

Sur le gérondif dans le français parlé et écrit¹

Floquet, Oreste^o, Escoubas-Benveniste, Marie-Pierre* & Bolasco, Sergio*

Sapienza, Università di Roma,

^oDSEAI, (Dipartimento di studi Europei, Americani e Interculturali), *MEMOTEF (Dipartimento di Metodi e Modelli per l'Economia, il Territorio e la Finanza)

oreste.floquet@uniroma1.it, mariepierre.escoubas@uniroma1.it, sergio.bolasco@uniroma1.it

1 Introduction

1.1 Le gérondif : de la forme à l'usage

Le plus souvent fondée sur une approche introspective – et donc sur les intuitions des linguistes – ou bien sur un petit échantillon d'exemples littéraires considérés comme paradigmatiques, la tradition des études sur le gérondif tourne essentiellement autour de trois axes : (a) la définition du gérondif par rapport au participe présent et à l'infinitif ; (b) le problème du sujet implicite et de son accord avec le sujet du verbe défini ; (c) les effets de sens virtuellement possibles². Une analyse fine des emplois concrets du gérondif dans les différents types de discours – aussi bien dans la modalité parlée qu'à l'écrit – reste à faire³. Nous voudrions présenter ici les résultats d'une première enquête sur les usages attestés du gérondif dans deux corpus homogènes du français parlé et du français écrit. Dans un premier temps, il s'agira de vérifier si l'opinion courante selon laquelle :

le gérondif est [...] 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)

est bien fondée. Sans vouloir remettre en cause le fait que :

structure incontournable, le gérondif se rencontre dans tous les types de texte et à tous les niveaux de langue : presse quotidienne et hebdomadaire, littérature, langues de spécialité (juridique, administrative, technique, modes d'emploi, recettes de cuisine, etc.). (Halmøy 2003 : 165)

Nous essaierons toutefois de montrer que sa distribution est si peu homogène qu'on a souvent l'impression que sa soi-disant «vivacité» ait pu être quelque peu surévaluée, notamment dans le discours oral spontané, où il semble occuper une position somme toute marginale. Une analyse quantitative nous permettra de fonder sur des données authentiques la présence statistique de cette forme non finie du verbe en fonction des situations discursives parlée et écrite. Dans un deuxième temps, nous passerons au peigne fin les attestations du gérondif dans les deux corpus dans le but de comprendre si les tendances d'emploi de ces formes verbales se distinguent d'un point de vue qualitatif. Une analyse linguistique des usages attestés nous permettra ainsi de faire apparaître certaines caractéristiques lexicales, morphosyntaxiques et sémantiques des formes qui y sont majoritairement employées selon le corpus considéré.

1.2 Le corpus Phonologie du français contemporain

Le projet international PFC (Phonologie du Français Contemporain) est l'un des plus importants corpus de français parlé qui sont actuellement disponibles pour mener des recherches phonologiques, morphosyntaxiques et discursives⁴. Son but explicite est 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)

Sa caractéristique principale, par rapport à d'autres corpus oraux (généralement de taille plus réduite) est que l'enquêteur :

équilibre au mieux les sexes, les âges et les catégories socioprofessionnelles. (Laks 2011: 11)

Ce qui signifie que :

PFC n'est donc pas un corpus sociologiquement construit pour être représentatif. Son extension à la francophonie et surtout le caractère coopératif et cumulatif du travail d'enquête qui voit un très grand nombre d'enquêteurs travailler en des points très différents interdisait de viser une telle représentativité par quota. PFC vise ainsi à rendre compte de la diversité des usages sans en pondérer le poids les uns par rapports aux autres. (Laks 2011 : 11, n.22)

Chaque enquête suit un protocole commun et est à son tour divisée en quatre parties : une discussion libre, une discussion guidée, une lecture de texte et de liste de mots (toujours les mêmes). Intéressés que nous sommes par l'incidence des syntagmes gérondifs dans le français tel qu'il est parlé spontanément, nous avons choisi d'analyser la discussion libre et la discussion guidée⁵. Par souci d'homogénéité avec la variété diatopique de français du corpus écrit (presse française) nous nous sommes limités aux enquêtes menées sur le territoire métropolitain (15 départements répartis sur tout le territoire, pour un total de 280 dialogues analysés ; la partie de PFC que nous avons analysée atteint 2,57 MB de texte).

1.3 Le corpus LH2010

Pour l'analyse du français écrit, nous avons choisi d'explorer un extrait du corpus de presse, mis à notre disposition par Jean-Marie Viprey qui regroupe les dix dernières années du *Monde*, du *Figaro*, de *Libération*, *l'Humanité* et de *La Croix*. Le corpus de presse est une ressource qui appartient au «Scriptorium numérique FANUM» de l'Université de Franche-Comté, sous la responsabilité de Jean-Marie Viprey. Afin de réduire la taille considérable du corpus de français écrit (2,2 GB correspondant à plus de 400 millions de mots) sans risque majeur d'altérer les résultats de l'exploration quantitative, nous avons effectué une première recherche visant à examiner la distribution des formes en /-ant/ sur l'ensemble des publications constitutives du corpus. Une exploration de chacun des quotidiens pour la même année de parution a montré que les formes en /-ant/ étaient quantitativement stables indépendamment du journal considéré. Cela nous a permis de formuler l'hypothèse que les journaux s'équivalent en ce qui concerne l'usage statistique du gérondif, qui représente une manifestation possible de forme verbale en /-ant/. Par conséquent, nous avons limité l'analyse à une année et à un journal et nous nous donnons la possibilité de compléter les mesures quantitatives par une analyse linguistique plus fine effectuée sur un corpus de concordances de taille raisonnable. Plus précisément, nous nous sommes limités à l'ensemble des articles de *L'Humanité* de l'année 2010 (ci-après LH2010). La taille du corpus LH2010 est de 40 MB et de 6 873 329 mots, ce qui demeure une dimension suffisante pour garantir la stabilité des fréquences.

1.4 Méthode d'extraction des formes du gérondif

Le repérage des données de l'analyse linguistique – formes verbales et contexte d'apparition – est le résultat d'une analyse semi-automatique assistée au moyen du logiciel TaLTaC2⁶ qui opère sur corpus non lemmatisés. Il permet, entre autres fonctions, l'extraction et la catégorisation de formes graphiques et de séquences, la lexicalisation des séquences extraites en tant que formes graphiques complexes (*multiword expressions*) ainsi que l'extraction des concordances. Le gérondif étant une "structure" discontinue qui autorise l'insertion d'éléments entre /en/ et la forme en /-ant/ du verbe, il faut avant tout dresser la liste des éléments susceptibles d'être insérés afin de pouvoir effectuer une requête complexe des segments bornés à gauche par /en/, à droite par /-ant/ et permettant l'insertion de 0, 1, 2, ou 3 éléments syntaxiques parmi les 18 recensés (clitiques et négation ainsi que leur forme élidée)⁷. Une

requête par expression régulière permet, après avoir catégorisé (selon TaLTaC2) les formes en /-ant/ (CATSEM(ant)) et les 18 candidats à l'insertion (CATSEM(insert)), d'extraire les locutions de ce type. La grammaire locale opérationnelle utilisée pour la recherche avancée est la suivante :

"en CATSEM(ant)" OR "en CATSEM(insert) CATSEM(ant)" OR "en CATSEM(insert) CATSEM(insert) CATSEM(ant)" OR
"en CATSEM(insert) CATSEM(insert) CATSEM(insert) CATSEM(ant)"

Un tel modèle permet de saisir les faux négatifs. Par la suite, il faut procéder à un nettoyage des données ainsi extraites pour éliminer les attestations non vraies (faux positifs), c'est-à-dire les suites syntaxiques qui ne sont pas des gérondifs (i.e. *en avant*, *en tant*, *en autant* etc.) ou des gérondifs grammaticalisés (*en admettant que* etc.). Finalement, il faut récupérer les vraies attestations qui pourraient être mal catégorisées en raison d'une imperfection dans la segmentation du corpus (i.e. *en enachetant un*). Après cette démarche, les gérondifs simples et avec insertions sont « lexicalisés » afin de permettre l'extraction de leurs concordances. Au moment où nous écrivons, la procédure de contrôle manuel a été en partie automatisée grâce à l'établissement de la liste extensive des lexèmes en /-ant/ qui ne sont pas des verbes. Restent cependant les formes ambiguës (p.e. *participant*, *restaurant*) à contrôler manuellement à partir des concordances. Observons que ces listes sont à concevoir et manier avec précaution compte tenu de l'évolution du respect de la norme orthographique dans l'écrit de presse (p.e. *fatiguant/fatigant*). Pour terminer, il faut rappeler que cette étude est une première ébauche d'un travail plus ample que nous sommes en train de mener ; par conséquent, nous n'avons volontairement pas opéré une distinction, du moins pour l'heure, entre le gérondif concessif introduit par *tout* et le gérondif simple, si bien que dans les statistiques que nous allons présenter les deux formes seront toujours confondues⁸.

2 Analyse comparée des aspects quantitatifs et qualitatifs dans les deux corpus

2.1 Aspects quantitatifs

Jusqu'à plus ample informé, il n'existe aucune recherche empirique concernant le gérondif. Les toutes récentes analyses sur les modes verbaux indéfinis de Rossi-Gensane (2011 : 88) sont très génériques et ne prennent en considération que le participe présent :

On remarquera enfin 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 représentés⁹.

Voici les premiers résultats de notre enquête sur les deux corpus. Un premier résultat concerne le nombre de formes gérondivales qui ont été repérées par rapport à toutes les autres formes. Dans PFC, on dénombre 182 syntagmes gérondifs sur un total d'environ 529 000 formes graphiques. Dans LH2010, on dénombre 8170 syntagmes gérondifs sur un total d'environ 6,8 millions de formes graphiques. N'ont pas été comptabilisées les formes qui ne sont en réalité que des configurations où le gérondif est désormais grammaticalisé dans un emploi quasi-figé de nature adverbiale (le type *en attendant*), prépositionnelle (le type *en allant à*) ou conjonctionnelle (le type *en supposant que*)¹⁰ :

(tableau 1)

	PFC	LH 2010
total types du corpus	18.806	148.989
total tokens du corpus	529.543	6.873.329
total tokens extraits	200	9657
total tokens de gérondif	182	9038
grammaticalisations	18	619
N gérondifs pour 10000 mots	3,4	13,1

On en infère aisément qu'une telle construction n'occupe pas une place centrale dans la morphologie du type d'oral que nous avons explorée (la langue parlée dans la discussion libre et guidée). En ce qui concerne le corpus de presse (LH 2010), en revanche, la présence du gérondif est tout autre. En données normalisées (rapportées à la taille respective des corpus), nous pouvons donc affirmer que, sans conteste, la présence du gérondif est plus importante à l'écrit, puisqu'on y rencontre environ 4 fois plus de gérondifs qu'à l'oral.

Par ailleurs, du point de vue de la richesse lexicale illustrée par les gérondifs, et exprimée en nombres de lexèmes verbaux distincts, les résultats de l'exploration sont assez prévisibles :

(tableau 2)

	PFC	LH2010
total classes lexématiques	72	1282
total tokens	182	9038

Les occurrences du gérondif sont l'expression de 70 classes distinctes de verbes dans le corpus PFC et de 1077 dans le corpus LM2010¹¹, dont respectivement 51% à l'oral et 45% à l'écrit sont réduites à des hapax. La richesse lexicale supérieure de l'écrit sur l'oral spontané est illustrée au niveau de l'emploi du gérondif, de même que les proportions de hapax rencontrées restent dans la norme compte tenu des volumes différentiels des corpus analysés¹². Ci-après le hapax du verbe /vouloir/:

(1)

50aev1_GG
EV : (rires) je suis partie au lycée **en voulant** euh tout recommencer.

2.2 Aspects morphosyntaxiques

Du point de vue morphosyntaxique, à l'exception des structures impersonnelles pour lesquelles on admet qu'elles ne possèdent pas de gérondif¹³, toutes les constructions sont représentées : transitives, intransitives (verbes inaccusatifs et inergatifs). Les trois exemples suivants, extraits du corpus PFC, en donnent respectivement une illustration :

(2)

13aid1_GG
ID1: Il me faut un quart, un quart d'heure à peu près à pied pour aller travailler donc euh, voilà ça me permet de, de marcher, de pas être stressée **en prenant** la voiture tous les matins, de trouver une place, pour se garer dans le centre ville

(3)

13aag1_LG
AG1: Et puis **en arrivant** à la maison alors il a essayé de faire ce que Pierre lui avait dit

(4)

11anb1_GG
NB: Moi j'aurais cru qu'il serait resté parce que pffft, il y a la mer euh c'était, bon, mais tu as les constructions. Oh peut-être oui oh, c'est un sauvage lui non, je sais pas, oui peut-être **en vieillissant** je crois (rires). C'est un sauvage mon frère il est très connu ici parce que quand il était jeune il, c'est un peu le Tarzan hein, il partait euh, de bonne heure le matin il faisait euh, La Grâce-Carcassonne à pied par les, par l'Alaric.

Une autre tendance qui caractérise du point de vue morphosyntaxique le gérondif à l'oral concerne le nombre d'éléments pouvant s'intercaler entre le morphème /en/ et la forme en /-ant/. Dans un premier temps, nous avons répertorié les différentes formes de gérondifs en termes quantitatifs d'éléments insérés (insert), sans distinction de leur fonction syntaxique : *en, la, le, l', les, leur, lui, me, m', ne, n', te, t', se, s', nous, vous, y*. Font partie de la catégorie 0-insert les gérondifs ne présentant pas d'insertion (le type *en donnant*) ; font partie de la catégorie 1-insert les gérondifs du type *en lui donnant, en ne donnant pas, en se mariant, en y mettant, en en mettant* ; et de la catégorie 2-insert : *en le lui donnant, en ne lui donnant pas, en ne se mariant pas, en n'y mettant pas, en n'en mettant pas*. Dans un deuxième temps, nous avons catégorisé ces formes en fonction de leur typologie structurelle, c'est-à-dire selon le statut grammatical (ou lexical) de ces insertions. Nous avons ainsi repéré les gérondifs de verbes pronominaux dont le pronom fait partie de la structure lexicale du verbe¹⁴. Ainsi parmi les gérondifs à insert (13 occurrences pour PFC et 578 occurrences pour LH2010) nous avons comptabilisé les pronominaux non ambigus (p.e. *en se mariant*), et les non pronominaux avec complément cliticisé (p.e. *en lui parlant*). Le tableau 3 illustre ces résultats ainsi que la densité de chaque catégorie selon les corpus et le nombre d'occurrences de gérondifs négatifs.

(tableau 3)

	PFC	% PFC	LH2010	%LH2010
total tokens de gérondif	182	100%	9038	100%
gérondif 0-insert	169	93%	8445	93%
gérondif 1-insert	13	7%	578	6,8%
gérondif 2-insert	0	0%	15	0,2%
gérondif 3-insert*	-	-	-	-
	PFC	% de x-insert	LH2010	% de x-insert
gérondif pronominal non ambigu	5	38%	216	36%
gérondif non pronominal à complément clitique	8	62%	377	64%
négation : en + ne+ (insert) + V-ant + pas	0	0%	53	9%

* *jamais actualisé dans nos corpus.*

À l'oral comme à l'écrit, nous pouvons voir que la forme sans insertion est la plus fréquente. Bien que le syntagme gérondif admette que l'on intercale jusqu'à trois éléments (le type *en ne le lui donnant pas*), dans les faits, les attestations de plus d'un clitique sont inexistantes à l'oral et tout à fait négligeables à l'écrit. La négation aussi, demeure un phénomène marginal quoique non négligeable à l'écrit¹⁵. Quant aux verbes pronominaux, leur proportion par rapport au nombre total de gérondifs extraits varie de manière significative selon les variétés de discours. Ils semblent nettement plus fréquents à l'écrit. Cette observation nécessitera d'étendre l'analyse aux formes ambiguës afin de pouvoir tirer des conclusions complètes¹⁶.

2.3 Gérondif et participe présent

Pour terminer par une observation sur la valeur des emplois du gérondif à l'oral, nous signalerons qu'une partie des occurrences du corpus PFC semblent avoir un comportement non canonique¹⁷. Il y a quelques gérondifs qui sont syntaxiquement autonomes et dont le fonctionnement n'est pas sans rappeler celui d'une construction absolue (ce qui est d'ordinaire l'apanage du participe présent¹⁸). Il s'agit souvent de

verbes d'état qui ont une présence très limitée dans le corpus. Dans (5), le gérondif *en étant* pourrait, par exemple, être glosé par une tournure causale (p.e. *puisque je suis Président...*) :

(5)

42arf1_LG

E2: Et alors maintenant, la retraite?

R: Ah, ben la retraite, c'est bon. Ca a été très dur quand même, ça a été très dur hein, j'ai eu six mois qui ont été très durs. Mais bon maintenant ça va, mais je pense toujours hein, j'y pense toujours, puis bon ben maintenant. Donc il y a la pétanque qui me prend beaucoup hein. **En étant** président de la société, donc ça fait quand même des années.

Au-delà de la ponctuation, qui demeure un choix subjectif du transcripteur¹⁹ et restant tout à fait conscients que l'analyse syntaxique de l'oral spontané requiert une méthodologie spécifique qui rende compte de phénomènes non linéaires propres à l'oral, il nous semble que la phrase introduite par le gérondif est le prédicat d'une construction absolue, détachée, par conséquent, tantôt de la phrase qui la précède tantôt de la phrase qui la suit. On a donc affaire à deux événements contigus plutôt qu'à deux actions simultanées²⁰.

(6)

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.

Ici encore, le gérondif *en ayant* introduit un événement qui n'est pas parfaitement concomitant avec ce qui précède, mais en représente plutôt un présupposé : le fait d'avoir vécu trois ans ailleurs peut servir d'explication aux difficultés actuelles. Exemple encore plus intéressant, dans (7), nous avons aussi bien un premier gérondif de coordination (*en passant*) - qui ne semble ni contemporain ni précédent à ce qui précède (ce qui contrevient à la règle selon laquelle l'action exprimée par le gérondif ne peut pas être postérieure à l'action exprimée par le verbe défini²¹ - qu'un deuxième gérondif (*en étant*) qui tolère soit une interprétation absolue, soit une interprétation « sémantiquement emboîtée » (on aurait alors affaire à un gérondif de gérondif, voire encore l'emboîtement d'un gérondif de type méta-discursif, qui viendrait expliquer en contexte la métaphore (*en passant...barrière*) :

(7)

13adp1_LG

PD1: Parce que j'ai découvert après soixante huit euh l'enseignement dans des conditions difficiles. Parce que euh il y a/ il y a eu euh, c'était par rapport à ce que j'ai connu moi quand j'étais euh, élève, et **en passant** de l'autre côté de la barrière **en étant** surveillant, conditions très difficiles où le pouvoir était pris par les parents d'élèves, pour pas dire par les élèves souvent.

Un cas, peut-être assez clair, de gérondif de gérondif est (8) :

(8)

21ash1_LG

SH: Puis avec les officiers supérieurs et puis, et puis le, le peuple citronné pour faire des sous pour les, les supérieurs, oui enfin quoi, il a vu beaucoup de choses. Si bien qu'**en voyant** toutes ces choses **en discutant**, eh ben, du reste **en discutant** avec mon grand-père Sillon quoi, eh ben, il est devenu communiste.

où le verbe /discuter/ semble dépendre du verbe /voir/. Tout semble se passer comme si la différence entre participe présent et gérondif tendait à s'estomper, une tournure telle que *voyant toutes ces choses en*

discutant étant, à nos yeux, plus proche de la grammaire normative. Même phénomène dans (9), où se pose un problème de temporalité. Le verbe /acheter/ est à interpréter comme antérieur au verbe /apprendre/ ²². Cependant, ce qui est remarquable est que /acheter/ soit un présupposé de /apprendre/, la relation temporelle étant pratiquement inexistante si ce n'est par le fait qu'une cause est forcément précédente à un effet :

(9)

12atp1_GG

TP: Mais il parlait très bien le russe parce qu'il avait appris lui-même. **En achetant** un livre et. Mon papa était très, il aimait apprendre.

Encore une fois, nous constatons un usage du gérondif dont la fonction est proche de celle d'un participe présent ou d'une subordonnée causale. À l'inverse, on peut aussi signaler des cas où le participe sans complément semble fonctionner comme un gérondif :

(10)

69ajd1_GG

E: Ouais, parce que je me suis paumé un peu en euh, **arrivant**, puis euh, je me/ je , je suis tombé sur la Tête d'Or, bon euh.

Il semble donc qu'à l'oral ait lieu comme une sorte d'effacement des frontières d'emploi entre les formes en verbe + /-ant/, ce qui nous conduit à penser qu'il convient que nous élargissions la recherche en incluant à l'analyse comparative celle des emplois du participe présent oral/écrit.

L'exemple (11), en revanche, pourrait témoigner d'une grammaticalisation en cours. Le gérondif *en parlant* peut être en effet glosé par une locution prépositionnelle du type *à propos de*²³ :

(11)

44ajn1_LG

E : Ah **en parlant** de BD, elle était très bien la BD de, ouais, c' est super sympa.

Finalement, nous signalons aussi que le gérondif à valeur disjonctive (ou concessive), le type *tout en faisant*, peut ne pas être introduit par /tout/:

(12)

13aog1_LG

E1: tu/ penses avoir suffisamment de congés ou pas?

OG1: Ben en fait j'ai une semaine à poser pour l'hiver **en sachant** qu'ils considèrent que l'hiver ça va du novem/ de début novembre jusqu'au trente décembre, euh trente décembre moi, jusqu'au trente juin, c'est la semaine d'hiver.

Dans ce cas-ci, le gérondif pourrait également être remplacé par un participe présent. Concernant les aspects quantitatifs, somme toute modestes, du gérondif de concession, rappelons brièvement que le type *tout* + gérondif est présent 15 fois dans PFC et 372 fois dans LH2010.

3 Conclusions

Nous voudrions conclure avec des considérations portant sur la rareté vs abondance du gérondif dans les deux types de discours examinés. Est-il possible de fournir une explication au fait que le gérondif soit si peu présent à l'oral ? Sans oublier que nous avons affaire à un corpus oral qui est d'une taille bien inférieure au corpus écrit et qui en outre n'est pas forcément représentatif de l'ensemble des

manifestations orales du français, force est d'admettre l'absence presque totale du recours au gérondif par les locuteurs. À nos yeux, une telle situation mérite une réponse différente de celle proposée par Rossi-Gensane (2011 : 100) concernant toutes les formes verbales indéfinies :

A l'oral, et contrairement à ce qui se produit à l'écrit [...] la succession linguistique (dans les passages narratifs) correspond dans une très grande mesure au déroulement chronologique des événements, qui sont rapportés de manière linéaire – caractéristique sans doute motivée par des contraintes cognitives. [...] Cette « iconicité » repose notamment sur un abondant recours à des verbes conjugués (c'est-à-dire à des modes tels que l'indicatif [...], le subjonctif et l'impératif) et, partant, sur un faible recours à des verbes non conjugués (participes dans des appositions), dits encore participes détachés, gérondifs et infinitifs dans des compléments circonstanciels).

D'une part, nous ne sommes pas tout à fait convaincus que le gérondif soit en lui-même moins iconique qu'un subjonctif, par exemple, ou n'importe quel autre mode défini. En effet, du moins sur le plan temporel, le gérondif exprime de manière non équivoque et succincte la simultanéité et cela pourrait bien être un atout plutôt qu'un handicap dans la narration. D'autre part, il nous semble que la différence d'usage soit à rechercher dans l'instabilité sémantique intrinsèque aux effets de sens supplémentaires qu'il peut véhiculer. Comme le dit clairement Zanola (1998 : 258) à propos de constructions participiales en général :

L'un des points qui doivent être approfondis est la possibilité de la pluralité de sens coexistant dans ces constructions, le recours à ces formes engendrant par conséquent la possibilité d'exprimer des rapports syntaxiques polysémiques (cause + condition, cause + temporalité, par exemple).

Autrement dit, nous faisons l'hypothèse que c'est parce qu'il est en même temps polyvalent et implicite que ce type de forme pourrait avoir été marginalisé dans le discours oral de type narratif et que sa richesse sémantique, en revanche, est un potentiel énorme largement exploité dans la prose journalistique plutôt axée sur un style argumentatif. Or, il est à propos de souligner que le fait d'avoir mis en évidence une propriété pouvant expliquer les usages attestés du gérondif ne nous dit rien quant au statut grammatical du gérondif à l'oral. S'agit-il uniquement d'une différence de styles communicatifs qui sont à la disposition de n'importe quel locuteur – ce qui semble être la position de Rossi-Gensane (2011) - ou bien a-t-on affaire à une fracture grammaticale plus profonde ? Depuis les recherches de Marchello-Nizia (2006) et Badiou-Monferran (2008), une linguistique de la disparition refait surface, qui prend en considération les raisons sous-jacentes aux pertes et non pas seulement aux acquisitions des formes. On est en droit, alors, de se demander si l'aspect moribond du gérondif ne serait pas dû au fait que, telle qu'elle est utilisée à un niveau plus normatif, cette forme est en train de sortir de manière plus ou moins définitive de la grammaire de l'oral et que les constructions « aberrantes » (mi-gérondivales mi-participiales) que l'on a commentées n'en seraient qu'un dernier avatar²⁴. Une telle hypothèse, encore que provisoire, et qui, de toute évidence, mérite de plus amples confirmations empiriques, pourrait s'inscrire dans le cadre plus général proposé par Zribi-Hertz (2011) où le français oral est désormais considéré comme un système de plus en plus indépendant du français normé (attesté à l'écrit et véhiculé par l'école). Il s'agit d'un modèle explicatif extrêmement innovant et qui est centré sur l'hypothèse diglossique selon laquelle le français contemporain devrait être décrit comme présentant deux grammaires en intersection : grammaire standard protocolaire (GS) et grammaires dialectales informelles (GD). Prenons un exemple. Alors que, par exemple, la grammaire standard prévoit la négation complexe /ne ... pas/ - ce qui est conforme à la grammaire de l'écrit véhiculée par l'enseignement scolaire - la forme dialectale (au sens de non standard) prévoit uniquement le morphe /pas/²⁵ :

L'activation de GS, qui implique de la part du locuteur une vérification de la conformité du discours qu'il produit à la norme standard, est socialement attendue dans divers types d'activités passant souvent, quoique non nécessairement par l'écriture. (Zribi-Hertz, 2011 : 237-238)

Dans la grammaire protocolaire, le lien entre la forme écrite et la norme étant très étroit, nous pouvons par contrecoup définir les grammaires dialectales comme l'expression d' :

une activité dans toute situation où le locuteur ne cherche pas à se conformer à la norme. (Zribi-Hertz, 2011 : 237-238)

Comme le rappelle Massot (2010), l'existence d'une zone intermédiaire incluant ces locuteurs à même d'utiliser l'une et l'autre variantes n'exclut pas que d'un point de vue formel les deux grammaires sont à distinguer. Cela expliquerait bien l'apparition sporadique du gérondif à l'oral dans le discours de quelques locuteurs (en particulier en relation avec certaines classes lexicales). Pour notre part, nous serions enclins à considérer que dans les GD, le gérondif est une forme dont les caractéristiques sont de moins en moins nettes (ce que la superposition avec le participe présent semble démontrer) et dont l'usage se raréfie à telle enseigne que – si nos résultats quantitatifs étaient confirmés pour d'autres types d'oral – on pourrait imaginer sinon sa disparition à moyen terme, au moins sa limitation à quelques verbes fréquents, donnant lieu à des gérondifs appelés, selon notre hypothèse, à une grammaticalisation probable. Reste à comprendre la raison profonde d'une telle transformation. Nous nous bornons à signaler que le caractère implicite et le contenu sémantique à géométrie variable du gérondif pourraient avoir un rôle majeur dans ce changement. Terminons avec trois questions ouvertes qui sont autant de pistes de recherche que nous sommes en train de poursuivre. D'une part, il y a le problème du gérondif et de son rapport avec le participe présent : quelle est la présence de ce dernier dans les corpus oraux et écrits? quel est le taux de superposition de ces deux formes ? De plus amples confirmations sont nécessaires pour mieux étayer notre hypothèse d'une augmentation tant des coalescences entre ces deux formes que des gérondifs grammaticalisés. Un deuxième problème concerne la fiabilité de nos données quantitatives. L'écart entre la dimension de notre sous-corpus PFC et celle de LH2010 est-il trop important pour arriver à des conclusions si générales ? Cela nous oblige d'une part à considérer nos résultats toujours avec une certaine prudence, d'autre part nous incite aussi bien à enrichir le corpus oral qu'à diminuer le corpus écrit afin que les deux soient plus proches en termes de taille. Reste par ailleurs à compléter l'analyse lexicale de façon à pouvoir mieux décrire les préférences sémantiques : y a-t-il des lexèmes qui seraient plus attestés (tantôt au gérondif tantôt au participe présent) ? Y a-t-il des causes sous-jacentes ? Nous sommes, partant, tout à fait conscients du caractère somme toute provisoire de notre démarche, mais en même temps convaincus que notre contribution quali-quantitative propose des observations empiriques assez insolites tout en soulevant un certain nombre de questions cruciales qui, de toute façon, restent opaques dans les analyses purement spéculatives.

Références bibliographiques

- 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.
- Badiou-Monferran, Cl. (2008). Les disparitions de formes sont-elles des épiphénomènes ?, In Durand J., Habert, B., Laks, B., *Congrès Mondial de Linguistique française, CMLF '08*, 147-158.
- Bolasco, S. (2010). *TaLTaC2.10, Sviluppo, esperienze ed elementi essenziali di analisi automatica dei testi*. Milano : LED.
- Chierchia, G. (1997). *Semantica*. Bologna : il Mulino.
- Combettes, B. (2003). L'évolution de la forme en *-ant* : aspects syntaxiques et textuels. *Langages*, 149, 6-24.
- Gettrup, H. (1977). Le gérondif, le participe présent et la notion de repère temporel. *Revue Romane*, 12, 211-270.

- Halmøy, O. (2003). *Le gérondif en français*. Paris-Gap : Ophrys.
- Herlsund, M. (2000). Le participe présent comme co-verbe. *Langue française*, 127, 86-94.
- Kleiber, G. (2008). Le gérondif : de la phrase au texte, In Bertrand, O., Prevost, S., Charolles, M., François, J., Schnedecker, C., *Discours, diachronie, stylistique du français (études en hommage à Bernard Combettes)*. Berne : Peter Lang, 109-124.
- Kleiber, G. (2011). Gérondif et manière, *Langue française*. 171, 117-134.
- Kleiber, G., Theissen, A. (2006). Le gérondif comme marqueur de cohésion et de cohérence, In Messili-Ben Aziza, Z., Tullon H., *Cohérence et discours*. Paris : PUPS, 173-184.
- Laks, B. (2011). La phonologie du français et les corpus. *Langue française*, 169, 3-17.
- Marchello-Nizia Ch. (2006). *Grammaticalisation et changement linguistique*. Bruxelles : De Boeck - Duculot.
- Massot, B. (2010). *Le patron diglossique de variation grammaticale en français*. *Langue française*, 168, 87-106.
- Reichler-Béguelin, M.-J. (1995). Les problèmes de l'emploi du gérondif et des participiales en français contemporain. In Zaleska K. et Cataldi A., *Le français langue étrangère à l'Université : théorie et pratique*. Actes du Colloque de Varsovie, Varsovie, Uniwersytet Warszawski, Institut Romanistyli, 243-260.
- Rihs, A. (2009). Gérondif, participe présent et expression de la cause. *Nouveaux cahiers de linguistique française*, 29, 197-214.
- Rossi-Gensane, N. (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*. Paris : Ophrys, 83-106.
- Zanola, M.T. (1998). Gérondif et constructions participiales dans la perspective de l'analyse sémantico-textuelle. In Cigada, S., *Studi di Linguistica francese in Italia 1960-1996*. Brescia : La Scuola, 255-266.
- 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.

¹ Cette étude est le fruit d'une collaboration, aussi bien théorique que pratique, des trois auteurs ; plus précisément, Oreste Floquet a rédigé le paragraphe 1.1, et les chapitres 2 et 3, Marie-Pierre Escoubas-Benveniste les paragraphes 1.2, 1.3, Sergio Bolasco, à qui l'on doit tous les calculs statistiques, le paragraphe 1.4. Un vif remerciement à Maria Teresa Zanola pour ses précieuses suggestions et à Jean-Marie Viprey qui a généreusement mis à notre disposition son corpus de presse.

² Pour une revue des différentes positions concernant la nature du gérondif et l'interprétation de son statut morphématique ou bi-morphématique, voir Arnavielle (2010).

³ Notre étude s'inscrit dans le droit fil des travaux de Zanola (1998), Combettes (2003), Kleiber (2008) et (2011) et Kleiber et Theissen (2006) qui analysent de manière attentive le rapport entre le gérondif et les paramètres discursifs : thème, rhème, co-texte, contexte, etc., tantôt en diachronie tantôt en synchronie. Par rapport à ces études, toutefois, nous essaierons de présenter aussi bien un examen statistique des formes (dépassant ainsi le cadre strictement qualitatif qui les distingue), qu'une comparaison entre les différents styles dans le but de démontrer que la variation joue un rôle majeur dans le fonctionnement du gérondif.

⁴ Le corpus est disponible en ligne à l'adresse : www.projet-pfc.net.

⁵ Quand nous parlons d'oral spontané, nous entendons les manifestations orales s'opposant à toutes ces formes d'oral formel (radio, télévision, conférence, etc.) qui ne sont le plus souvent que de l'écrit oralisé. Pour ce qui est du protocole adopté dans toutes les enquêtes du PFC, nous rappelons brièvement que ce qui différencie la discussion libre et la discussion guidée c'est que dans le premier cas, il n'y a aucune spécialisation des rôles, l'enquêteur participant, lui-même, à la conversation. La discussion guidée, en revanche, met face à face un enquêteur et un enquêté. Soulignons aussi que dans le choix des informateurs, on suit toujours la méthode dite de la proximité sociale et de la familiarité. Les interviewés sont des personnes déjà connues de l'enquêteur (amis, parents, etc.), si

bien que l'échange qui en résulte est moins formel et plus spontané que s'il s'agissait de personnes qui n'entretiennent aucun rapport entre elles.

⁶ www.taltac.it, voir Bolasco (2010).

⁷ Par souci de complétude, nous citons ci-après les formes graphiques des insertions pronominales possibles : *en, la, le, l', les, leur, lui, me, m', te, t', se, s', nous, vous, y*. Précisons, en outre, que ces formes clitiques peuvent être au nombre de 2 (*en l'y invitant*) ce qui amène à 3 le nombre de places disponibles entre /en/ et /-ant/ si l'on ajoute aux pronoms les formes graphiques de la négation (*ne, n'* ; p.e. *en ne le leur ayant jamais dit*).

⁸ Dans le paragraphe 2.3, nous donnons quelques détails quantitatifs concernant cet aspect.

⁹ Gras et italique de l'auteur. Les analyses de Rossi-Gensane se basent, elles aussi, sur le corpus PFC.

¹⁰ Les formes grammaticalisées du corpus PFC ne concernent que les verbes /aller/ et /attendre/ (12 occurrences de /aller/ dont 1 seule pleinement gérondive, 10 occurrences de /attendre/ dont 3 pleinement gérondives). Pour ce qui du verbe /partir/, les formes attestées ont toutes une valeur pleinement verbale. Plus en général sur la grammaticalisation des gérondifs, voir Halmøy (2003 : 77-79). Rappelons brièvement les grammaticalisations possibles : (a) adverbiales (*en attendant, en passant*), prépositionnelles (*en allant à/sur/vers, en partant de, en passant par*), conjonctionnelles (*en admettant que, en attendant que/de, en supposant que*).

¹¹ Cet écart est parfaitement illustré par la classe des gérondifs de verbes de parole (réduite à *dire, parler, expliquer, résumer* à l'oral) et qui s'élève à plusieurs dizaines à l'écrit (*admettre, énoncer, déclarer, annoncer, etc.*).

¹² Les chiffres relatifs aux hapax font en effet partie des valeurs considérées comme normales en matière de statistiques textuelles opérant sur des échantillons de textes non particulièrement contraints. D'autre part l'écart entre ces valeurs pour l'oral et l'écrit est à relativiser par rapport à la différence de volume des deux corpus, si bien que nous pouvons affirmer que dans les deux cas la loi de Zipf semble bien être confirmée. Pour affiner l'analyse il conviendra donc de la reconduire sur un corpus oral plus ample et d'équilibrer les tailles des corpus.

¹³ Soulignons toutefois que dans certains idiolectes le gérondif de certains verbes impersonnels, le type *en pleuvant*, est possible, encore que très marqué. Nous en voulons pour preuve le distique d'Alphonse de Lamartine : *Chaque goutte en pleuvant remontait en poussière / sur l'herbe, et s'y roulait en globes de lumières* (Jocelyn, quatrième époque).

¹⁴ Les pronominaux présentent l'avantage de fournir des formes non ambiguës au niveau même du gérondif (par le biais de *se* et *s'*). Il convient de préciser deux points à leur endroit : a) ce ne sont naturellement pas les seuls cas de pronoms faisant partie intégrante de la structure lexicale (*en* et *y* sont aussi des éléments de locutions verbales du type *s'en prendre à, s'y mettre*) ; b) le pronom *se, s'*, n'épuise pas les occurrences de pronominaux (p.e. *en nous interrogeant, en me réclamant*) mais ces formes-là nécessitent une analyse de la valence verbale à partir des concordances.

¹⁵ Dans le cadre de cette recherche, nous ne nous intéressons qu'à la négation du syntagme gérondif. Or la négation du type /sans/ + infinitif, et qui entre dans certains cas en concurrence avec la précédente, est également à considérer afin de dresser un cadre complet de l'usage actuel différentiel du gérondif à l'écrit et dans l'oral spontané. Nous nous promettons d'y revenir par la suite.

¹⁶ Cette première enquête étant essentiellement de nature statistique et morphosyntaxique, nous n'avons pas analysé dans le détail tous les aspects sémantiques. Signalons toutefois que du point de vue aspectuel, les verbes atéliques (a) monopolisent le devant de la scène (environ 50 % des lexèmes) ; suivent les verbes téliques (b) qui sont aussi assez bien représentés, alors que les statifs (c) sont tout à fait minoritaires :

(a)

31atc1_GG

E: Et donc euh, oui alors bon, famille d'origine portugaise, donc vous avez été élevé
en parlant portugais ou en parlant un peu, oui ?

(b)

50app1_LG

PP : Elle est bien rentrée oui il vous a appelé hier soir **en rentrant** ?

(c)

31age1_LG

E2: Il était, il a été, il était beaucoup plus proche de nous **en restant** dans son rôle de professeur.

¹⁷ Nous ne nous référons pas au problème du sujet zéro. Comme l'ont déjà souligné différents auteurs, les sujets implicites du gérondif peuvent ne pas être coréférentiels au sujet du verbe défini. Voir Reichler-Béguelin (1995) et Halmøy (2003).

¹⁸ Voir Arnavielle (2003).

¹⁹ Le protocole PFC prévoit une transcription orthographique standard sans normalisation des éléments morphémiques pouvant être absents (le type *y a pas* pour *il n'y a pas*), et la ponctuation reflète, par ailleurs, l'interprétation syntaxique du transcripteur.

²⁰ A ce propos, voir Gettrup (1977) et Rihs (2009). Pour un point de vue différent voir Herslund (2000).

²¹ Voir Gettrup (1977).

²² Sur la possibilité que seulement la fin de l'action véhiculée par le gérondif (et non sa totalité) coïncide avec le début de l'action du verbe défini, voir Gettrup (1977).

²³ La locution gérondive *en parlant de X* ne figure pas parmi les types de grammaticalisation répertoriés par Halmøy (2003).

²⁴ À ce propos, il faut rappeler qu'à l'oral le participe présent est encore moins présent que le gérondif. Nous en repérons une centaine dans le corpus PFC. Cet aspect-là déborde du cadre de cette étude et fera l'objet d'une communication qui actuellement est en cours de préparation.

²⁵ Au sujet de la négation, voir, dans la même veine, Barra-Jover (2004).