinctes en fonction des corpus : une analyse assistée par ordinateur et une exploration des corpus au moyen des outils de recherche mis à disposition par les auteurs sur les plateformes respectives des corpus.
- 12 L'extraction des formes a été effectuée au moyen du logiciel TaLTaC2 (Bolasco, 2010) qui opère sur le corpus de français écrit (L'Humanité 2010) et sur les transcriptions écrites de deux des corpus oraux (Phonologie du français contemporain et Corpus du français parlé parisien), non lemmatisés, à l'aide des fonctions suivantes : extraction des formes graphiques par requêtes élaborées au moyen de grammaires locales, catégorisation des formes selon des critères méthodologiquement pertinents, lexicalisation des segments répétés et extraction des concordances. Nous n'entrerons pas ici dans le détail de la procédure. Celle-ci a fait l'objet d'une description précise dans nos précédents travaux (Escoubas-Benveniste *et al.* 2012 ; 476-477) concernant notamment la grammaire locale utilisée pour l'extraction semi-automatisée des formes gérondivales (dont la structure syntaxique discontinue permet de capturer les clitiques arguments du verbe et la modalité négative) ainsi que le protocole utilisé pour le repérage des formes participiales dont le fort degré d'ambiguïté a nécessité une analyse manuelle des concordances.
- 13 Pour l'extraction des formes des trois corpus oraux (Enquête socio-linguistique d'Orléans, Corpus du français parlé du Québec et le Corpus oral de français de Suisse romande) nous avons utilisé les outils de recherche mis à disposition par les auteurs sur les plateformes respectives. Quoique non standardisés, ces outils proposent de manière générale une recherche avancée dans la transcription écrite (recherche d'un mot, d'une partie de mot, d'un segment borné à droite ou à gauche). Ils permettent presque toujours d'accéder au contexte élargi de la forme extraite telle qu'elle apparaît dans la transcription écrite⁴ ainsi qu'à l'enregistrement de l'extrait. Une recherche complexe peut généralement être effectuée par filtrage des corpus en fonction de l'identité socio-linguistique du locuteur, du type de situation langagière lorsqu'il en existe plusieurs. Lorsque le paramètre était pertinent (cf. corpus de français parlé de suisse romande) nous avons effectué un filtrage ne retenant que les locuteurs de langue maternelle. Le repérage des données – élimination des non-verbes en *-ant* et des faux-positifs (*en*

autant, en avant), désambiguïsation (*sachant que vs le sachant*), les analyses quantitative et qualitative – a été effectué manuellement à partir des concordances des formes en *-ant*.

3. Les corpus de français parlé et de français écrit

3.1. Corpus de français oral

- 14 Pour le présent travail nous avons eu recours à 5 corpus oraux. Aux deux ressources utilisées dans notre précédente étude pour illustrer le français oral spontané – le corpus Phonologie du français contemporain (dorénavant PFC) et le Corpus du français parlé parisien (CFPP) ont été ajoutées les données de 3 corpus oraux actuellement disponibles⁵ : l'Enquête socio-linguistique à Orléans (ESLO), le Corpus de français parlé au Québec (CFPQ), le Corpus oral de français de Suisse romande (OFRM). L'accroissement du volume de données de l'oral nous permet de vérifier la validité des hypothèses qui ont été formulées dans le premier et le deuxième volet de nos précédentes études. En outre, afin d'enrichir l'étude variationniste de l'usage des deux modes verbaux dans le français parlé, il nous a paru intéressant d'ajouter à la dimension diamesique des travaux précédents la dimension diatopique en comparant l'usage que les locuteurs en font dans le français parlé sur l'ensemble du territoire français (PFC, CFPP, ESLO), au Québec (CFPQ) et en Suisse romande (OFRM).
- 15 Le projet international PFC (Phonologie du français contemporain) a pour but explicite de « construire un important corpus de référence permettant de rendre compte de la diversité des usages oraux du français sur le territoire national, mais aussi dans l'ensemble de la francophonie » (Laks 2011 :10). Nous reprenons ici les données issues de nos deux explorations précédentes portant exclusivement sur les enquêtes menées sur le territoire métropolitain (15 départements répartis sur l'ensemble du territoire pour un total de 280 dialogues analysés). Chaque enquête suit un protocole commun et se subdivise en quatre parties : une discussion libre, une discussion guidée, une lecture de texte et de liste de mots. Pour l'analyse, qui porte sur l'usage des formes verbales en *-ant* dans le français tel qu'il est parlé spontanément, nous avons retenu la discussion libre et la discussion guidée. La partie du corpus PFC analysée représente 529 543 occurrences.
- 16 Le Corpus du français parlé parisien (CFPP) est composé d'un ensemble d'interviews d'habitants de la région Île-de-France portant sur la ville : les quartiers de Paris et sa proche banlieue. Le corpus est constitué d'enregistrements longs (entre 60 et 105 minutes, soit en moyenne 20 000 mots) et a une durée globale de 36,6 heures d'enregistrement soit 483 152 occurrences.
- 17 L'Enquête socio-linguistique à Orléans (ESLO)⁶ est une initiative entreprise en 1969 consistant à collecter, à des fins didactiques, des documents sonores illustrant le français parlé au quotidien, dans les années soixante-dix, dans cette région. L'enquête (en voie d'enrichissement) comprend environ 200 enregistrements réalisés dans différentes situations, soit au total plus de 300 heures de parole incluant des interviews en face à face et tout une gamme d'enregistrements variés (conversations téléphoniques, réunions publiques, transactions commerciales, repas de famille, entretiens médico-pédagogiques, etc.). Les documents sonores et leurs transcriptions sont accessibles et il est possible de rechercher un mot ou une partie de mot (chaîne de

caractères) en faisant une requête sur tout ou partie du corpus. Pour affiner la recherche, sont disponibles des filtres liés à certaines caractéristiques sociodémographiques des locuteurs (sexe, âge, profession, niveau d'étude) et à la typologie des situations de communication enregistrées, pour la langue parlée. Le corpus analysé est composé de 1 597 273 segments.⁷

- 18 Le Corpus de français parlé au Québec (CFPQ) initié en 2006 offre un échantillon du français québécois en usage dans les années 2000 et entend refléter au plus près le français québécois spontané, c'est-à-dire tel qu'il se pratique « dans les interactions informelles où la parole a libre cours ».⁸ La situation communicative prise en considération est toujours l'interaction verbale en face-à-face, ce qui caractérise le corpus comme multimodal car il codifie graphiquement la dimension verbale, la dimension paraverbale (éléments prosodiques et événements vocaux qui accompagnent la parole) et la dimension gestuelle associée aux flux de la parole. Le corpus est subdivisé en sous-corpus quantitativement homogènes qui résultent de la transcription des propos tenus pendant environ 1h30 par trois ou quatre personnes se connaissant bien (parents ou amis) qui connaissent également l'étudiant responsable de l'enregistrement. Les sujets ne sont pas planifiés et varient selon les intérêts des personnes en présence (les adolescents parlent de l'école, de leurs amis ; les jeunes adultes discutent de leur travail, de la vie à deux, etc.). Au moment de la recherche il est possible de sélectionner les sous-corpus en fonction de trois paramètres sociodémographiques caractérisant les locuteurs : le sexe, l'âge, le niveau de scolarisation. Le CFPQ se compose de 427 667 occurrences.
- 19 Le Corpus oral de français parlé en Suisse Romande OFROM⁹ a été réalisé à partir de 2008 et se compose en majeure partie d'extraits d'entretiens guidés, à dominante monologique, dans lesquels l'interviewé (un locuteur né en Suisse, et vivant en Suisse romande) est sollicité pour répondre à des questions nécessitant des réponses plus ou moins longues posées par l'intervieweur (le responsable de l'enquête) et de quelques interactions impliquant au moins deux personnes qui parlent à bâtons rompus. Les thèmes abordés concernent les métiers, les voyages, les passe-temps des locuteurs, leurs relations de voisinage, leurs projets ou les situations auxquelles ils ont été confrontés dans leur vie. Elles peuvent également être en rapport avec le système politique ou la situation linguistique de la Suisse, etc. L'ensemble des enregistrements disponibles au moment de la recherche compte 156 911 occurrences.

3.2. Corpus du français écrit

- 20 Pour la confrontation des usages de l'oral spontané et de l'écrit nous nous référerons à nos analyses précédentes d'un corpus de presse daté de 2010. Le sous-corpus analysé est constitué de l'ensemble des articles de L'Humanité publiés cette année-là (ci-après LH 2010). Il est composé de 6 873 329 occurrences¹⁰ et il appartient au « Scriptorium numérique FANUM » de l'Université de Franche-Comté, qui regroupe l'ensemble des articles publiés de 2006 à 2010 par cinq quotidiens nationaux français. Ce que nous comparerons au corpus d'oral spontané sera donc un corpus homogène exclusivement représentatif du type particulier d'écrit que constitue l'article de presse. Sans préjuger de l'éventualité que ce type de corpus d'écrit puisse présenter de réelles spécificités en ce qui concerne l'emploi des verbes en *-ant*, par rapport à d'autres types d'écrit (i.e. textes littéraires, textes juridiques, forums internet...), nous avons postulé, pour les

besoins de l'étude, qu'il constitue un échantillon représentatif de la norme actuelle du français standard avec laquelle il est possible de confronter les usages des verbes en -ant que la parole spontanée manifeste.

4. Analyse quantitative selon la modalité orale/écrite¹¹

4.1. Distribution des gérondifs / participes présents et variété lexicale

- 21 Le tableau 1 présente les résultats de l'analyse quantitative qui met en évidence la distribution des formes gérondivales et des formes participiales¹² dans l'ensemble des données de l'oral et de l'écrit. Le corpus d'oral pris dans son ensemble représente un volume d'environ 3,2 millions de mots, soit près de la moitié du volume du corpus de langue écrite (6,8 millions de mots). Ce rapport de 1 à 2 rend plus significative la comparaison quantitative des résultats selon les deux modalités expressives.¹³
- 22 N représente le nombre d'occurrences de chacun des modes verbaux (*tokens*) et V(LE), le nombre de lemmes verbaux distincts (*types lexicaux*) employés au participe présent et au gérondif.

Tableau 1. Distribution des Gér/PP^{ant} selon la modalité orale/écrite : densité et types lexicaux

	Corpus oraux confondus		LH 2010	
Nombre de FG	3 157 152		6 873 329	
Densité relative Gér	2,4 x 10 ⁻⁴		14,0 x 10 ⁻⁴	
Densité relative Ppant	1,2 x 10 ⁻⁴		33,8 x 10 ⁻⁴	
	V(LE)	N	V(LE)	N
Gér	314	764	1 282	9 657
Ppant	93	374	1 901	23 239

- 23 Un rapide examen du tableau 1 permet de confirmer les tendances observées dans nos premières études empiriques. D'une part les données manifestent la raréfaction extrême du participe présent dans les corpus oraux, alors qu'il est proportionnellement bien plus présent dans le corpus écrit : la probabilité de rencontrer cette forme est de 1,2 pour 10 000 (374 occurrences sur 3 millions de mots) dans l'oral spontané et de 33,8 pour 10 000 dans l'écrit journalistique, soit 30 fois plus élevée (23 239 occurrences sur 6,8 millions de mots).
- 24 D'autre part, la même observation vaut pour le gérondif à l'oral, même si elle est moins marquée. Cette raréfaction du gérondif (2,4 gérondifs pour 10 000 mots), déjà décrite pour un corpus d'entretiens libres et guidés (Floquet *et al.* 2012), contraste avec un usage qui reste bien établi dans le corpus de presse écrite (14,0 gérondifs pour 10 000 mots soit 7 fois plus). Par conséquent, le fait d'avoir étendu l'analyse à des données

orales issues de différents types de situation de communication (cf. ESLO qui en offre une multiplicité), ne contredit pas les résultats de nos études antérieures.

- 25 Au-delà d'un usage quantitativement très différent, en liaison avec les types de corpus considérés, la tendance mise en évidence dans notre précédente étude (Escoubas-Benveniste *et al.* 2012) se confirme également : la fréquence d'usage des deux modes s'inverse selon le type de corpus. En effet si dans l'écrit de presse, les formes participiales sont beaucoup plus fréquentes (2,4 fois plus) que les formes gérondivales, dans l'oral spontané ce sont les gérondifs qui sont 2 fois plus nombreux que les participes présents.
- 26 Cette observation vaut également pour chacun des corpus d'oral pris séparément, à quelques légères nuances près :¹⁴

Tableau 2. Occurrences (N) et types lexicaux (V(LE)) des Gér/PP^{ant} dans les corpus oraux

	PFC		CFPP		ESLO		CFPQ		OFROM	
N	529 543		483 152		1 559 879		427 667		156 911	
	V(LE)	N	V(LE)	N	V(LE)	N	V(LE)	N	V(LE)	N
Gér	72	200	66	194	79	178	54	115	43	77
PPant	12	104	24	108	29	109	15	27 ¹⁵	13	26

- 27 Systématiquement plus nombreuses que les participes, les formes gérondivales le sont dans une proportion sensiblement variable selon les corpus oraux (PFC et CFPP : environ 2 fois plus ; ESLO : 1,6 fois plus ; CFPQ : 4,25 fois plus ; OFROM : 2,9 fois plus).
- 28 L'analyse de la dimension lexicale des verbes montre que la variété du lexique, dont on sait qu'elle est bien supérieure dans l'écrit formel par rapport à l'oral spontané, présente cette même variation au niveau des formes gérondivales et, dans une plus grande mesure encore, au niveau des participes : selon le tableau 1, les formes *en V^{ant}* sont l'expression de 314 classes distinctes dans le corpus oral et de 1282 (4 fois plus) dans le corpus LH2010. Les formes participiales *V^{ant}* sont l'expression de 93 classes distinctes dans le corpus oral et de 1901 (10 fois plus) dans le corpus écrit.
- 29 Si l'on compare la variété lexicale manifestée à l'oral dans chaque mode, on constate que le gérondif s'actualise globalement à travers une plus grande variété de verbes que le participe (6 fois plus (PFC), environ 3 fois plus (CFPP, ESLO, CFPQ, OFROM)). Or les occurrences de participes sont moitié moins nombreuses que les occurrences de gérondif. Par conséquent il est prévisible que certains types lexicaux employés soient particulièrement récurrents, tout particulièrement dans le mode participe. La proportion des hapax, c'est-à-dire des types lexicaux n'ayant qu'une occurrence dans le corpus, contribue à moduler cet aspect.¹⁶

4.2. Verbes les plus fréquents dans les deux modes et les deux modalités

- 30 L'examen des lemmes les plus fréquemment associés aux deux modes verbaux dans les deux modalités discursives (tableaux 3 et 4) montrent que, à l'oral, *dire*, *faire*, *être* sont régulièrement les plus fréquents dans les emplois gérondivaux tandis que *être*, *avoir*, *parler*, *suivre* le sont dans les emplois participiaux – résultat qui est également en harmonie avec les fréquences des verbes en *-ant* (les deux modes confondus) issus du corpus d'Aix (*disant*, *étant*, *suivant*, *parlant*, *ayant*, *faisant*).¹⁷ Certains de ces verbes sont aussi parmi les plus fréquents dans le corpus de presse (*en faisant*, *en étant* et *ayant*, *étant*, *faisant*). Ajoutons que les trois lexèmes gérondivaux les plus fréquents dans les données orales coïncident avec les trois lexèmes verbaux les plus fréquents du Français fondamental : 1-*être*, 2-*faire*, 3-*dire*, (Gougenheim et al. 1958, 1977) à quelque nuance près toutefois concernant leur rang puisque le gérondif de *dire* s'affirme comme le plus fréquent des trois. Cette forte présence de *en disant* semble être l'une des caractéristiques des emplois gérondivaux dans nos données de l'oral spontané.

Tableau 3. Lexèmes les plus fréquemment associés aux Gér. (c. oraux vs c. écrit)

Gér	PFC et CFPP	ESLO	CFPQ	OFROM	LH2010
1	en disant (56)	en disant (22)	en faisant (8)	en disant (13)	en faisant (369)
2	en étant (27)	en mettant (12)	en revenant (8)	en faisant (9)	en mettant (139)
3	en faisant (21)	en faisant (11)	en vieillissant (7)	en étant (4)	en prenant (137)
4	en rentrant (21)	en allant (6)	en attendant (6)	en attendant (4)	en étant (105)

Tableau 4. Lemmes les plus fréquemment associés aux Pp^{ant} (c. oraux vs c. écrit)

Pp ^{ant}	PFC et CFPP	ESLO	CFPQ	OFROM	LH2010
1	étant (84)	étant (41)	étant (9)	étant (7)	ayant (853)
2	ayant (19)	suivant (22)	parlant (4)	ayant (4)	étant (814)
3	sachant (14)	ayant (11)	travaillant (3)	concernant (3)	faisant (489)
4	parlant (7)	voulant (6)	descendant (3)	suivant (3)	visant (483)

- 31 Les attestations de l'oral montrent par ailleurs que certains verbes sont utilisés dans les deux formes (*pouvoir*, *vouloir*, *savoir*, *rester*, *suivre*, *attendre*, *être*) ; d'autres, seulement au Pp^{ant} (*avoir*, *concerner*, *s'adresser*, *comprendre*), d'autres encore, seulement au gérondif (*comparer*, *donner*, *demandeur*, *employer*, *mettre*, *partir*, *passer*, *rentrer*, *revenir*, *sortir*, ...), parmi lesquels on remarque une nette proportion de verbes de mouvement. (cf. § 6.4 pour l'analyse des attestations).

- 32 Les données issues des nouveaux corpus oraux confirment qu'à eux seuls, *être* et *avoir* représentent les formes de Pp^{ant} les plus fréquentes de l'oral et de l'écrit. Celles-ci sont grammaticalement très ambiguës puisqu'il peut s'agir de participes simples des verbes sémantiquement pleins *être* et *avoir* (i.e. *les gens étant là ; mon frère ayant lui aussi une moto*), de l'auxiliaire du passif (i.e. *une espèce étant appelée à disparaître*) ; de l'auxiliaire exprimant l'aspect accompli du « participe composé » (cf. Arnavielle, 2003 : 40) (i.e. *étant arrivé en retard ; s'étant levé de bonne heure ; ayant marché longtemps ; ayant été admis*) dont la valeur temporo-aspectuelle en fait un objet linguistique bien différent du participe présent. Nos données confirment que les participes composés sont plus fréquents que les gérondifs composés. Cependant l'oral spontané n'est toutefois pas dépourvu de certains usages de gérondifs composés qui n'ont pas cours dans le corpus d'écrit (cf. § 6.5 pour analyse des attestations).

5. Analyse qualitative des formes

5.1. Formes libres et formes grammaticalisées

- 33 Un examen des contextes d'usage des formes gérondivales et participiales fait apparaître deux types de gérondifs et de participes présents : des formes qui semblent libres et des formes qui semblent résulter d'un processus de lexicalisation, voire de grammaticalisation. Les premières ne sont pas figées, il s'agit de faits de discours : le participe présent ou le gérondif résultent d'une construction syntaxique élaborée par le locuteur (ex : *en n'insistant absolument pas*,) ; les secondes sont des usages plus ou moins figés (ex : *moyennant quoi*), il s'agit de faits de figement, à savoir d'usages de ces constructions verbales qui sont actuellement répertoriées et attestées en langue française comme adverbe, locution conjonctive ou prépositionnelle.
- 34 Pour le recensement des formes grammaticalisées du gérondif nous faisons référence à l'inventaire proposé par Halmøy (2003 : 77-79)¹⁸ afin de rendre compte de ces formes qui ne semblent pas totalement libres. Les cas ne sont pas aisés à départager si l'on pense par exemple aux verbes de mouvement et de déplacement et au système complexe de complémentation qu'ils régissent. Par conséquent l'analyse requiert une étude approfondie et l'identification de critères syntaxiques précis, dans un contexte où les indices sont ténus, afin de délimiter objectivement les étapes du figement (cf. l'étude récente du processus de grammaticalisation de *en passant par*).¹⁹ Nous optons donc, dans le droit fil de nos précédentes études empiriques, pour une solution maximaliste qui consiste à considérer comme grammaticalisées toutes les occurrences de l'inventaire (cf. note 14), afin de permettre d'estimer la probabilité d'occurrence des formes réellement libres.
- 35 Pour les formes grammaticalisées du participe présent, nous nous appuyons sur les attestations recensées dans le dictionnaire de langue (Petit Robert 2008) de participes qui assurent le rôle d'une autre partie du discours que le verbe (i.e. *concernant, suivant, donnant-donnant, (professionnellement) parlant, soi(-)disant, étant donné, étant donné que, ceci étant, moyennant*), auxquelles nous ajoutons *sachant que* et *pensant que* même si ces syntagmes ne sont pas explicitement catégorisés comme des conjonctions par le dictionnaire.²⁰

Tableau 5. Formes libres et grammaticalisées des Gér./ Pp^{ant} à l'oral et à l'écrit

	PFC	CFPP	ESLO	CFPQ	OFROM	Corp. oraux	LH 2010
N	529 543	483 152	1 559 879	427 667	156 911	3 157 152	6 873 329
N Gér gramm	18 (9%)	24 (12%)	14 (8 %)	9 (8%)	5 (7%)	70 (9 %)	619 (6%)
N Gér. libres	182 (91%)	170 (88%)	167 (92%)	107 (92%)	71(93%)	697 (91%)	9 038 (94%)
Total des f. gérondivales	200 (100%)	194 (100%)	178 (100%)	115 (100%)	76 (100%)	763 (100%)	9 657 (100%)
Densité Gér ¹	3,4	3,5	1	2,5	4,5	2,2	13,1
N Pp ^{ant} gramm	51 (49%)	39 (36%)	38 (53%)	11 (41%)	9 (35%)	148 (40%)	138 (1%)
N Pp ^{ant} libres	53 (51%)	69 (64%)	71(47%)	16 (59%)	17 (65%)	226 (60%)	23 101 (99%)
Total des f. participiales	104 (100%)	108 (100%)	109 (100%)	27 (100%)	26 (100%)	374 (100%)	23 239 (100%)
Densité Pp ^{ant}	1,0	1,4	0,4	0,3	1,0	0,7	33,6

¹. Exprimée par le nombre de formes libres rencontrées tous les 10 000 mots.

- 36 Le tableau 5 présente le pourcentage des formes libres et grammaticalisées pour chacun des modes verbaux par rapport au total des formes gérondives et participiales de l'oral et de l'écrit. Il permet d'illustrer la densité des gérondifs et participes libres par rapport au nombre total de mots du corpus. Ce chiffre, qui représente le nombre de formes libres rencontrées tous les 10 000 mots, constitue une indication sur la probabilité que le mode verbal soit attesté. En confrontant ces valeurs, on remarque que :
- hormis dans les corpus ESLO (cf. note 7) et CFPQ, les densités respectives des formes verbales libres en *-ant* relèvent du même ordre de grandeur pour chacun des corpus oraux ;
 - la tendance observée pour la répartition des formes à bases gérondive et participiale se vérifie pour les formes libres : le gérondif est 3 fois plus présent que les formes libres de participe présent à l'oral et cette répartition est inversée à l'écrit ;
 - le fait que les formes grammaticalisées sont beaucoup plus représentées dans le mode participe que dans le mode gérondif à l'oral (cf. Escoubas-Benveniste *et al.* 2012 : respectivement 42,5 % vs 10,7 %) n'est pas démenti ici (Pp^{ant} grammaticalisés : 40 % vs Gér. grammaticalisés : 9 %). Cette tendance oblige à relativiser encore la réelle disponibilité du participe présent, déjà rare, comme mode verbal du parlé ;
 - le participe reste en revanche une construction libre de l'écrit avec à peine 1 % de ses occurrences qui sont des participes grammaticalisés.

5.2. Gérondifs, clitiques et négation

- 37 Afin de compléter l'analyse qualitative différentielle des formes du gérondif, il convient d'examiner le degré de complexité syntaxique de la structure en liaison avec les emplois de l'oral.²¹ Nous posons que cette complexité est proportionnelle au nombre d'éléments insérés au sein de la structure. En tant que structure discontinue à deux constituants, le gérondif peut porter la trace de la modalité négative ainsi que d'un processus de pronominalisation. Les formes clitiques s'insérant entre *en* et la forme verbale en *-ant* peuvent être au nombre de deux ce qui amène à 1, 2, ou 3 le nombre d'éléments potentiellement insérés si l'on tient compte de la négation (i.e. : *en le leur disant ; en ne le leur ayant jamais dit*). Les formes graphiques des insertions sont au nombre de 18 (pronoms : *en, la, le, l', les, leur, lui, me, m', te, t', se, s', nous, vous, y* ; et négation : *ne, n'*).²²

Tableau 6. Nombres d'éléments insérés dans les formes verbales au Gér. à l'oral et à l'écrit

	PFC	CFPP	ESLO	CFPQ	OFROM	Oral	LH 2010
N Gérondif	182	170	167	107	71	697 (100 %)	9 038 (100 %)
Gér. 0-insert	169	137	131	91	55	588 (84 %)	8445 (93 %)
Gér 1-insert	13	32	36	13	15	109 (16 %)	578 (6,4 %)
Gér 2-insert	0	0	0	3	1	4 (0,6 %)	15 (0,2 %)
Gér négation ²³	0	0	0	0	1	1 (0,1 %)	53 (0,6 %)

- 38 Proportionnellement les gérondifs à 1 insertion sont près de 3 fois plus nombreux dans les corpus oraux que dans le corpus de presse (16 % vs 6,4 %). En revanche les gérondifs négatifs sont presque inexistantes dans l'oral spontané (0,1 %) alors qu'ils sont 6 fois plus présents à l'écrit, même si la proportion reste très faible. Il n'est peut-être pas sans intérêt de souligner que la seule attestation de gérondif à négation complète sur l'ensemble des données de français parlé analysées coïncide avec un inachèvement de la construction :

(4)

unine08-yba : on a en en ouais *en n'ayant pas* après deux deux mois avec elle j'arrivais vraiment à [Source OFROM]

- 39 L'observation selon laquelle « les attestations de plus d'un clitique sont inexistantes à l'oral et tout à fait négligeables à l'écrit » (Floquet *et al.* 2012 : 2147) n'est pas contredite par les nouvelles données de l'oral. Les gérondifs à deux éléments insérés y sont en effet extrêmement rares.²⁴ Qui plus est, à en juger par l'exemple suivant, l'utilisation de cette construction syntaxique avec pronoms clitiques peut être amorcée puis abandonnée pour une expression en extension de l'objet verbal (*aux aux élèves*), sans toutefois occasionner de rupture dans le flux de la parole spontanée :

(5)

j' trouve que la pédagogie devrait quand même + + devrait plus évoluer + *en les en permettant* aux aux élèves de réfléchir par eux-mêmes d'avoir des outils de recherche de + + de travailler entre pairs [Source CFPP]

- 40 Quant aux 3 occurrences à 2 insertions attestées dans le français parlé québécois (*en s'en allant, en s'en revenant, en m'en venant*), leur statut de constructions résultant d'une construction syntaxique cliticisée est douteux puisqu'il s'agit de verbes lexicalement pronominaux qui, à l'instar du verbe de mouvement *s'en aller*, ont un pronom *en* lexicalisé représentant la source du mouvement. Ces trois formes gérondivales ne sont donc pas le résultat d'opérations syntaxiques de pronominalisation. Sur 4 occurrences à double insertion, une seule, sur l'ensemble des corpus oraux, illustre une construction gérondivale à 2 clitics. Il s'agit d'une construction causative en *faire* + *VInf* associée à

une mise en relief d'un des arguments par reprise pronominale en post-position (*me/moi*) :

(6)

unine08-aca : qu'il m'a transmis et puis euh tout au long *en me les faisant travailler moi* donc euh [Source OFROM]

6. Analyse de certains usages de l'oral

- 41 Le gérondif, indiscutablement plus présent à l'oral que le participe, manifeste deux types d'usages : ceux qui appartiennent à la norme du français écrit standard, tel que l'écrit de presse, et qui offrent quelques variantes (§ 6.1 et 6.2) et ceux qui constituent des spécificités de l'oral. Nous examinerons ainsi les gérondifs juxtaposés et les relations syntactico-sémantiques sous-jacentes (§ 6.3) ; la récurrence de motifs associés au mode gérondif (expression du mouvement et de la localisation) (§ 6.4), la présence de gérondifs composés comme prédication seconde à valeur causale (§ 6.5) et enfin la tendance qu'ont les frontières des emplois des forme en *-ant* à s'estomper dans les usages de l'oral spontané (§ 6.6).

6.1. Emploi disjonctif : *tout + en V^{ant}* et quelques variantes

- 42 La construction disjonctive (ou concessive) *tout + en V^{ant}* est bien présente à l'oral (hormis dans le corpus de français parlé québécois), avec des verbes très divers (états, actions, verbes psychologiques...). Les attestations ci-après illustrent combien les nuances sémantiques de cette construction sont variables, allant de la concession (où *tout* peut être remplacé par *même* (7)) à l'opposition qui peut être explicitée par un marqueur (8), en passant par l'expression purement temporelle d'une simultanéité dans la durée (9, 10) :

(7)

j'avais déjà compris que on pouvait être de nationalité française *tout en étant* intérieurement camerounais et français aussi [Source CFPP]

(8)

11amg2_GG.txt : a acheté cette maison justement pour pouvoir se la retaper, mais *tout en gardant* son style ancien . MG2 : C' est , c' est ce qui lui plaît , sinon elle [Source PFC]

(9)

unine09-lba : *tout en parlant* il y en a un qui avait mis une une cassette vidéo sur une télé de de catch [Source OFROM]

(10)

31afr1_GG.txt : interne de la banque . FR : donc il a gravi les échelons euh , *tout en travaillant* à la banque . FR : Mais , parce que , ses études n' ont pas été fructueuses [Source CFPP]

- 43 La construction peut être contrôlée par un verbe principal impersonnel, le point d'ancrage sémantique de son sujet syntaxique zéro²⁵ est alors à rechercher dans l'énoncé précédent voire dans le contexte pragmatique :

(10')

419PERS : et tout en se connaissant

419PERS : y a une rivalité qui des fois est même euh cinglante enfin méchante même des fois [Source ESLO]

- 44 Le référent du sujet \emptyset peut être indéfini, mais quantifié (*chacun* sujet principal) et faire l'objet d'un réajustement sémantique avec une reprise accentuée en post-position (*tous*) qui permet d'assurer une plus grande cohérence au sens concessif de la construction :

(11)

euh c' qu'on avait à faire on l' faisait donc j' crois quelque part voilà *chacun* menait un peu sa vie séparément tout en habitant tous sous l' même toit [Source CFPP]

- 45 À l'oral, la valeur concessive est également exprimée par *même + gérondif* ou *même + participe présent* :

(12)

comment dire à valoriser les diplômés + et euh *même en étant pas diplômés* ils ont plus de facilités qu' nous [Source CFPP]

(13)

424PERS : *même étant petit* la structure était la même que pour euh un atelier plus important [Source ESLO]

(14)

E : c'est pas ça la question c'est une a- c'est une manière de dire que *MÊME en leur donnant des subventions aux écoles privées* même si elles peuvent BAISSER leur prix un peu <f-ça reste quand même une élite là ↑ le trois quarts du monde peuvent pas envoyer leurs enfants là pareil là voyons donc surtout quand t'as deux trois enfants là/>> «<419704> [sous-corpus 2, segment 5, page 62, ligne 1] [Source CFPQ]

6.2. *Aller + en V^{ant}* et quelques variantes

- 46 La construction gérondivale *aller + en V^{ant}* où le verbe *aller*, suivi d'un gérondif de verbe de changement d'état, fait office de verbe support d'aspect progressif est bien attestée à l'oral :

(15)

417PERS : et euh si cette situation va en s'atténuant un petit peu [Source ESLO]

(16)

S : non parce que ça du un excès de poids là ça te suit toute ¹ta vie là ça va pas ²euh :: : *en diminuant* là <resc<plus que tu vieillis ³justement plus qu'il faut que tu

fasses attention>> «<118033> [sous-corpus 1, segment 7, page 89, ligne 15] [Source CFPQ]

- 47 Avec des variantes liées à l'usage d'un synonyme du verbe support d'aspect *aller* : *avoir été+ en V^{ant}* :

(17)
417PERS : les écarts de salaire entre la région parisienne et la province *ont été en se*
417PERS : *en s'amenuisant* depuis quinze ans [Source ESLO]

- 48 Et en français québécois, *s'en aller + en V^{ant}*, où l'on remarque que la préposition du gérondif se maintient, malgré l'effet possible de cacophonie avec le pronom clitique :

(18)
J : 'parce que ça s'en va pas en s'améliorant on dirait «<110131> [sous-corpus 6, segment 6, page 81, ligne 9] [Source CFPQ]

(19)
G : {Ø ;ben} je croirais que ça s'en va en améliorant [sous-corpus 1, segment 4, page 53, ligne 20] [Source CFPQ]

6.3. Accumulations de gérondifs et relations syntaxiques

- 49 On observe à l'oral des accumulations de gérondifs comme en (20), qui pourrait indiquer une parenthèse (*rien qu'en y pensant*) ou une reformulation (*en se disant*), voire la succession logique de procès psychologiques. Selon cette dernière interprétation, elle ferait alors office d'une coordination implicite, c'est-à-dire d'une expression dont le sens peut être paraphrasé par la conjonction de deux éléments coordonnés (*on y pense juste un instant*) et *on se dit*) :

(20)
Unine08-mba même rien qu' *en y pensant en se disant* ben tiens euh [Source OFROM]²⁶

- 50 Mais la juxtaposition n'est pas toujours interprétable comme une coordination. Les trois exemples suivants ne peuvent s'analyser comme la succession de procès placés au même niveau syntaxique de l'énoncé. Il s'agit vraisemblablement d'enchâssements de procès, plus ou moins facilement démontrables, selon le contexte et selon la nature sémantique des lexèmes verbaux.
- 51 Dans le premier exemple, le contexte fait immédiatement apparaître que les deux gérondifs juxtaposés appartiennent à deux niveaux syntaxiques distincts et assurent des fonctions différentes.

(21)
64 - L : ³c'est parce que tu t'en vas pas *en maigrissant en vieillissant* ⁴là/ tu tu prends du poids ⁴en vieillissant là\ [sous-corpus 1, segment 7, page 90, ligne 5] [Source CFPQ]

- 52 Le gérondif *en maigrissant* fait partie intégrante de la prédication principale (construction marquée associant variation et aspect duratif : *aller* + gérondif) et n'est pas détachable. La forme négative revêt une valeur rhétorique (litote ou euphémisme), comme le prouve la reformulation immédiatement consécutive (*tu prends du poids*). Le gérondif *en vieillissant* constitue en revanche une prédication seconde détachable et antéposable : « *c'est parce que en vieillissant tu t'en vas pas en maigrissant* » voire « *c'est en vieillissant que tu t'en vas pas en maigrissant* ». Elle est donc régie par la construction *aller* + gérondif : son sujet syntaxique zéro s'inscrit dans une relation de coréférence avec le sujet du verbe fini et la référence de l'agent de *en vieillissant* s'établit dans le cadre syntaxique du domaine de la rection du verbe fini *tu t'en vas pas en maigrissant*.
- 53 Le deuxième exemple se distingue du précédent par un paramètre syntaxique :

(22)

É : pensez-vous ↑ que ça s'en va en amé- en s'améliorant ¹en regardant ²peut-être le le chum de Valérie euh mon frère «<407428> [sous-corpus 1, segment 4, page 53, ligne 12]
[Source CFPQ]

- 54 Ici, le gérondif *en regardant* n'est pas détachable et antéposable (**c'est en regardant que ça s'en va en amé- en s'améliorant*). Il se situe donc hors du domaine de la rection du verbe principal. De plus, du point de vue sémantique, la coréférence du sujet zéro de *en regardant* au pronom neutre *ça*, le sujet principal, est impossible pour d'évidentes raisons liées aux propriétés de sélection des sujets des verbes de perception. Par conséquent le référent de l'agent de *en regardant* est établi par inférence pragmatique, hors du domaine syntaxique de la rection : il a pour référents les acteurs de l'interaction. Le procès exprimé par *en regardant* se situe au niveau de l'énonciation et l'élément qui le contrôle sémantiquement est à rechercher dans les traces du sujet parlant qui se manifeste dans la suite de l'interaction en tant que sujet de l'énonciation.
- 27
- 55 En revanche, l'analyse du troisième exemple est beaucoup plus difficile à justifier car elle ne peut s'appuyer que sur des considérations sémantiques (V d'action / V d'état) et pragmatiques (*l'école est un lieu où des personnes gardent des enfants*).

(23)

tu vois on peut xx ça § ça *en gardant des enfants en étant à l'école* j' ai rencontré énormément de familles juives [Source CFPP]

- 56 L'énoncé est ambigu et interprétable de deux manières : comme une coordination implicite d'une activité et d'un état (*en gardant... et en étant...*) qui ont été la cause d'un résultat (*j'ai rencontré*) exprimé par le prédicat principal ; ou, ce qui nous paraît plus réaliste, comme une prédication seconde de cause contrôlée par (*j'ai rencontré*) exprimée par un premier procès au gérondif (*en gardant...*) qui est enchâssée dans un cadre temporel exprimé par le deuxième gérondif (*en étant...*). Celui-ci constitue alors un troisième niveau de prédication – contrôlé par la prédication seconde – qui peut s'interpréter comme une proposition circonstancielle à valeur causale (*puisque j'étais...*) ou temporelle (*pendant que j'étais...*) et pourrait fonctionner comme une glose explicative.

6.4. « Gérondifs de mouvement » et « gérondifs de localisation »

- 57 Nombre de verbes de mouvement sont utilisés au gérondif à l'oral alors qu'ils n'apparaissent pas au participe présent. Ce phénomène peut s'expliquer par l'existence d'une « construction obligée » en français dont la “nature « libre » [...] est à relativiser, précisément parce que, dans ce cas, sa présence est comme régie par une loi pour ainsi dire « systémique » du français” (Escoubas-Benveniste et al. 2012 : 484) : pour exprimer simultanément dans cette langue le déplacement et la manière de se mouvoir le gérondif de manière de mouvement est le schéma syntaxique par défaut (cf. *partir en courant* vs *to run away*). Au-delà de ces emplois, l'oral spontané présente des gérondifs de mouvement associés à certains prédicats de localisation spatiale (*être là, il y a là*) :

(24)

F : ouais on a marché là pendant nos trois jours on a fait du vélo il y a des vélos partout ¹même là ²tu vois depuis deux ans ³là tu tu peux faire même : à l'hôtel où on était là *il y avait direct les vélos en sortant là* que tu pou ⁴vais avoir fait que ⁵c'était le fun [*sous-corpus 18, segment 8, page 82, ligne 1*] [Source CFPQ]

(25)

Anatole France § Anatole France *c'est en allant + en allant à Franprix + où y a l' IUT maintenant + vous savez* [Source CFPP]

- 58 Où il est clair que *en allant à X*, qui peut être remplacé par l'adverbe *là* (*il y avait les vélos là ; c'est là où y a l'IUT maintenant*) constitue le cadre locatif, l'itinéraire vers une destination sur lequel se situe le point à localiser. Dans le processus de localisation spatiale, le gérondif, en tant que procès virtuel de déplacement, sert à fournir un repère en référence duquel doivent être interprétées les autres indications spatiales.
- 59 La localisation peut être également celle d'un événement dans un cadre temporel :

(26)

seize ans ouais § mais c' était des § *c'était en rentrant* un soir il faisait un peu nuit et puis + mais c' est [Source CFPP]

- 60 Syntactiquement ces emplois gérondivaux sont régis : *c'est où ? en allant... ; c'était quand ? en rentrant...*

6.5. Participes et gérondifs composés

- 61 Tout comme le participe présent simple (du type *Mangeant peu, il ne grossissait pas* (cf. §1), le participe présent composé (27-31) peut exprimer une prédication secondaire à valeur causale, à laquelle s'ajoute l'expression de l'antériorité du procès second sur le procès principal, à la voix active ou passive (28, 29) :

(27)

unine08-jla et puis ben le village *ayant perdu sa poste* [Source OFROM]

(28)

407PERS : le français *ayant été passé* en première

HM : hm hm

407PERS : donc c'est déjà quand même spécialisé [Source ESLO]

(29)

unine08-aba : et *étant guillotiné ayant été guillotiné* assez rapidement [Source OFROM]

(30)

421PERS : je m'excuse en revenant souvent à au Mans parce que

421PERS : *ayant passé* vingt ans dans l'usine je connais mieux que celle-ci [Source ESLO]

(31)

unine11-eja *étant né* au Portugal j'ai appris le portugais euh [Source OFROM]

- 62 Du point de vue aspectuel, le « participe composé » (*ayant perdu, ayant été passé*) marque l'accompli et cumule la valeur sécante de l'auxiliaire à celle, résultative, du participe passé, maintenant une valeur sécante à l'ensemble (cf. Arnavielle, 2003 : 40), tandis que « le rôle de *en* est d'apporter une surcharge aspectuelle qui rend malaisée l'expression [...] de l'accompli » (*Ibidem* : 44). De sorte que, en toute logique et comme d'aucuns l'ont affirmé, l'usage du gérondif composé serait très rare en français standard où l'antériorité est plutôt exprimée à l'aide du participe composé (Halmøy 2003).
- 63 Nos données confirment ces affirmations. Cependant, l'oral spontané n'est pas exempt de gérondifs composés :

(32)

mais j' crois que je suis comme j'ai jamais fait quasi d' vélo d' ma vie euh en tant que vraie parisienne et *en étant très peu partie* à la campagne j' suis pas + pas très à l'aise mais pour faire les petites distances [Source CFPP]

(33)

? § jusqu' au bacca- § -lauréat et puis § *en ayant* quand même *redoublé* deux classes justement dû aussi [Source CFPP]

- 64 On aurait donc ici affaire à des emplois où le conflit aspectuel présumé, dû à la rencontre de *en* et de l'aspect accompli du participe composé, semble s'effacer à la faveur de l'expression d'une prédication seconde exprimant la cause (i.e. *comme / étant donné que j'étais très peu partie...*) (cf. infra). On pourrait alors faire l'hypothèse que, en vertu de l'économie de moyens syntaxiques offerte par le gérondif composé (parataxe vs hypotaxe et forme non finie vs conjuguée du verbe), la violation du conflit aspectuel supposé, loin de constituer un obstacle, permette d'enrichir le gérondif d'une nouvelle valeur sémantique : l'aspect accompli du procès antérieur (*étant parti, ayant redoublé*) qui est à l'origine d'un état qui dure (double aspect itératif et statif de 32 ; durativité de l'état résultant en 33) jusque dans la temporalité du verbe principal (par le jeu de *en*) confère à la construction une valeur causale qui se surajoute aux valeurs temporo-aspectuelles et à la signification englobante du gérondif.

6.6. Des frontières incertaines dans les emplois

- 65 Participe présent ou gérondif ? Les frontières de leurs territoires d'action ne sont pas aisées à tracer²⁸ et le tableau d'ensemble se complique encore dès lors que l'on considère les données du français parlé. Certes, certains emplois connus et décrits pour la langue écrite sont illustrés à l'oral, mais il en est d'autres qui révèlent des phénomènes spécifiques de la parole spontanée et qui se caractérisent souvent par leur résistance à une analyse syntaxique traditionnelle et univoque.

6.6.1. Gérondif en lieu et place du Pp^{ant} ?

- 66 Arrêtons-nous un instant sur l'exemple que nous avons cité en exergue de notre travail.

(0)

Ce livre est une espèce d'ouverture *en disant* eh bien ça peut être ça aussi la philosophie. [Michel Onfray, Questions du public, Conférence Université populaire de Caen, 29.07.2011]

- 67 Il n'appartient pas aux corpus analysés, mais représente un événement de parole spontanée puisqu'il apparaît dans une réponse aux questions adressées par les auditeurs au conférencier. Il se situe donc non plus dans l'écrit oralisé de la conférence mais bien dans la spontanéité d'une interaction verbale directe avec des interlocuteurs situés dans un cadre formel.
- 68 Cet usage du gérondif n'appartient pas à ceux que préconise la norme écrite française. En effet, la prédication seconde de cause au gérondif est généralement placée en tête de l'unité énonciative, comme dans l'exemple suivant :

(34)

aussi donc euh forcément euh *en ayant* des enfants et + *en voulant* essayer d'écourter les horaires de (mm mm) de enfin ... [Source CFPP]

- 69 Et c'est le participe présent et non le gérondif qui est habilité à exercer la fonction de modifieur du nom (*personnel ; mail*) :

(35)

417PERS : très curieux moi je j'avais pas connu mille neuf cent trente-six
417PERS : euh j'ai connu cette période ça a été très curieux on voyait arriver des des du personnel euh
417PERS : tout le personnel d'une entreprise ou euh *disant* on veut se syndiquer parce que ils savaient
417PERS : que dans la mesure où ils voulaient discuter avec leur patron
417PERS : y a que si l'accord était signé avec une organisation syndicale
417PERS : qu'il était valable [Source ESLO]

(36)

m' a renvoyé + *un mail* au bout de trois semaines *me disant* peut-être que j' vais l' retrouver j' vous l' enverrai [Source CFPP]

- 70 L'usage illustré en (0) peut donc s'analyser de deux façons : 1) comme un gérondif utilisé syntaxiquement en lieu et place un participe présent, dans la fonction de modifieur du nom *ouverture*, interprétable comme un prédicat de parole qui introduit le discours direct, comme le ferait la relative (Ce livre est une espèce d'ouverture *disant* eh bien ça peut être ça aussi la philosophie) ; 2) comme un gérondif de prédication seconde à valeur causale, qui, pour des raisons liées à la spontanéité de l'énonciation, aurait été produit a posteriori, et non dans la zone pré-verbale – son site syntaxique canonique – (*En disant* eh bien ça peut être ça aussi la philosophie ce livre est (constitué) une espèce d'ouverture). Selon cette deuxième interprétation, l'expression spontanée de la démonstration obéirait à un ordre chronologique inverse des propositions – thèse puis argument qui l'appuie – par rapport à l'ordre syntaxique de la norme écrite. Dans les deux cas d'analyse, la proposition (*ce livre est une ouverture*), qui occupe la première place, représente l'information saillante de l'énoncé.
- 71 Les corpus oraux offrent divers exemples d'attestations qui sembleraient compatibles avec la première analyse, où le gérondif se situe sous la portée du nom (cf. (37) *mentalité*), à l'instar du participe présent :

(37)

unine08-sma : cette mentalité *en disant* non il y a pas c'est parce qu'on est âgé qu'on est automatiquement souffrant [Source OFROM]

- 72 ou bien avec la deuxième analyse, où le gérondif employé comme prédication seconde de causalité se trouve en position post-verbale et non en tête d'énoncé. Il obéit ainsi à une économie de l'information propre à la parole en construction (hésitation puis explication a posteriori *euh en me trouvant*) et occupe une place que l'usage normé réserve plutôt au participe :

(38)

RF211 : non euh faut dire que les concerts à Orléans

RF211 : [rire :noise :previous]

RF211 : il m'est arrivé d'aller au concert par exemple à Bourges euh

RF211 : *en me trouvant* au festival du théâtre de Bourges la maison de la culture de Bourges qui est un

RF211 : disons un

RF211 : un outil merveilleux pour les les comédiens et pour les acteurs [Source ESLO]

- 73 L'exemple (0) est également intéressant en ce que, au-delà des aspects de l'analyse syntaxique, il illustre l'un des verbes les plus fréquents au gérondif à l'oral : *dire*. L'oral spontané recèle nombre d'emplois bien connus de *en disant* comme gérondif de manière (et/ou moyen et simultanéité) régi par le verbe principal (i.e. je terminerai [...] *en disant* ; c'est *en disant* [...] que je terminerai) :

(39)

417PERS : alors qu'avant euh c'était euh la porte tout de s- euh

417PERS : à à terme bien sûr puisqu'on met pas les gens dehors

417PERS : *en disant* vous avez fait grève on trouve toujours un autre prétexte
[Source ESLO]

(40)

FA192 : on s' était fait la réflexion justement à une équipe UFCS et
FA192 : *en disant* euh ben c' est intéressant on on on a créé quelque chose [Source
ESLO]

- 74 Les traces de la valence verbale (*en (leur) disant*, *en (se) disant*) ont disparu et *en disant* semble parfois devenir une sorte de construction fixe qui permet, à distance de plusieurs tours de parole, d'assurer une continuité prédicative avec le verbe principal (*ils disent [...] en disant*) :

(41)

OH593GRM : ils disent pas beau temps et
OH593GRM : oh il fait
OH593GRM : temps encore *en disant* jeune homme quoi
MD : oui
OH593GRM : ouais il en
MD : qu'est-ce qu'ils disent encore madame ? [Source ESLO]

- 75 Dans d'autres cas, le sémantisme du verbe principal s'oppose à l'interprétation temporelle (simultanéité) et impose une interprétation du gérondif comme procès consécutif (*et me disent ; et je lui ai demandé*) ou comme finalité (*pour me dire, pour lui demander*) par rapport au procès principal :

(42)

441PERS : ou alors les les hommes qui s'en sont occupés les sapeurs pompiers par exemple
441PERS : me té- me téléphonent tout de suite *en disant* en voilà ce qui se passe
441PERS : et à une heure une heure du matin faut que je sois levé pour remédier à la catastrophe [Source ESLO]

(43)

BA725 : et alors à ce moment-là je suis allé voir le le directeur de du lycée
BA725 : *en lui demandant* ce que euh mon aîné ce qu'il pourrait faire il m'a dit écoutez c'est un élément [Source ESLO]

- 76 Ces gérondifs ne sont pas syntaxiquement détachables et doivent être nécessairement post-posés. Ils fonctionnent comme des marqueurs de discours rapporté soit indirectement (43 ... *en demandant ce que...*) soit directement (42 ... *en disant voilà...*).
- 77 Enfin il est des cas extrêmes où le gérondif est utilisé, même si l'aspect lexical et grammatical du verbe principal (télique et accompli) ne permettrait pas l'usage du gérondif (inaccompli) selon la norme de l'écrit standard, qui lui préfère alors le participe présent, passé ou encore l'infinitif (*je me suis retrouvé constatant / à constater ; je me suis réveillé Pp^{ant}/Pp^e*) :

(44)

unine08-tsa : donc en fait euh je me suis retrouvé à la fin de l'année *en constatant* que [Source OFROM]

(45)

unine08-jga : ben j'ai voulu euh je me suis senti mal je me suis réveillé *en étant* complètement [Source OFROM]

- 78 Du point de vue syntaxique, ces gérondifs sont régis par le verbe principal (cf. *je me suis retrouvé/réveillé comme ça/ en constatant.../ en étant complètement ...*).

6.6.2. Participe présent en lieu et place du gérondif

- 79 En revanche les productions orales du français québécois et de Suisse romande présentent différentes attestations du Pp^{ant} là où l'on attendrait le gérondif. Ces cas sont nous semble-t-il suffisamment nombreux pour ne pas être considérés comme des épiphénomènes (telle l'ellipse occasionnelle du *en*) :

(46)

VE : [1vous êtes combien de journalistes/ (0 :07 :34.2) <P10,L7>

VI : quatre (0 :07 :35.1) <P10,L8>

VE : *comptant le rédacteur/* (.) <all<ou non/>> (0 :07 :36.9) <P10,L9>

VI : *comptant le rédacteur* (0 :07 :37.9) <P10,L10> [Source CFPQ]

(47)

unine08-yba : et euh ils s'arrangent avec elle ils euh *disant* voilà on paie euh [Source OFROM]

(48)

unine08-rza : *grandissant* il devient toujours plus autonome [Source OFROM]

- 80 Cette tendance s'observe aussi pour le gérondif (*en parlant de*) dont on a suggéré qu'il pourrait présenter un emploi en voie de grammaticalisation, lorsqu'il peut être glosé par la locution prépositionnelle à *propos de* (Floquet, 2012 : 2149). Il est bien présent (49, 50) mais la préposition *en* semble assez instable dans le français parlé québécois (51, 52) :

(49)

unine08-ada : eh *en parlant de* fast food ils ont des... [Source OFROM]

(50)

G : et pis euh les cours *en parlant de* cours aussi je sais pas à l'école s'il y a des cours de : de 1diététique ou n'importe quoi de la manière qu- qu'ils mangent ou euh : : :

≡<548445> [sous-corpus 1, segment 8, page 121, ligne 11] [Source CFPQ]

(51)

J : heille/ *parlant de ça* j'ai ben aimé le gâteau l'autre fois [sous-corpus 17, segment 8, page 106, ligne 11] [Source CFPQ]

(52)

SO : heille *parlant d'ados* 1là j'en ai vu deux là «<515996> [sous-corpus 7, segment 1, page 11, ligne 14] [Source CFPO]

- 81 De toute évidence, il s'agit là d'emplois de gérondifs et de participes non régis, qui ont pour fonction pragmatique d'assurer la continuité thématique de l'énonciation.

6.6.3. Participe et adjectif : la fonction d'attribut

- 82 On sait que la séparation entre adjectif verbal et participe présent semble très mince et artificiellement entretenue puisque, entre autres spécificités, les deux formes peuvent être coordonnées à l'écrit (Wilmet :2003 : 312). Cela est également vrai dans l'oral spontané illustré par notre corpus : l'expression de la valence verbale (*s'*) ne laisse aucun doute sur la nature verbale de la forme en *-ant*.²⁹

(53)

441PERS peut-être pas très imagitatif

441PERS mais positif sérieux

441PERS et *s'appliquant* surtout [Source ESLO]

- 83 On sait également que le participe peut assurer la fonction d'attribut du sujet, qui lui est d'ordinaire déniée.³⁰ Cela ne se fait pas sans contrainte syntaxique, puisque la présence d'un « médiateur »³¹ est alors nécessaire entre *être* et le participe :

(54)

...qui *sont* souvent *roulant* des mécaniques (Le Canard enchaîné, 2000 cité par Arnavielle, 2003)

- 84 Les productions orales semblent le confirmer :³²

(55)

unifr11-gaa : *c'est* principalement *dépendant* des profs il y a des mauvais profs chez Steiner il y a des bons profs chez Steiner [Source OFROM]

(56)

401PERS : mais on ne peut pas dire

401PERS : que des monuments

401PERS : ne devant plus *rester comme étant* des monuments d'Eglise [Source ESLO]

- 85 L'exemple précédent nous conduit à aborder la fonction d'attribut de l'objet, que revêt également le participe, en particulier de l'objet de verbes psychologiques qui régissent la préposition *comme*. Si le participe *étant* assure la prédication adjectivale sur l'objet (*l'*, *c'*) et semble en quelque sorte faire partie intégrante d'un syntagme prépositionnel introduit par *comme* :

(57)

pas enfin pff + disons que j' l' *ressens pas comme étant* traditionnel j' le j' l' vis pas comme ça [Source CFPP]

(58)

441PERS : et que c'est *considéré comme étant* de première qualité [Source ESLO]

- 86 d'autres lexèmes sont attestés dans la forme participiale suivie de sa complémentation en fonction d'attribut de l'objet après *comme* :

(59)

419PERS : euh l'Etat actuellement ne nous c- *ne considère pas* les fabricants de vinaigre *comme ani- assainissant* le marché du vin

419PERS : alors si bien que d'un aut-

419PERS : du fait qu'on n'est pas *considéré comme* euh419PERS : *pouvant assainir* le marché de vin euh on est

419PERS : on n'a pas l- on comment vous expliquer ça exactement c'est assez c'est assez confus parce que vous savez c'est très compliqué c'est le marché du vin où l'Etat s'en mêle [Source ESLO]

- 87 Avec l'introduction de la variation diatopique, les pistes se brouillent et obligent à prendre en considération les autres formes verbales en *-ant* ainsi que la question de leur recatégorisation syntaxique. Les exemples ci-après montrent qu'en français parlé québécois la forme *travaillant* fluctue dans la parole du locuteur R où elle passe du statut syntaxique de nom à celui d'adjectif et de participe, si l'on s'accorde à considérer comme tels les cas où s'intercale un adverbe « médiateur » (TROP, BEN). Notons que la présence d'un adverbe ne semble pas indispensable. La construction qui en résulte n'est pas sans rappeler la structure du présent progressif de l'anglais.

(60)

R : ayoyoye pauvre Joël je l'aimais ben Joël *il était un bon travaillant* à part de çaH : Joël oui TROP *il est TROP travaillant* là ¹t'sais/ [sous-corpus 4, segment 9, page 105, lignes 18 et 19] [Source CFPQ]

(61)

R : ¹ah *il était BEN travaillant* de la manière ²que je l'ai connu pour euh quand il travaillait avec nous-autres ³dans la napping c'était numéro un là/ ⁴c'était un bon employé ça

[...]

A : ⁴pis com- comme boss faut tu fasses plus attention parce que tu peux *même si t'es travaillant* tu peux pas pousser les employés comme tu veux [sous-corpus 4, segment 9, page 106, lignes 1 et 4] [Source CFPQ]

- 88 Enfin, cette dernière attestation (62) illustre bien, à notre sens, comment le participe présent – mode supposé compatible avec tout type de verbe qui ne soit pas impersonnel – peut s'avérer une construction utile au locuteur qui peut assurer ainsi une continuité au flux de la parole, tout en évitant de rompre la structure syntaxique qu'il a ébauchée (i.e. *tout ce qu'il y a de [...] tout ce qu'il y a d'existant / tout ce qui existe*) :³³

(62)

431PERS : euh j'ai j'ai été à New York et je crois avoir vu en meubles français à peu près *tout ce qu'il y a de*431PERS : euh *tout ce qu'il y a d'existant* et je peux dire que je n'en ai pas trouvé en

dehors de ceux qui sont chez quelques antiquaires je n'ai pas trouvé un meuble français correct [Source ESLO]

7. En guise de conclusion provisoire

- 89 Confrontée à l'épreuve des faits, l'intuition de certains linguistes, selon laquelle le gérondif serait très vivace à l'oral et à l'écrit, est donc démentie ou tout au moins à nuancer.
- 90 Cette intuition pourrait cependant s'expliquer par deux phénomènes que cette étude empirique confirme et précise, à savoir, 1) la disponibilité plus grande du mode gérondif en regard du mode participe dans la parole orale spontanée 2) la récurrence d'un nombre relativement restreint de constructions incluant le gérondif, voire de certaines classes lexicales au gérondif, qui pourrait donner l'illusion d'une forte présence de ce mode verbal, à l'oral.
- 91 Pour ce qui est de la répartition comparée des deux modes dans les corpus oraux et écrit, le gérondif joue un rôle prééminent par rapport au participe dans la parole spontanée : lorsqu'il résulte d'une construction libre, le gérondif est 3 fois plus fréquent que le participe à l'oral et cette répartition est inversée à l'écrit. Dans la parole spontanée illustrée par les corpus, le gérondif est également plus souvent une forme libre (91 % des occurrences) que ne l'est le participe présent (60 % des occurrences). Dans l'écrit de presse, en revanche, le participe présent est plus fréquent et correspond dans 99 % des cas à une forme libre.
- 92 Quant aux emplois gérondivaux des corpus oraux, on ne peut dissocier la disponibilité du mode verbal de la nature lexicale des verbes actualisés. On remarque alors deux types de constructions spécifiques de l'oral spontané qui semblent fortement corrélés à des fonctions discursives bien précises : 1) *en disant* marque le changement de niveau énonciatif en introduisant un discours rapporté (d'un tiers humain ou pas, ou du locuteur monologuant) et tend à substituer le participe comme modifieur du nom ; 2) les gérondifs de verbes de mouvement, très fréquents, qui, au-delà de la « construction obligée » exprimant la manière de mouvement, entrent dans des constructions du type *X est là en entrant ; il y a Y en descendant* et expriment la localisation spatiale sur une trajectoire représentée. En instaurant un repère dynamique virtuel – dont l'agent implicite est un humain engagé dans un mouvement orienté, qu'il soit identifié pragmatiquement ou simplement indéfini – le gérondif de mouvement a vraisemblablement pour fonction de faciliter, dans le cadre d'une interaction, l'opération discursive de localisation spatiale d'un objet, pour le locuteur, et son interprétation, pour l'interlocuteur. En d'autres termes, il ancre en quelque sorte la description verbale dans l'expérience cognitive spatiale de l'interlocuteur afin de la rendre plus concrète donc plus déchiffrable. Ces catégories de constructions gérondivales constituent des motifs syntactico-sémantiques associés aux genres discursifs descriptif et narratif généralement bien représentés dans l'oral spontané.
- 93 Du point de vue syntaxique, les emplois de gérondifs de l'oral, sont détachés de toute contrainte concernant le sujet zéro. L'agent du procès gérondival peut être explicite (par coréférence au sujet du verbe principal) ou implicite (et exprimé hors du cadre de l'unité énonciative) ou encore non exprimé linguistiquement mais interprétable pragmatiquement. Quant aux autres arguments du verbe, ils sont rarement exprimés

au sein de la structure gérondivale : les formes de gérondifs cliticisés en effet sont en nombre négligeable à l'oral. De plus, elles sont parfois accompagnées d'une reprise redondante de l'argument en syntagme post-verbal, et elles semblent constituer un lieu potentiel d'interruption du flux de la parole. Une exploration systématique et approfondie de ces phénomènes pourrait probablement éclairer le degré de complexité (en production et en compréhension) que représente cette structure dans la parole spontanée et la manière dont les locuteurs gèrent cet obstacle éventuel.

- 94 L'oral spontané offre des attestations de gérondifs (simples ou composés) de prédication seconde à valeur causale que la norme de la langue écrite réserverait plutôt au participe. Il semble qu'une des raisons du caractère étrange de ces constructions soit à rechercher dans un conflit de position qui dérive de l'ordre syntaxique et de la position que la parole en train de se construire décerne aux unités d'information. Il n'est pas rare en effet à l'oral que le gérondif en construction absolue à valeur causale apparaisse hors de la zone pré-verbale – son site syntaxique canonique – mais bien dans la zone post-verbale souvent réservée au participe présent. Il nous semble par conséquent nécessaire, afin de mieux comprendre ces emplois de l'oral, d'intégrer dans une analyse systématique les différents niveaux structurels de la description linguistique.
- 95 Enfin, en français parlé du Québec et de Suisse romande, au contraire, le participe tend à occuper des emplois de gérondifs et présente des cas où les frontières entre les usages des formes prédicatives en *-ant* semblent encore plus perméables. Les raisons de cette tendance inverse restent à élucider.
- 96 Pour tenter d'expliquer une telle opposition dans les usages de l'oral spontané et de l'écrit élaboré de la presse, nous évoquerons à nouveau ici l'idée que chacune des expressions considérées – la parole composée du registre formel journalistique distincte de la parole immédiate et en devenir du registre moins formel de conversations privées –, est régie par une grammaire distincte.³⁴ Selon cette perspective, le gérondif et le participe présent seraient des formes dont l'usage serait rare dans la grammaire du français parlé pour se limiter à quelques verbes fréquents, à quelques motifs sémantico-syntaxiques particulièrement fonctionnels selon les genres discursifs de l'oral spontané, et susceptibles de donner lieu à des gérondifs et participes présents appelés, selon notre hypothèse, à un figement probable.

BIBLIOGRAPHIE

- Avanzi, M. & Béguelin, M.-J. (2012). Présentation du corpus OFROM – corpus oral de français de Suisse romande. Université de Neuchâtel, <http://www.unine.ch/ofrom>.
- Arnavielle, T. (2003). Le participe, les formes en *-ant* : positions et propositions. *Langages*, 149 : 37-54.
- Arnavielle, T. (2010). Le gérondif : nouvelle définition d'un objet étrange. *Cahiers AFLS*, 16(1) : 6-24, <http://www.afls.net/cahiersrevue.php>.

- Barra-Jover, M. (2004). Interrogatives, négatives et évolution des traits formels du verbe en français parlé. *Langue française*, 141 : 110-125.
- Bergougnieux, Gabriel. (1996). Etude Socio-Linguistique sur Orléans (1966-1970), *Revue Française de Linguistique Appliquée*, I-2, 87-88.
- Blanche-Benveniste (1990). Grammaire première et grammaire seconde : l'exemple de EN, *Recherches sur le français parlé*, 10, 51-73.
- Bolasco, S. (2010). *TalTac2.10, Sviluppi, esperienze ed elementi essenziali di analisi automatica dei testi*. LED
- Branca-Rosoff S., Fleury S., Lefeuvre F., Pires M., 2012, Discours sur la ville. Présentation du Corpus de Français Parlé Parisien des années 2000 (CFPP2000), <http://cfpp2000.univ-paris3.fr/CFPP2000.pdf>.
- Combettes, B. (2003). L'évolution de la forme en -ant : aspects syntaxiques et textuels. *Langages*, 149 : 6-24.
- Escoubas-Benveniste, M.-P., Floquet O., & Bolasco S. (2012) Contribution empirique à l'étude du gérondif et du participe présent en français parlé et écrit, in *JADT 2012 : 11èmes Journées internationales d'Analyse statistique des Données Textuelles*, Anne Dister, Dominique Longrée, Gérald Purnelle (eds) Actes du colloque JADT 2012, Liège , 13-15 juin 2012, pp. 473-485, <http://lexicometrica.univ-paris3.fr/jadt/jadt2012/tocJADT2012.htm>.
- Floquet, O. (2011). Notes sur la traduction du « gerundio », *Plaisance, Sezione lingua*, 2013-228.
- Floquet, O., M.-P. Escoubas-Benveniste, S. Bolasco (2012) Sur le gérondif dans le français parlé et écrit, in *Actes du IIIe Congrès Mondial de Linguistique française*, éd. par F. Neveu, V. Muni Toke, P. Blumenthal, Th. Klingler, P. Ligas, S. Prévost et S. Teston-Bonnard, EDP Sciences, 2012, pp. 2142-2154, <http://dx.doi.org/10.1051/shsconf/20120100247>.
- Floquet, O., (2013, à paraître). Participe présent et gérondif alternant dans un même contexte : une enquête par questionnaire, in A.M. Scaiola (a cura di), *Un'idea di Francia. Scritti per Gianfranco Rubino*, Roma: Vecchiarelli Editore, pp. 301-308
- Gettrup, H. (1977). Le gérondif, le participe présent et la notion de repère temporel. *Revue Romane*, 12: 211-270.
- Giordano R. and Voghera M. (2002). Verb system and verb usage in spoken and written Italian. In Morin A. and Sébillot P., editors, *JADT 2002 : 6èmes Journées internationales d'Analyse statistique des Données Textuelles*. Irista-Inria, pp. 289-300, <http://lexicometrica.univ-paris3.fr/jadt/jadt2002/tocJADT2002.htm>.
- Gougenheim G, Dictionnaire fondamental de la langue française, Didier, Paris, 1958. – (3 500 mots). Nouvelle éd. revue et augmentée, Didier, Paris, 1977 (ISBN 2-208-00133-8)
- Halmøy, O. (2003). *Le gérondif en français*, Paris-Gap : Ophrys.
- Herlsund, M. (2000). Le participe présent comme co-verbe. *Langue française*, 127 : 86-94.
- Laks, B. (2011) La phonologie du français et les corpus. *Langue française*. 169, 3-17.
- Massot, B. (2010). Le patron diglossique de variation grammaticale en français. *Langue française*, 168 : 87-106.
- Meillet A. (1912). L'évolution des Formes Grammaticales. *Scientia*, 12, 384-400, réimprimé in Meillet, A., editor (1948), *Linguistique historique et linguistique générale*, 1, Champion, pp. 130-149.

Reichler-Béguelin M.-J. (1995) Les problèmes d'emploi du gérondif et des participiales en français contemporain, in K. Zaleska et A. Cataldi, édés, *Le Français Langue Etrangère à l'Université : théorie et pratique*, Actes du Colloque de Varsovie, 25-26 novembre 1993, Varsovie, Uniwersytet Warszawski, Instytut Romanistyki, 1995, p. 243-260.

Rossi-Gensane (2011). Oralité, syntaxe, discours, In Detey S., Durand J., Laks B., Lyche Ch. *Les variétés du français dans l'espace francophone : ressources pour l'enseignement*. Ophrys, pp. 83-106

Stosic, D. (2012), En passant par : une expression en voie de grammaticalisation ?, CORELA / Numéros thématiques / Langue, espace, cognition, <http://corela.edel.univ-poitiers.fr/index.php?id=2844>.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical form. In Shopen T., editor, *Language Typology and Syntactic Description*. Cambridge University Press, vol. 3: 57-149.

Wilmet (2003). *Grammaire critique du français*, 3^{ème} édition, Editions Duculot, Bruxelles.

Zanola, M.T. (1999). La diachronie des constructions participiales dans une perspective sémantico-textuelle, *L'analisi linguistica e letteraria*. 2 : 333-384.

Zribi-Hertz, A. (2011). Pour un modèle diglossique de description du français : quelques implications théoriques, didactiques et méthodologiques, *Journal of French Language Studies*, 21(2) : 231-256.

NOTES

1. Nous tenons à exprimer nos plus vifs remerciements à Oreste Floquet et Frédéric Sabio pour leur relecture attentive et leurs précieuses suggestions.
2. Pour une approche comparative de l'italien et du français en liaison à des questions de traduction voir (Floquet, 2011).
3. Italiques et caractères gras de l'auteur.
4. Les protocoles de transcription des corpus oraux utilisés prévoient une transcription orthographique standard sans normalisation des éléments morphémiques (i.e. *y a pas* vs *il n'y a pas*) de l'oral. Chaque transcription adopte son propre codage des phénomènes suprasegmentaux et nous avons maintenu ces codes dans les citations. Le CFPQ est le seul des corpus oraux à avoir incorporé à la transcription un codage des gestes accompagnant la parole, par didascalie. Précisons enfin que, pour certains exemples que nous citons, une extension du contexte, coupé un peu abruptement, serait utile à l'analyse. Il n'a malheureusement pas toujours été techniquement possible, dans le cas des corpus directement consultés en ligne, d'accéder au contexte d'usage élargi.
5. Nous tenons à remercier vivement M. Olivier Baude pour l'aide précieuse qu'il nous a apportée dans le repérage des ressources illustrant le français parlé qui sont disponibles pour la recherche au moment où nous écrivons.
6. Le corpus ESLO se compose de deux volets (ESLO1 et ESLO2), la seconde partie de l'enquête initiée dans les années 2000 visant à poursuivre et amplifier le projet initial. Les matériaux (enregistrements et transcriptions écrites) ont été et sont encore mis en ligne par le laboratoire CORAL qui s'est donné pour objectif d'en permettre l'exploration au moyen d'outils de recherche (<http://eslo.tge-adonis.fr/>).
7. Au moment où nous avons effectué la recherche, la plateforme ESLO était en voie de finalisation. Il est par conséquent probable que les valeurs concernant le nombre de formes gérondivales et participiales présentes sur la base soient sous-estimées par rapport au volume

total déclaré du corpus. Cet aspect est à prendre en considération dans l'analyse quantitative, et tout particulièrement dans les tableaux 2 et 5.

8. Voir le texte « Présentation » accessible à l'adresse : <http://recherche.flsh.usherbrooke.ca/cfpq/> du site du corpus, hébergé par le Centre d'analyse et de traitement automatique du français québécois, Faculté des lettres et sciences humaines, Université de Sherbrooke (Québec), consulté le 10 juillet 2013.

9. Accessible à l'adresse suivante : <http://www11.unine.ch/>.

10. Nous remercions M. Jean-Marie Viprey d'avoir gracieusement mis à notre disposition ce précieux corpus de presse de 400 millions d'occurrences. Nous avons fait le choix méthodologique d'explorer une partie du corpus de presse après avoir vérifié la stabilité des fréquences des formes en *-ant* sur l'ensemble du corpus et sur chacun des sous-corpus constitués par les cinq quotidiens (cf. Floquet et al. 2012).

11. Ainsi, contrairement à ce que pourrait laisser entendre un usage un peu schématique des termes de « modalité orale/modalité écrite », il ne s'agit donc pas tant de confronter des usages selon le canal de production du message que selon le mode de planification de l'expression, en comparant donc les données issue d'une parole différée et formelle avec celle de la parole instantanée et située dans un cadre plus familial. La première est le résultat d'une expression élaborée, revue, remaniée et perfectionnée ; la seconde est une parole spontanée, c'est-à-dire immédiate et saisie dans tous les aspects sensibles de sa construction.

12. Dans cette phase du décompte nous qualifions de formes gérondivales et formes participiales toutes les formes en *-ant* qui ne sont pas lexicographiquement attestées comme des noms, des adverbess ou des adjectifs et toutes les formes grammaticalement ambiguës dont le contexte oral nous confirme qu'il s'agit de verbes et non pas de noms (i.e. *restaurant, montant*), d'adverbess (i.e. *maintenant, durant*), ni d'adjectifs (i.e. *épuisant, amusant*). Pour la méthode d'extraction concernant les données de l'écrit voir Escoubas-Benveniste *et al.* (2012 : 477).

13. Ce rapport était de 1 à 7 environ (volume du corpus d'oral 7 fois inférieur à celui du corpus d'écrit) dans notre précédente analyse (Escoubas-Benveniste *et al.*, 2012 : 478).

14. Voir les considérations de la note 7 concernant la probabilité d'une sous-estimation du nombre d'occurrences de formes gérondivales et participiales dans le corpus ESLO.

15. Précisons que 7 occurrences de PP^{ant}, soit près d'1/4, correspondent à des emplois de langue écrite récitée ou citée :

• poésie

560 - P : **•étant ainsi resté** loin des autres élèves TOUT un mois/ <len<à vivre au gré des rêves>> craintif\ comme j'entraiss\ devant le crucifix où sa lèvres se colle ma mère\ était en pleurs\ oh mes <f<ardents>> regrets depuis il fut toujours le premier à l'école° [sous-corpus 13, segment 10, page 145, ligne 3] (source CFPQ) ;

• définitions lexicographiques de mots

M : juif ils marquent [euh::: (en regardant le dictionnaire) <P26,L6>] [...]

576 - M : ça fait que c'est ça ils marquent •personne **appartenant** à la communauté ju- euh israéliite° [sous-corpus 12, segment 2, page 26, ligne 17] (source CFPQ).

16. Pour illustrer le raisonnement, prenons l'exemple du CFPQ : 54 classes lexématiques verbales au gérondif sont représentées par 115 occurrences (dont 30 hapax, c'est-à-dire en termes lexicaux, 30 classes « singletons », soit plus de la moitié des 54). Les 24 classes restantes actualisent 85 occurrences de gérondif, ce qui correspond à une fréquence moyenne de 3,5 occurrences par type lexical. Pour les participes, 9 hapax représentent 9 classes lexicales « singletons » sur 15, soit près des 2/3. Par conséquent 6 classes actualisent 18 occurrences avec une fréquence moyenne de 3 occurrences par classe.

17. Cette liste de formes en *-ant* (tous types grammaticaux confondus) rangées par ordre de fréquence décroissante (27376 > F > 10) a été établie par Jean Véronis, Professeur de Linguistique et Informatique à l'Université de Provence, à partir de Corpaix, corpus de français parlé d'un

million de mots (version mai 2000). Elle est consultable à l'adresse suivante <http://sites.univ-provence.fr/~veronis/data/freq-oral.txt>.

18. La liste retenue par l'auteur est la suivante : *en attendant, en passant, en allant (à sur vers), en partant de, en passant par, en admettant que, en attendant que/de, en supposant que*.

19. Dans son étude, Stosic (2013) analyse en diachronie et en synchronie l'évolution des emplois du gérondif *en passant par* et décrit le processus de grammaticalisation qu'a subi cette construction pour aboutir, en français contemporain à la coexistence d'emplois libres (a) et d'emplois où le syntagme fait office de locution prépositionnelle (b) : (a) *J'ai vu une boulangerie en passant par cette rue* (b) *Il a tout lu de Montaigne à Gary en passant par Gautier et Balzac*.

20. Voir pour cet aspect (Escoubas-Benveniste et al. 2012).

21. Les attestations de participes avec pronom clitiques et/ou négation sont rares à l'oral : 441PERS le secrétaire général n'a aucune responsabilité politique le maire seul *en ayant* [Source ESLO].

22. Nous n'avons pas tenu compte de l'insertion de « quoi » (*en quoi faisant*) dont nous n'avons rencontré aucune attestation dans les corpus oraux.

23. Ne sont comptabilisées ici que les négations complètes, avec *ne/n'* explicité au sein de la structure gérondivale.

24. Afin de comprendre si cette rareté des usages des séquences de plus d'un clitique est ou pas le fait du gérondif en tant que tel, il serait éclairant de confronter ce résultat avec la fréquence d'usage des séquences à plus d'un clitique des verbes conjugués, dans l'oral spontané.

25. Nous admettons l'existence d'un sujet syntaxique zéro, ou sujet \emptyset selon la terminologie utilisée par M.-J. Reichler-Béguelin (1995), en vertu du fait que l'usage du gérondif peut s'accompagner, comme dans l'exemple 10, de « traces » analysables comme les expressions indirectes d'un sujet syntaxique dans le cadre d'un verbe non fini. Les pronoms clitiques des gérondifs de verbes réfléchis ou réciproques en fournissent selon nous de nombreux exemples. Halmøy (2003 :112) observe en effet « [...] si le sujet n'est pas exprimé en tant que tel dans le SG, il peut cependant y figurer sous différentes formes, pronoms réfléchis et expansions diverses, qui sont autant de traces [gras de l'auteur] contribuant à la clarté de l'énoncé. ». Exemples de l'auteur : *Je lui prouverai en y allant moi-même. Sara était assise à côté de lui, et fumait, tout en regardant le fleuve, elle aussi. ... dit-il en se levant. (Ibid.)*.

26. Le contexte élargi, qui pourrait peut-être permettre de préciser l'analyse, n'est pas disponible au moment de la recherche.

27. En particulier dans les énoncés métadiscursifs que nous soulignons en italiques : *G : moi je croirais S : ben moi je dis que ça dépend <P53,L15> L : ah ton frère il est parfait (SOUPIR) (en s'adressant à Émilie) S : il est parfait ton frère (.) [1on l'aime ASSEZ (en s'adressant à Émilie) G : [1{Ø;ben} je croirais.*

28. Nous renvoyons à ce propos à une étude récente basée sur questionnaire (Floquet, 2013 à paraître) qui semble indiquer que, dans les contextes où gérondifs et participe présent peuvent alterner (*Reentrant chez lui/En reentrant chez lui, Paul a mangé une pomme*), les locuteurs natifs optent plus volontiers pour le gérondif. Invités à choisir une seule des deux formes possibles pour différents contextes, 37 répondants de langue maternelle française, dans la majorité des cas (¾ vs ¼), proposent le gérondif au détriment du participe, indépendamment de leur âge, sexe et origine géographique.

29. Nous savons en effet que les adjectifs issus de verbes pronominaux ont perdu la trace du pronom clitique (cf. *une chaise pliante (qui se plie), un parking payant (qui se paie)*).

30. (Wilmet : 2003, 313)

31. Voir (Arnavielle, 2003 : 41) dont nous reprenons l'exemple issu de la presse écrite.

32. Soulignons toutefois que, à la différence de 54, 56, dans lesquels *roulant* et *étant* sont morphologiquement invariables, 55 met en jeu une forme participiale susceptible de porter

l'accord (« c'est principalement dépendant des profs » vs « elle est dépendante des profs »). Cet indice morphosyntaxique semble indiquer une évolution en direction de l'adjectif et invite donc à faire une distinction entre ces cas, dans une analyse plus approfondie.

33. Précisons que l'emploi adjectival de « existant » confirmé par l'accord (existante/s) est attesté dès le XVII^{ème} siècle dans le Dictionnaire de l'Académie.

34. Cette observation a été théorisée par Claire Blanche-Benveniste (1990) qui propose pour en rendre compte les notions de « grammaire première » et de « grammaire seconde », cette dernière étant enseignée en tant que norme langagière par l'institution scolaire. Plus récemment, les travaux de Barra Jover (2004), Massot (2010) et Zribi-Hertz (2011) analysent le français parlé comme un système de plus en plus indépendant du français normé, véhiculé par l'école et attesté dans les écrits formels.

RÉSUMÉS

Notre recherche vise à approfondir la question, peu étudiée, de la présence du gérondif et du participe présent en français parlé et écrit. Troisième volet d'une analyse quantitative et qualitative effectuée sur les énoncés authentiques dans les deux modalités expressives, elle s'appuie sur un corpus de français écrit d'environ 6 800 000 mots et de français oral d'environ 3 200 000 mots. Par français écrit, nous entendons la langue de la presse française contemporaine ; par français parlé, les productions orales spontanées – entretiens libres et guidés, conversations familiales, conversations téléphoniques, interactions verbales variées – de locuteurs dont la langue maternelle (ou l'une des langues maternelles) est le français. L'analyse reconduite sur un échantillon d'oral spontané plus étendu confirme les observations de nos deux études empiriques antérieures concernant a) la rareté relative du gérondif en français parlé par rapport au français écrit et sa tendance à entrer dans des constructions absolues (Floquet *et al.* 2012) ; b) la rareté des emplois « libres » (i.e. non grammaticalisés) des modes verbaux non finis en *-ant* dans l'oral spontané par rapport à l'écrit de presse, ainsi que la disponibilité plus importante du gérondif à l'oral tandis que c'est le participe présent qui prédomine sur le gérondif dans le corpus journalistique (Escoubas-Benveniste *et al.* 2012). Une analyse qualitative fine des aspects lexicaux et morpho-syntaxiques des formes gérondivales et participiales libres ainsi que de leurs emplois à l'oral fait apparaître, à côté des emplois canoniques de l'écrit, d'autres cas qui montrent que les frontières entre les usages des deux modes par les locuteurs tendent à s'estomper.

The aim of the research is to better understand the use of the French *gérondif* (i.e. *en parlant*) and present participle (i.e. *parlant*) in spoken and written contemporary French. This is the third part of an earlier data-driven quantitative and qualitative analysis. It is based on a written corpus of about 6,800,000 words and an oral corpus of approximately 3,200,000 words. Written French refers to the language of contemporary French newspapers (i.e. national daily quality press), whereas spoken French is to be understood as spontaneous oral productions - free and guided interviews, family conversations, telephone conversations, various daily life verbal interactions – between speakers whose mother tongue (or one of their native languages) is French.

Our first objective is to verify the reproducibility of the results obtained in our previous empirical studies (Floquet *et al.* 2012, Escoubas-Benveniste *et al.* 2012) by further reducing the potential bias that might have been introduced by the difference in the size of the oral and the

written corpora analyzed so far. Our second goal is to enrich the description of linguistic facts by introducing a diatopic dimension of two varieties of French into the spoken data. Consequently the previous sample of spoken data was enlarged by adding one more corpus of the French spoken in France and two corpora of, respectively, Switzerland and Quebec spoken French language.

We applied the previously developed quantitative analysis (Escoubas-Benveniste *et al.* 2012), which combines computer assisted TaLTaC2 (Bolasco, 2010) and manual exploration performed on a written newspaper corpus (*L'Humanité*, 2010) and on two spoken corpora – corpus *Phonologie du français contemporain* (PFC); *Corpus du français parlé parisien* (CFPP) –, to the three additional speech corpora of spontaneous oral data – *Enquête socio-linguistique à Orléans* (ESLO), *Corpus oral de français de Suisse romande* (OFRM); *Corpus du français parlé au Québec* (CFPQ). A qualitative analysis followed, whose purpose was to describe lexical, morphological and syntactical aspects of both non-finite verbal uses from a differential point of view (written vs spoken corpora; *gérondif* vs present participle).

In order to examine the reproducibility of our earlier results on the status of the two verbal forms in *-ant*, we evaluate (a) the rarity of free (as opposed to lexicalized) verbal forms in *-ant* in the spontaneous spoken vs newspaper written French corpora; (b) the differential distribution of *gérondifs* and present participles in both types of data; (c) the presence of both verbal modes in spontaneous spoken French, taking into account the phenomena of grammaticalization.

The quantitative and qualitative analysis performed on the much larger sample of spoken French confirms our previous results as for (a) the relative rarity of *gérondifs* in spoken data as compared to journalistic written data and the tendency shown by these forms to be used as “*constructions absolues*” (Floquet *et al.* 2012) (i.e. often initial clause syntactically unbound constructions); (b) the rarity of “free” uses (i.e. non grammaticalized uses) of both non finite verbal forms in *-ant* in the oral as compared to the written data and also the greater frequency of the *gérondif* in the spoken corpora, whereas the present participle is much more frequent than the *gérondif* in the written one (Escoubas-Benveniste *et al.* 2012). The lexical and morpho-syntactical analysis of the free non finite verbal oral occurrences shows, apart from some canonical uses also present in the written corpus, other interesting cases indicating that the distinction between the uses of these non finite forms might be diminishing in the spoken language.

In quantitative terms and with respect to the comparative distribution of both verbal non-finite forms in the written and oral corpora, the data show that the *gérondif* plays a pre-eminent role in relation to the participle in spontaneous speech: when it is the result of a free construction, the *gérondif* is 3 times more common than the participle in oral data and this distribution is reversed in written French. In spontaneous speech, as illustrated by the oral corpora, the *gérondif* is also more often the result of a free construction (91% of occurrences) than the present participle (60% of occurrences). Written data, by contrast, show that present participles are more common than gerunds and correspond to a free form in 99% of cases.

From a qualitative standpoint, the greater frequency of the *gérondif* in oral data could well be correlated with the lexical nature of the verbs. In fact, the data show two frequent types of spontaneous oral specific constructions that seem strongly associated with specific discursive functions: 1) *en disant* that introduces reported speech/thought, marking therefore changes in enunciative levels; it tends to override the participle as a noun modifier; (2) the *gérondifs* of moving verbs that occur in two types of constructions: moving manner constrained constructions (i.e. *sortir en courant* vs *run out*) and spatial location constructions (i.e. *X est là en entrant*; *il y a Y en descendant*). These types of *gérondifs* are closely related to descriptive and narrative discourses that are generally well represented in spontaneous speech.

Syntactically, the spoken uses of the *gérondifs* are free of any constraint on their syntactic empty subject, whose referent (the agent of the process) may be either explicit or implicit (and expressed in some way outside the utterance frame) or even not expressed but pragmatically

identifiable. As for the other predicate arguments of the verb, they are rarely expressed within the *gérondif* structure: indeed, *gérondifs* with more than one cliticised argument are extremely rare.

The spontaneous speech data show uses of simple and compound *gérondifs*, that standard rules of written language would normally reserve to the present participle, such as second level causal predicates. One explanation for these uses might be found in a position conflict deriving from the canonical syntactic order for these constructions (in the preverbal zone) and the position they effectively receive during the real-time elaboration of spontaneous speech. Therefore a systematic analysis of the various structural levels is needed in order to better describe these phenomena.

Finally, in the varieties of French that are spoken in Quebec and in Switzerland, the present participle tends to incorporate the uses of the *gérondif* and shows particular cases in which it is not easy to discriminate its exact nature. The reasons for this reverse trend remain to be elucidated.

In order to account for such an opposition between the uses of *gérondif* and present participles in spoken and written samples of French, we will evoke the idea that each particular form of expression - journalistic formal written register as opposed to spoken informal register of private/family verbal interactions- is ruled by a distinct grammar. According to this view, the use of *gérondifs* and present participles would be rare in the grammar of spoken French and limited to a reduced number of lexical verb types and syntactic-semantic patterns that might undergo a form of lexicalization.

INDEX

Mots-clés : gérondif - participe présent - français parlé - français écrit - linguistique de corpus

Keywords : gérondif - French gerund - present participle - spoken French - written French - corpus driven analysis

AUTEUR

MARIE-PIERRE ESCOUBAS-BENVENISTE

MEMOTEF, Sapienza, Università di Roma
mariepierre.escoubas@uniroma1.it