

***Je dirais que la grosse majorité sont là***  
**Questions d'accord(s) avec les noms collectifs dans les données *OFROM***

Claus D. Pusch

Albert-Ludwigs-Universität Freiburg  
claus.pusch@romanistik.uni-freiburg.de  
<https://orcid.org/0000-0002-4083-8663>

Reçu le 28/12/2019, accepté le 21/3/2020, publié le 5/11/2020 selon les termes de la licence  
*Creative Commons Attribution 4.0 International (CC BY 4.0)*

**Résumé :** Les noms collectifs du français (Ncoll), qui désignent une pluralité de référents par une forme lexicale au singulier, sont connus pour être soumis à un accord verbal (mais aussi pronominal) variable, notamment dans la langue orale. L'article présente une description de la distribution d'une sélection de Ncoll, établie sur la base des échantillons de Tristram (2014) et de Mougeon & Mougeon (2017a), dans le français parlé en Suisse romande tel qu'il est documenté dans le corpus *OFROM*, et le comportement de ces Ncoll face à l'accord. Les résultats globaux indiquent que les Romands se conforment assez fidèlement à la norme et emploient majoritairement l'accord grammatical. Une partie des occurrences à accord sémantique s'explique par l'effet d'accord par proximité après des Ncoll assortis d'un complément, alors que les résultats concernant l'effet de l'éloignement syntagmatique et structural entre le Ncoll et les éléments avec lesquels il est censé être en relation d'accord, ne s'avèrent pas probants.

**Abstract:** In French, collective nouns (Ncoll), which refer to a plurality of objects or individuals through a singular form, are known to be subject to variable verbal (but also pronominal) agreement, particularly in spoken language. The study presented here analyzes the *OFROM* corpus to provide a distributional description of an array of Ncoll, selected on the basis of the samples used by Tristram (2014) and Mougeon & Mougeon (2017a), in the variety of French spoken in Switzerland. Specifically, it focuses on the agreement patterns associated with these nouns. The results indicate that Swiss speakers of French tend to closely follow the norm and mostly apply grammatical agreement. A part of the diverging tokens with semantic agreement attested in the data can be explained as agreement attraction by a pluralized nominal complement within the NP, while the results concerning the effect of syntagmatic distance and intervening syntactic boundaries between the Ncoll and the target(s) for agreement turn out to be inconclusive.

## 1 Introduction

[1] Notre contribution est consacrée aux noms collectifs (Ncoll) comme *famille*, *équipe*, *majorité* ou *monde* qui désignent des ensembles ou des regroupements de référents. La spécificité de ces noms est qu'ils expriment une pluralité référentielle à travers une morphologie au singulier. Ce fait a des répercussions morpho-syntaxiques notamment dans la modalité orale de la langue dans la mesure où l'accord avec des éléments (verbaux ou pronominaux) avec lesquels ces noms entretiennent une relation de dépendance syntaxique peut se faire au singulier, comme dans (1), ou au pluriel, comme dans (2)<sup>1</sup> :

- (1) tu vas dans l'administration communale *tout le monde* se connaît (unine11-vra, unine11c23m)
- (2) ben *tout le monde* au bout de la troisième attaque *font* la faute (unine18-010, unine18a15d)

[2] Notre propos est de fournir une approche descriptive d'un échantillon raisonné de ces noms et des syntagmes nominaux dont ils constituent le noyau, tels qu'ils se trouvent distribués dans le corpus *OFROM* (Avanzi, Béguelin & Diémoz 2012-2019). Je m'intéresserai particulièrement à l'accord variable, illustré précédemment, et aux facteurs structuraux susceptibles de favoriser ou de freiner l'accord sémantique au pluriel. Cette description (§ 4 et 5) est précédée d'une synthèse des approches définitoires des noms collectifs (§ 2). Dans le § 3, je résumerai, d'un côté, les opinions exprimées dans la littérature spécialisée par rapport au statut de la variabilité de l'accord après les noms collectifs en français et, de l'autre côté, les résultats de deux travaux empiriques récents sur le sujet qui, par leur approche basée sur corpus oraux, se prêtent à une comparaison avec les résultats obtenus à partir des données d'*OFROM*.

## 2 Les noms collectifs : approches définitoires

[3] En partant du terme même, on peut définir les noms collectifs comme des noms qui « désignent des collections ou collectivités d'entités isolables » (Riegel, Pellat & Rioul 2018 [1994] : 325). Ce sémantisme, qui implique une opposition entre des noms communs à référence individuelle vs non-individuelle et qui en même temps concerne la relation sémantique partie-tout (cf. Kleiber, Schnedeker & Theissen 2006), paraît évident à première vue, mais il est loin d'être banal et une vaste bibliographie est consacrée à délimiter l'inventaire des noms qui y entrent (cf., pour le français, les références mentionnées dans Flaux 1999, Lam-

<sup>1</sup> Les exemples numérotés proviennent tous du corpus *OFROM* ; ils sont identifiés par le code du locuteur ou de la locutrice, suivi du code de l'enregistrement. Les conventions de transcription d'*OFROM* sont conservées, mais les Ncoll et d'autres éléments sur lesquels porte l'analyse sont en italique.

mert 2010 et Lammert & Lecolle 2014, Lecolle 2019<sup>2</sup>, entre autres). Certains auteurs vont jusqu'à inclure les formes pluralisées des noms individuels ; ainsi, *les maisons* dans *les maisons de la Grand'Rue* serait, dans l'optique très 'généreuse' de l'approche de Cabezas Holgado (2017), une *expression collective syntaxique*, alors que ce linguiste considère un nom comme *sofa* comme *expression collective composée* car englobant les différentes pièces qui constituent ce meuble. D'autres chercheurs plaident en faveur d'une application plus restreinte du terme. Flaux (1999 : 422) définit les Ncoll comme « des noms dénotant des entités composées de parties distinctes préalablement constituées, homogènes entre elles, mais autonomes et hétérogènes par rapport au tout », c'est-à-dire : par rapport à la collection. Par conséquent, *le bouquet* est accepté par cette auteure comme Ncoll, alors qu'elle refuse ce statut à *le mobilier*. Dans un sens semblable, Joosten (2006) distingue les 'vrais' noms collectifs tels *club* ou *armée*, dont les parties constitutives sont liées entre elles par contiguïté, de ce qu'il propose d'appeler *noms agrégatifs* (*aggregate nouns*), comme *lingerie* ou *feuillage*, dont les éléments constitutifs se trouvent essentiellement dans une relation de similitude.

[4] Que les Ncoll puissent être rapprochés des noms individuels pluralisés découle de leur « pluralité ontologico-sémantique interne » que Lammert (2010 : 71-72) considère « être une caractéristique décisive pour différencier ces noms des autres et cela quelles que soient les théories prises en compte ». Cependant, cette auteure souligne qu'on est devant « une discordance morpho-sémantique : la morphologie d'un Ncoll ne reflète en rien la sémantique plurielle de cette catégorie de noms ». À côté de l'opposition entre singularité et pluralité sémantiques, par rapport à laquelle les Ncoll s'avèrent problématiques, il faut aussi évoquer leur statut quant à la distinction sémantico-lexicale entre noms comptables (qui « renvoient à des segments discontinus de la réalité » (Riegel, Pellat & Rioul 2018 [1994] : 323)) et noms massifs (qui « dénotent des substances continues qui obéissent aux principes de la référence cumulative homogène » (Riegel, Pellat & Rioul 2018 [1994] : 323)). Si certains linguistes (comme Bußmann 2002 : 423) rapprochent les Ncoll des noms massifs en les considérant comme un sous-type de ces derniers, d'autres ne soutiennent pas un tel rapprochement ; c'est le cas de Riegel, Pellat & Rioul (2018 [1994] : 325) qui soulignent que les Ncoll « tiennent à la fois des noms comptables et des noms massifs ». Laycock (2006 : 534) rejette la notion de *noms massifs* et plaide en faveur d'une simple distinction entre *noms comptables* (*count nouns* / CN) et *noms non comptables* (*non-count nouns* / NCN) : « CNs, or their occurrences, are semantically either singular or plural, to be non-count is simply to be neither singular nor plural ». Laycock (2006) montre que les noms comptables pluralisés et les NCN (massifs et autres) partagent des traits essentiels ; d'autre part, certains noms comptables pluralisés comme *clothes* semblent dénotativement proches de noms collectifs comme *clothing*. Mais cet

<sup>2</sup> Cet ouvrage récent ne m'était pas accessible au moment de la rédaction du présent article.

auteur de préciser : « while collective nouns like *clothing* and *furniture* might be said to be **ontologically** equivalent to cognate CNs – such NCNs are no less **semantically** non-count than non-collective nouns like *water* and *mashed potato* » (Laycock 2006 : 537 ; souligné dans l'original). Laycock (2006) situe les Ncoll dans une position intermédiaire entre les noms comptables et les noms massifs ou, dans son approche, entre CN et NCN, lesquels se laissent caractériser, comme il l'a souligné, par les paramètres de singularité et de pluralité. On peut donc conclure que cette distinction singularité vs pluralité joue un rôle de première importance dans la description sémantique des Ncoll, qui autrement s'avère difficile à cause de l'hétérogénéité des éléments lexicaux censés en faire partie.

[5] Un facteur contribuant à cette hétérogénéité est le degré de détermination référentielle qu'ont les items généralement inclus dans le groupe des Ncoll. Il paraît évident que les individus qui constituent la 'collectivité' désignée comme *gouvernement* sont ses ministres ou ses conseillers, et un nom collectif comme *armée* identifie à lui seul ses 'entités constitutives', de sorte qu'au moins dans un contexte neutre, il ne faut pas préciser qu'il s'agit d'une *armée de soldats*. D'autres Ncoll comme *un bouquet* apparaissent préférablement accompagnés d'un complément qui les détermine. Finalement, des Ncoll comme *la majorité* ne peuvent normalement pas se passer de complément. Cabezas Holgado (2017 : 18) appelle ce dernier type *noms collectifs indéterminés (colectivos indeterminados)*, par opposition aux noms collectifs déterminés du type *gouvernement*, et l'associe aux numéraux et, plus généralement, aux quantificateurs. L'inclusion de ces noms quantifiants dans le groupe des Ncoll ne fait pas l'unanimité ; Michaux (1992) et Lamert (2010) ne les considèrent pas comme de vrais noms collectifs et les excluent de l'inventaire.

[6] Étant donné que la définition des Ncoll sur la base de critères sémantiques s'avère problématique, il paraît intéressant de chercher des critères morpho-syntaxiques qui puissent contribuer à circonscrire cette classe de noms, d'autant plus que l'opposition sémantique 'singularité' vs 'pluralité', si essentielle pour saisir la spécificité des Ncoll, se manifeste dans la morphologie et la syntaxe à travers le trait flexionnel du marquage du nombre. Selon Joosten (2006 : 73), c'est cette perspective morpho-syntaxique qui caractérise l'approche anglo-saxonne du domaine, où « it is customary to define collective nouns fairly strictly, on the (primarily) syntactic basis of variable concord ». En effet, en anglais l'équivalent de la phrase *La police a arrêté cinq suspects*, avec le Ncoll en position de sujet, serait *The police has arrested five suspects* mais peut aussi prendre la forme *The police have arrested five suspects*, les deux versions étant courantes et acceptées par la norme, au moins dans une partie des variétés de cette langue<sup>3</sup>. Le focus sur l'ac-

3 Si une phrase *La police ont arrêté cinq suspects* serait surprenante en français, la modalité orale de cette langue peut facilement introduire la pluralité morpho-syntaxique moyennant une construction à dislocation avec pronom de relais (*La police ils ont arrêté cinq suspects*). Cette configuration, évoquée sous § 5.4, est illustrée par les ex. (6), (9), (14), (16-17) et (19).

cord en nombre variable comme critère-clé pour la définition des Ncoll réduit du coup l'inventaire de ces noms en anglais car les données empiriques montrent que seuls les collectifs animés – et préférablement humains – montrent l'accord variable de façon récurrente (Levin 2001 : 11-14). Qu'en est-il de cet accord variable en français ?

### 3 La variabilité de l'accord avec les Ncoll en français

[7] Dans son étude détaillée des Ncoll du français, Lammert (2010 : 90) souligne que les questions d'accord « ne se posent pas directement pour le français », au moins en ce qui concerne des phrases comme *La police a arrêté cinq suspects*. Avec ces Ncoll déterminés (dans la terminologie proposée par Cabezas Holgado (2017) que nous adoptons) en position sujet, « le verbe sera toujours conjugué au singulier, puisque la morpho-syntaxe l'emporte sur la sémantique » (Lammert 2010 : 91). En même temps, elle reconnaît la possibilité d'un accord verbal au pluriel pour les SN constitués d'un Ncoll suivis d'un complément, comme dans son exemple *Une foule de touristes visiteront Paris* (Lammert 2010 : 91). Lammert (2010 : 91) propose une explication syntaxique en disant que, dans ce cas de figure où le Ncoll ne fonctionne plus comme « nom plein », mais plutôt comme déterminant du nom au pluriel qui le suit sous forme de complément prépositionnel, c'est ce deuxième nom qui est la tête du SN binominal et qui déclenche l'accord au pluriel. Étant donné que l'accord du verbe au pluriel après ce type de SN binominal avec Ncoll sous-déterminé (c'est-à-dire : fréquemment assorti de complément, comme *foule* ou *équipe*) ou Ncoll non déterminé (c'est-à-dire : normalement suivi par un complément, comme *plupart* ou *majorité*) suivi d'un complément qui, lui, est au pluriel, n'est pas systématique, cette auteure conçoit l'accord variable comme résultat d'une ambiguïté structurale.

[8] Riegel, Pellat & Rioul (2018 [1994]) présentent le phénomène de l'accord de la même façon mais proposent une explication plutôt cognitive pour sa variation :

[...] lorsque le nom collectif au singulier est suivi d'un complément de nom au pluriel, le verbe se met au singulier ou au pluriel, selon la manière d'envisager le référent. [...] L'accord se fait au singulier avec le nom collectif qui privilégie la vision collective des référents, considérés globalement [...] L'accord se fait au pluriel quand le nom collectif est un simple quantificateur qui dénote la diversité plurielle des individus auxquels réfère le groupe nominal [...] (Riegel, Pellat & Rioul 2018 [1994] : 498)

Et Riegel, Pellat & Rioul (2018 [1994] : 498-499) d'ajouter plusieurs alinéas sur des Ncoll particuliers, les oscillations auxquelles peut être soumis l'accord du verbe qui les suit, et sur les « tolérances officielles » de la norme à propos de ces oscillations, en laissant entendre que parfois les locuteurs sont peu sûrs de l'usage

à suivre. Même constat de la part de Lammert & Lecolle (2014 : 212) : « on rencontre fréquemment, dans les faits, hésitations et fluctuations portant sur des phénomènes d'accord », notamment dans la langue parlée.

[9] De ce qui vient d'être présenté, il s'ensuit que l'accord variable existe en français mais qu'il paraît (a) limité *grosso modo* aux Ncoll à trait sémantique [+animé] voire [+humain] (à l'instar de l'anglais) et (b) restreint à des Ncoll suivis de complément comme *une équipe de cyclistes* ou *une foule de touristes* (ce qui diffère de l'anglais). Lammert (2010 : 91) désigne ce phénomène comme *accord sylleptique* – terme issue de la tradition rhétorique –, mais il a reçu aussi d'autres noms comme *accord sémantique* (Corbett 2006, entre autres), *constructio ad sensum* (Koch & Oesterreicher 1990, entre autres), *accord associatif* (Berrendonner & Reichler-Béguelin 1995) ou *discordance*. Signalons enfin que Lammert (2010) et Riegel, Pellat & Rioul (2018 [1994]) – comme d'ailleurs de nombreux linguistes qui se sont penchés sur le sujet de l'accord variable – ne semblent avoir en vue que les relations entre un Ncoll en position sujet et le SV qui s'y réfère dans les limites d'une seule proposition.

[10] Si la description sémantique des Ncoll du français a attiré l'attention des linguistes, on trouve peu de travaux empiriques consacrés à l'usage effectif de ces noms et de la variabilité de l'accord, notamment dans la modalité orale de la langue où, selon l'avis général, les 'hésitations' et les 'fluctuations' sont particulièrement prononcées. Cependant, il existe deux références récentes sur le sujet : d'une part l'étude de Tristram (2010, 2014) sur l'accord variable dans le français européen de France, et d'autre part l'étude de Mougeon & Mougeon (2017a, 2017b) sur la variation de l'accord dans le français canadien d'Ontario. Cette dernière étude suit l'approche variationniste labovienne classique et utilise des données synchroniques issues d'entretiens sociolinguistiques avec plus de 240 témoins. Dans leur analyse statistique, les auteurs tiennent compte de quatre facteurs linguistiques internes et de trois à quatre facteurs externes selon les groupes de témoins. Étant donné qu'il s'agit de locuteurs ontariens qui vivent en situation de minorisation linguistique et donc en contact plus ou moins intense avec l'anglais, on trouve, parmi les facteurs externes, le degré de bilinguisme et celui de la restriction linguistique, fortement pertinents dans une telle situation sociolinguistique. Comme les Ncoll constituent un groupe de lexèmes et, par conséquent, un paradigme plutôt ouvert et difficile à délimiter, il se pose évidemment la question de l'échantillon pour la constitution du corpus de travail, c'est-à-dire celle de la liste des Ncoll à inclure dans l'analyse. Mougeon & Mougeon (2017a, 2017b) optent pour une définition restreinte et n'acceptent que de 'vrais' Ncoll, à savoir des noms déterminés capables de désigner des collectivités sans complément, du type *police*, des noms sous-déterminés du type *classe* qui font souvent, mais non pas impérativement appel à un complément pour être déterminé, et le nom *monde* (dans ses différentes déclinaison) sur lequel on reviendra plus loin. Par contre, ils excluent les Ncoll non déterminés du genre *plupart* ou *majorité* qui, eux, en prin-

cipe ne peuvent pas se passer de complément nominal par rapport auquel ils fonctionnent comme déterminant. Le corpus de travail ainsi constitué est assez impressionnant, le nombre d'énoncés à Ncoll analysé s'élevant respectivement à 1817 (pour Mougeon & Mougeon 2017a) et 1570 occurrences (pour Mougeon & Mougeon 2017b)<sup>4</sup>.

[11] La recherche de Tristram (2010, 2014) diffère considérablement dans sa méthode et sa portée de celle de Mougeon & Mougeon (2017a, 2017b). Elle étudie un choix de Ncoll en français européen en synchronie (sur corpus propre) et en diachronie (en utilisant les données de *Frantext*). Pour le volet synchronique, elle applique une triangulation méthodologique qui comprend des données orales spontanées issues d'entretiens d'une durée totale de 32 heures, et des données élicitées par questionnaire, les deux types de données étant produits par 42 locuteurs et locutrices habitant en France septentrionale. En accord avec son approche méthodologique, l'analyse de Tristram (2010, 2014) se veut quantitative et qualitative à la fois. Les données des entrevues, comparables à celles d'*OFROM*, lui fournissent un corpus de travail très limité (en tout cas en comparaison avec celui de Mougeon & Mougeon 2017a, 2017b) composé de 12 Ncoll dont un seul du type déterminé (*gouvernement*), trois du type sous-déterminé et le reste appartenant à la catégorie des Ncoll non déterminés ; de plus, elle y inclut l'expression fractionnelle *X sur N* (*un sur dix*) bien que celle-ci ne contienne pas d'élément nominal à proprement parler. Le nombre total d'occurrences qu'elle soumet à l'analyse ne s'élève qu'à 34<sup>5</sup>.

[12] Les travaux de Mougeon & Mougeon (2017a, 2017b) et de Tristram (2010, 2014) fournissent les balises pour notre étude des Ncoll dans le corpus *OFROM*, dont la constitution du corpus de travail et les critères d'analyse seront détaillés dans le paragraphe suivant.

#### 4 Corpus de travail et méthodologie

[13] Dans un souci de comparabilité des résultats obtenus à partir du corpus *OFROM* avec ceux de Mougeon & Mougeon (2017a) et Tristram (2014), l'échantillonnage des Ncoll à extraire – choix inévitablement sélectif pour un paradigme lexical relativement ouvert et donc de délimitation malaisée – s'est effectué par fusion des échantillons des auteurs cités. Rappelons que Mougeon & Mougeon (2017a) n'incluent que des Ncoll déterminés et sous-déterminés alors que Tristram (2014) étudie aussi – et même majoritairement – des Ncoll non déterminés. La liste des Ncoll cherchés dans *OFROM* se compose donc des items suivants (dans

4 La différence s'explique par le fait que dans Mougeon & Mougeon (2017a) les auteurs incluent des données provenant de témoins adolescents et de témoins adultes, alors que dans Mougeon & Mougeon (2017b) seules les entrevues avec les adolescents sont analysées. À part cela, les objectifs et les critères d'analyses diffèrent jusqu'à un certain point entre les deux articles.

5 Il faut concéder que la base empirique de Tristram (2014) est en réalité plus vaste et solide grâce aux données élicitées par questionnaire, qui pourtant ne permettent pas de comparaison.

l'ordre alphabétique) :

- (a) Ncoll déterminés (normalement sans complément) : *école, famille, gouvernement, police, université* ;
- (b) Ncoll sous-déterminés (fréquemment ou normalement avec complément déterminant) : *classe, entourage, équipe, groupe* ;
- (c) Ncoll non déterminés (normalement avec complément à déterminer) : *majorité, minorité, (la) moitié, moyenne, nombre, partie, (le) peu, plupart, série* ;
- (d) Ncoll construits avec *monde*, y compris *tout le monde*.

[14] Quelques items pris en compte dans les études de référence n'ont pas été inclus dans cette liste, pour des raisons diverses : d'un côté, l'expression fractionnelle *X sur Y* recueillie par Tristram, qui s'assimile fonctionnellement aux Ncoll non déterminés quantificateurs mais ne comporte pas d'élément nominal ; et, d'un autre côté, le Ncoll *parenté* (sans occurrences pertinentes dans *OFROM*) et les catégories *compagnies / entreprises* et *noms de pays* de Mougeon & Mougeon (2017a), où ce ne sont évidemment pas ces lexèmes qui ont été extraits mais des noms propres appartenant à ces catégories. Or, comme le corpus *OFROM* n'est pas étiqueté pour des noms propres, il n'aurait pas été possible de les identifier de façon efficace dans les données romandes. D'autre part, notre liste contient le Ncoll *plupart* dont Mougeon & Mougeon (2017a) ne parlent pas à cause de son caractère déterminant quantificateur mais qui est aussi écarté par Tristram (2014). En fait, dans une étude diachronique publiée ailleurs, Tristram & Ayres-Bennett (2012 : 374) observent que *plupart* n'est plus soumis à l'accord variable depuis le 18<sup>e</sup> siècle : « while many of these nouns [Ncoll. ; C.P.] exhibit variation between singular and plural verbal agreement, *la plupart* does not: in contemporary French it canonically occurs with a plural verb ». Si cette conclusion est en accord avec la norme, il nous a semblé prometteur de la mettre quand même à l'épreuve des données, vu que *plupart* est un Ncoll relativement fréquent. Un autre Ncoll d'une fréquence considérable est *monde* dans le sens de 'groupe de gens', pour lequel il nous paraît impossible de trancher s'il appartient à la catégorie (a) des noms déterminés ou à la catégorie (b) des noms sous-déterminés, d'autant plus que fonctionnellement il s'apparente à un pronom indéfini, notamment dans sa forme d'emploi la plus courante, *tout le monde*. Nous avons suivi Mougeon & Mougeon (2017a) en comptant séparément les occurrences de *tout le monde* et les autres occurrences de *monde* au sens cité, car déjà Blinkenberg (1950 : 67, en renvoyant à une étude antérieure de Høybye 1944) avait remarqué un comportement différent de la première forme, qui « a résisté à l'accord avec le sens », et des formes comme *bien du monde* où l'usage est variable<sup>6</sup>. Autre détail à remarquer : l'inclusion d'*école* et

6 Si *plupart* et *monde* sont des formes nominales à valeur sémantique collective (ce qui justifie leur inclusion dans cette étude), leur statut à côté de Ncoll comme *famille* ou *équipe* est indéniablement discutable. Comme on l'a dit supra, *plupart* (comme d'autres Ncoll quantificateurs) fonctionne plutôt comme le déterminant d'un nom dans sa suite ; *tout le monde*, en revanche, peut être classifié comme pronom indéfini, comme le souligne un évaluateur anonyme. Le degré de gram-



d'*université* dans une recherche sur les Ncoll ne se justifie que dans la mesure où ces noms désignant des institutions sont souvent employés métonymiquement pour désigner le collectif des personnes qui y sont rattachées (le corps des enseignants, le personnel, etc.). Cela vaut évidemment pour un grand nombre de noms d'institutions, et si *école* et *université* – et seulement ces deux noms-là – ont été pris en compte ici, c'est parce qu'ils apparaissent parmi les items de Mougeon & Mougeon (2017a).

[15] Afin de constituer notre corpus de travail, j'ai cherché les items de la liste par requête lexicale dans les transcriptions disponibles dans la version d'*OFROM* en ligne en octobre 2018. Comme les segments des transcriptions que montre le concordancier web du corpus ont été découpés selon des critères prosodiques plutôt que syntaxiques, il a fallu élargir parfois le contexte visualisé afin de s'assurer qu'il y ait une forme verbale ou autre susceptible de se trouver dans une relation d'accord avec le nom repéré. Or, cette vérification s'est avérée plus coûteuse que prévu, car si des phrases canoniques comme *Une foule de touristes visiteront Paris*, avec un accord intrapropositionnel facilement détectable, sont monnaie courante à l'écrit, les énoncés oraux ont souvent une structure syntagmatique plus complexe. L'ex. (3) illustrera ce point (le signe % indiquant un élément inaudible) :

- (3) par exemple si tu prends euh toute une *équipe* là qui est en | % | \_ | qui *embête* ou qui quoi qui *sont* euh un peu extrémistes dans leurs idées etcetera (unine15-006, unine15a07m)

Le concordancier d'*OFROM* montre seulement le segment *que par exemple si tu prends euh toute une équipe là qui est en*, avec un accord verbal non ambigu au singulier, donc un segment suffisant pour classer cette occurrence du Ncoll *équipe* (qui par ailleurs ne déclenche l'accord verbal qu'à travers le pronom relatif sujet ; j'apporterai des précisions sur cette constellation plus loin). Ce n'est qu'en élargissant considérablement le contexte droit qu'on détecte – outre le verbe *embête* avec un accord scripturalement singulier mais phoniquement ambigu – le verbe *sont* qui par lien anaphorique se réfère toujours à *équipe* mais qui montre un accord non ambigu au pluriel. Comme l'élargissement du contexte doit s'effectuer manuellement moyennant la fonction *Détails* pour chaque occurrence du Ncoll visualisée dans la concordance, il est plus que probable que de nombreux SV en relation d'accord interpropositionnel avec le Ncoll nous ont échappé. L'analyse des données ne peut donc pas prétendre à l'exhaustivité.

[16] Une autre facette du phénomène d'accord variable complique l'analyse. Comme souligné supra, les Ncoll donnent lieu à « une discordance entre leur pluralité sémantique [...] et leur morphologie singulière » (Riegel, Pellat & Rioul matisation de ces formes, lié à une tendance vers la recatégorisation, explique certainement leur fréquence élevée dans l'usage.

2018 [1994] : 325), et il paraît normal que cette tension entre singularité et pluralité se répercute avant tout sur l'accord flexionnel verbal en nombre. Cependant, à y regarder de près, les 'questions de l'accord' dépassent le seul accord verbal en nombre, et dans les faits on peut distinguer trois types d'accord susceptibles de variabilité : (a) au sein du SN, entre le Ncoll et un déterminant, un cas extrêmement marginal, le seul exemple trouvé dans *OFROM* étant (4) :

(4) regarde nos amis quoi *les les plupart* ils viennent euh de mon côté c'est les anciens élèves quoi (unifr11-aaa, unifr11a03d)

(b) entre le Ncoll et un (ou plusieurs) pronom(s) phorique(s) (5), qui lui (eux) peut (peuvent) contrôler les relations d'accord verbal (6) :

(5) en internat on accueillait les enfants le matin [...] la *plupart* faut *les* assister | \_ | donc faut *leur* brosser les dents faut *les* doucher faut *les* habiller (unine08-hda, unine08a13m)

(6) mais finalement l'école euh ben voilà *ils* doivent enseigner euh je verrais pas pourquoi *ils* mettraient | \_ | des noms neuchâtelois (unine15-901, unine15z01m)

et finalement (c) entre le Ncoll et le verbe, le cas le plus proéminent et le mieux décrit dans les études portant sur l'accord variable avec les Ncoll et déjà illustré dans (1-3)<sup>7</sup>. Cette variabilité concerne certes surtout l'accord en nombre, mais les traits flexionnels de la personne grammaticale et du genre peuvent aussi être concernés, notamment quand il s'agit de cas du type (b) impliquant des pronoms ana- ou (rarement) cataphoriques, où de surcroît la question de l'accord variable peut être pertinente en dehors des Ncoll en position sujet. Dans la présente étude, j'ai pris en compte dans la mesure du possible les trois types de constellations d'accord mais c'est uniquement sur le trait flexionnel du nombre que nous avons porté notre attention. Par rapport à cet accord en nombre, au moment du codage des occurrences, effectué sur feuille *Excel*, j'ai distingué entre un accord non ambigu singulier, un accord non ambigu pluriel et un accord ambigu, ces trois cas de figure ayant déjà été évoqués lors de la discussion de l'ex. (3). On sait qu'en français de nombreuses formes verbales de la 3<sup>e</sup> personne sont homophones et ne permettent pas de trancher, au moment de transcrire l'oral, s'il s'agit d'un singulier ou d'un pluriel. Évidemment, en cas de doute le transcripateur attribuera au locuteur une conscience normative et choisira par défaut la forme grammaticalement concordante ; mais dans une analyse consacrée à l'accord variable, il faut éliminer

<sup>7</sup> On peut considérer des occurrences du type (b) comme dans l'ex. (6) comme un sous-cas de (c) dans la mesure où le pronom clitique sujet tend à s'assimiler à une marque flexionnelle du verbe, en tout cas dans la modalité orale du français, où le paradigme des désinences verbales est réduit. Merci à Gilles Corminboeuf pour cette observation.

ou au moins traiter séparément ces formes ambiguës.

[17] Passons maintenant aux résultats obtenus à partir des données *OFROM* par ce procédé.

## 5 Les Ncoll dans *OFROM*

### 5.1 Résultats quantitatifs globaux et analyse comparative

[18] *OFROM* a fourni un total de 432 occurrences des Ncoll de notre échantillon qui se distribuent comme suit :

Ncoll	Accord ambigu	Accord non ambigu		Total
		singulier	pluriel	
Ncoll déterminés				
<i>école</i>	10	15	2	27
<i>famille</i>	20	19	10	49
<i>gouvernement</i>	1	3	0	4
<i>police</i>	2	0	0	2
<i>université</i>	2	4	0	6
Ncoll sous-déterminés				
<i>classe</i>	6	6	1	13
<i>entourage</i>	1	1	1	3
<i>équipe</i>	3	1	8	12
<i>groupe</i>	12	9	4	25
Ncoll non déterminés				
<i>majorité</i>	0	3	3	6
<i>minorité</i>	0	1	0	1
<i>(la) moitié</i>	1	8	1	10
<i>moyenne</i>	0	3	0	3
<i>nombre</i>	3	1	6	10
<i>partie</i>	2	15	1	18
<i>(le) peu</i>	0	0	0	0
<i>plupart</i>	9	1	37	47
<i>série</i>	2	0	2	4
Ncoll formés avec <i>monde</i>				
<i>tout le monde</i>	66	106	2	174
<i>autres</i>	8	6	4	18
<b>Total</b>	<b>148</b>	<b>202</b>	<b>82</b>	<b>432</b>

Tableau 1: Ncoll individuels dans *OFROM* par type d'accord (chiffres absolus)

Ncoll	Accord ambigu	Accord non ambigu		Total
		singulier	pluriel	
déterminés	35 / 40%	41 / 47,5%	12 / 13,5%	88 / 100%
sous-déterminés	22 / 41,5%	17 / 32%	14 / 26,5%	53 / 100%
non déterminés	17 / 17%	32 / 32,5%	50 / 50,5%	99 / 100%
formés avec <i>monde</i>	74 / 38,5%	112 / 58,5%	6 / 3%	192 / 100%
Total	148 / 34%	202 / 47%	82 / 19%	432 / 100%

Tableau 2: Types de Ncoll dans *OFROM* par type d'accord (chiffres absolus et relatifs)

[19] On voit que pour un tiers des occurrences, l'accord en nombre ne peut pas être déterminé avec certitude parce que les formes verbales du singulier et du pluriel sont homophones et qu'il n'y a pas d'autre indice pour les désambigüiser, de sorte que les transcrip-teurs d'*OFROM* ont opté pour la forme par défaut (et normative) du singulier. En même temps, on constate que pour tous les types de Ncoll on trouve une quantité non négligeable d'occurrences avec un accord au pluriel, hormis celui construit avec *monde*. La proportion élevée de l'accord sémantique pour les Ncoll quantificateurs s'explique par la haute récurrence du nom *plupart* qui, dans les données *OFROM*, représente à lui seul presque la moitié de ce type de Ncoll et pour lequel une seule occurrence non ambiguë avec accord au singulier a été détectée, et cela dans une configuration constructionnelle sur laquelle je reviendrai plus loin :

(7) c'est des bon c'est la *plupart* c'est des autistes (unine08-hda, unine08a13m)

Sans la forme *plupart*, dont le comportement dans *OFROM* confirme par ailleurs pleinement les affirmations de Blinkenberg (1950) et de Tristram & Ayres-Bennett (2012), la répartition pour les Ncoll non déterminés est la suivante : 32,5% avec accord ambigu, 59,5% avec accord non ambigu au singulier, 25% avec accord au pluriel. Cela veut dire que les Ncoll qui fréquemment ou obligatoirement ont besoin d'un complément déclenchent l'accord sémantique deux fois plus souvent que les Ncoll employés canoniquement sans complément ; là encore, les données *OFROM* confirment les généralisations avancées dans les travaux cités précédemment. Quelle image se dessine-t-elle quand on compare les données romandes *OFROM* aux études sur corpus portant sur le français de France et sur le français ontarien ?

[20] Dans le tableau 3, les cas d'accord non ambigu du tableau 1, réarrangés légèrement pour faciliter la comparaison, se trouvent confrontés aux résultats correspondants tirés de Tristram (2014) et de Mougeon & Mougeon (2017a, 2017b) ; les tirets indiquent que le Ncoll en question n'a pas été pris en compte dans l'étude concernée<sup>8</sup> :

<sup>8</sup> *Plupart* n'est pris en compte ni par Tristram (2014) ni par Mougeon & Mougeon (2017a, 2017b)

Corpus	OFROM		Tristram (2014)		Mougeon & Mougeon (2017a, 2017b)	
	Accord singulier	Accord pluriel	Accord singulier	Accord pluriel	Accord singulier	Accord pluriel
Ncoll déterminés						
<i>école + université</i>	19	2	---	---	4	37
<i>famille + parenté</i>	19	10	---	---	88	99
<i>gouvernement</i>	3	0	0	1	2	11
<i>police</i>	0	0	---	---	9	25
Ncoll sous-déterminés						
<i>classe</i>	6	1	1	0	3	17
<i>entourage</i>	1	1	0	0	---	---
<i>équipe + groupe</i>	10	12	0	1	11	39
Ncoll non déterminés						
<i>majorité</i>	3	3	0	3	---	---
<i>(la) moitié</i>	8	1	0	1	---	---
<i>moyenne</i>	3	0	0	1	---	---
<i>partie</i>	15	1	1	4	---	---
<i>(le) peu</i>	0	0	0	2	---	---
Ncoll formés avec <i>monde</i>						
<i>tout le monde</i>	106	2	---	---	310	88
<i>autres</i>	6	4	---	---	94	425
<b>Total</b>	<b>199</b>	<b>37</b>	<b>2</b>	<b>13</b>	<b>521</b>	<b>741</b>

Tableau 3: Ncoll individuels dans 3 corpus de différentes variétés du français par type d'accord (chiffres absolus)

[21] Ce tableau synoptique indique en premier lieu que les possibilités de comparaison entre le français de la Suisse Romande et le français de France sont très limitées à cause de la taille réduite du corpus conversationnel de Tristram (2014), qui de surcroît contient plus de 50% (18 sur 34) d'occurrences de Ncoll

et n'est donc pas repris. Les Ncoll *minorité*, *nombre* et *série*, inclus dans notre comptage par souci de conformité avec Tristram (2014), n'apparaissent plus dans ce tableau car toutes les occurrences de Tristram (2014) sont à accord ambigu. Pour le français ontarien, on a reproduit les chiffres du corpus partiel des adolescents à partir du tableau 7 de Mougeon & Mougeon (2017a : 151) en les contrôlant, le cas échéant, à l'aide des chiffres plus explicites du tableau 5 dans Mougeon & Mougeon (2017b : 15). Or, il y a de légers écarts entre les deux tableaux bien qu'ils se réfèrent aux mêmes données. Les chiffres provenant du corpus partiel des adultes n'ont malheureusement pas pu être inclus car ce corpus n'a pas été pris en compte dans Mougeon & Mougeon (2017b) et son tableau 7. À cause du nombre d'occurrences très bas pour certains Ncoll notamment dans *OFROM* et dans le corpus de Tristram (2014), on a renoncé, ici comme dans les tableaux suivants, à calculer des chiffres relatifs.

avec accord ambigu. Pour les 15 occurrences restantes, l'auteure constate une nette préférence pour l'accord sémantique au pluriel, qui s'oppose à la tendance que laissent entrevoir les données *OFROM* globalement, mais aussi lorsqu'on ne prend en compte que les Ncoll d'*OFROM* que ce corpus et l'échantillon de Tristram (2014) ont en commun (49 cas d'accord au singulier vs 19 cas d'accord au pluriel). Cette différence est particulièrement saillante pour *partie*, où 4 des 5 occurrences de Tristram (2014) entraînent l'accord au pluriel tandis qu'une seule des 16 occurrences d'*OFROM* montre ce type d'accord et cela à nouveau dans une configuration constructionnelle spécifique sur laquelle je reviendrai infra :

- (8) donc y a euh une *partie* des mots qu'on utilise dans ma famille qui *viennent* du | \_ | du dialecte vaudois (unine15-068, unine15b04m)

Cependant, la faible base empirique interdit des conclusions très poussées. Cette base quantitative est meilleure au moment de comparer les résultats romands avec les données franco-canadiennes de l'Ontario, où l'on peut déceler la même tendance par rapport aux résultats globaux mais aussi en ne tenant compte que des items partagés (31 occurrences sur 200, c'est-à-dire 15,5% avec accord pluriel dans *OFROM* contre 59% d'accord au pluriel chez Mougeon & Mougeon 2017a, 2017b). On peut en conclure que les locuteurs et locutrices romands, en privilégiant l'accord morphologique singulier sauf pour *plupart*, adhèrent plutôt à un usage sanctionné par la norme, alors que les locuteurs et locutrices ontariens s'en éloignent sensiblement<sup>9</sup>.

[22] Dans la suite de cet article, je vais tenter d'obtenir une image plus nuancée de l'emploi des Ncoll dans *OFROM*, en approfondissant certains aspects de leur distribution, entre autres ceux regardant les configurations syntaxiques dans lesquelles apparaissent ces noms.

## 5.2 Les Ncoll formés avec *monde*

[23] Les Ncoll qui font preuve de la plus haute récurrence dans les données *OFROM* sont ceux formés avec le nom *monde* dans son acception à trait [+humain] ; le même constat vaut pour les données des adolescents ontariens analysées par Mougeon & Mougeon (2017b), qui ont saisi séparément la forme définie *le monde*, la forme indéfinie *un monde / du monde* et la forme *tout le monde*, qui – rappelons-le – aurait « résisté à l'accord avec le sens » (Blinkenberg 1950 : 67). Cette affirmation est clairement confirmée par *OFROM* où seulement deux occurrences de *tout le monde* sur 108 sont attestées avec accord au pluriel, dont une (ex. 9) avec accord interpositionnel :

<sup>9</sup> Il convient de rappeler que les témoins de Mougeon & Mougeon (2017a, 2017b) font partie d'un groupe sociolinguistique bien spécifique. Il serait utile d'étudier les Ncoll dans l'usage de locuteurs franco-canadiens non restreints, notamment ceux du Québec.

- (9) ça fait toujours marrer *tout le monde* parce que *ils savent* que mon niveau d'allemand est pourri (unine15-062, unine15a90m)

alors que chez Mougeon & Mougeon (2017b), l'accord au pluriel pour *tout le monde* atteint 22%, dix fois plus que dans *OFROM*. Pour les autres emplois, ces auteurs constatent la répartition suivante : (a) emplois définis : accord au pluriel 79% (N = 286) ; (b) emplois indéfinis : accord au pluriel 86% (N = 234). Dans *OFROM*, ces autres emplois en position argumentale susceptibles de déclencher un accord sont rares, mais il est vrai qu'avec 4 des 10 occurrences, le SV s'accorde au pluriel ; dans tous ces cas, il s'agit du SN *beaucoup de monde* en relation d'accord interpropositionnel, comme dans (10) :

- (10) on a eu | \_ | *beaucoup de monde* qui ont qui sont en tout cas qui ont dit | \_ | euh apprécier euh la pièce (unine09-bfa, unine09a05m)

### 5.3 Les Ncoll suivis de complément et l'accord par proximité

[24] On a vu plus haut que Riegel, Pellat & Rioul (2018 [1994]) et Lammert (2010) considèrent admissible, voire peu surprenant l'accord au pluriel quand le Ncoll est suivi d'un complément, comme dans *Une foule de touristes visiteront Paris*, en alléguant des raisons cognitives ou syntaxiques.

Ncoll	Ncoll seul		Ncoll + Ncomp		
	Accord singulier	Accord pluriel	Accord singulier avec Ncoll	Accord pluriel avec Ncomp	Autres
<i>classe</i>	6	0	1	0	0
<i>entourage</i>	1	0	0	0	1
<i>équipe</i>	1	5	0	2	1
<i>groupe</i>	9	2	0	2	0
<i>majorité</i>	0	1	2	2	1
<i>minorité</i>	1	0	0	0	0
<i>(la) moitié</i>	3	0	0	1	5
<i>moyenne</i>	0	0	0	0	3
<i>nombre</i>	0	0	1	6	0
<i>partie</i>	5	0	1	1	9
<i>plupart</i>	1	16	0	21	0
<i>série</i>	0	0	0	2	0
<i>monde</i> <sup>10</sup>	6	4	0	0	0
Total	33	28	5	37	20

Tableau 4: Ncoll et types + portées d'accord en fonction de la présence d'un complément (chiffres absolus)

<sup>10</sup> Sans *tout le monde*.

Sur le plan syntagmatique, on peut parler d'un accord par proximité (Riegel, Pellat & Rioul 2018 [1994] : 898), phénomène performatif dû à un effet d'amorçage (*priming*). Si l'on passe au crible les Ncoll d'*OFROM* susceptibles d'être suivis par un complément, en vérifiant si le Ncoll apparaît seul ou avec un nom en position de complément (Ncomp) et si, dans ce deuxième cas de figure, l'accord (non ambigu évidemment) s'établit avec le Ncoll ou le Ncomp, on arrive aux résultats présentés dans le tableau 4. Ces chiffres nous apprennent que les noms concernés, appartenant aux Ncoll sous-déterminés et non déterminés, apparaissent dans *OFROM* autant avec que sans complément exprimé. Dans ce deuxième cas de figure, on trouve un certain nombre d'exemples où le Ncomp et l'élément (verbal) en relation d'accord sont tous les deux au singulier ; ces exemples ont été rangés sous *Autres*. Mais cette catégorie comprend aussi deux exemples remarquables : dans (11) le complément du Ncoll est lui-même un nom collectif, du type déterminé, et l'accord se fait au pluriel (le trait du nombre y est identifiable grâce au pronom *tous*, l'accord verbal étant indécidable à cause de l'ambiguïté de la forme verbale à l'oral) :

- (11) des propositions qui ont été faites par ma par euh l'*entourage* de ma *famille* qui me *disaient tous* | \_ | il faut que tu changes de j/ de boulot (unine08-jga, unine08a21m)

alors que dans (12), avec un Ncomp au pluriel, la première forme verbale qui suit le Ncoll est au pluriel tandis que la deuxième porte l'accord au singulier :

- (12) euh on cons/ on pense qu'à la fin du siècle y y aura la *moitié* des *langues* qui *sont* aujourd'hui parlées | \_ | qui *aura* disparu (unifr14-cba, unifr14a01m)

À travers cet accord variable dans (12) le locuteur est à même d'explicitier les relations entre les éléments du SN et les verbes correspondants (*les langues [...] sont parlées vs la moitié [...] aura disparu*) ; il vaudrait donc mieux parler d'accord différencié<sup>11</sup>.

[25] Au vu du total indiqué dans la dernière ligne du tableau 4, on trouve confirmé le fait que la présence d'un Ncomp pluriel encourage l'accord au pluriel. Si on déduit de ces chiffres les occurrences de *plupart*, où l'accord au pluriel est pratiquement grammaticalisé, la tendance évoquée se décèle encore plus nettement : l'accord au pluriel se fait alors dans 16 sur 21 cas (76%) avec les Ncoll assortis de complément exprimé contre 12 sur 44 cas (27%) lorsque le Ncoll n'est pas suivi d'un complément. Ceci dit, il convient d'admettre que la limitation – délibérée pour des raisons méthodologiques – aux seules occurrences où le Ncomp suit le Ncoll appauvrit jusqu'à un certain point l'image des SN collectifs à complé-

11 Je remercie l'évaluateur anonyme pour cette observation.



ment. L'ex. (13) illustre ce point (le signe @ marque une intervention de l'enquêteur non transcrite, alors que *natel* est un mot suisse romand pour désigner le téléphone portable) :

- (13) alors déjà en troisième primaire la *plupart* en ont | @ | ouais en tout cas ben ceux de Zermatt la *plupart* avaient un natel (unine11-csa, unine11c02m)

Cet extrait d'*OFROM* contient deux occurrences du Ncoll quantificateur *plupart* (avec accord pluriel canonique), les deux non suivis de complément. Alors que pour la première occurrence, il apparaît qu'il faut inférer le nom sur lequel *plupart* quantifie – inférence facile vu le contexte scolaire dans lequel se meut le récit –, pour le second *plupart* ce complément (*ceux de Zermatt*) se trouve en fait verbalisé dans le cotexte *antérieur* immédiat, comme c'est également le cas dans (14) :

- (14) le problème avec les *chevaux* la *plupart* | \_ | ils ils expriment pas très fort les les douleurs au début (unine08-fja, unine08a24m)

#### 5.4 Les Ncoll en position disloquée

[26] L'ex. (14) nous servira de point d'appui pour mettre en évidence la nécessité de ne pas limiter l'analyse de la variation au seul accord verbal mais d'inclure l'accord pronominal, surtout quand on se penche sur l'oral. Il est vrai que des phrases 'prototypiques' (en termes de grammaire traditionnelle), avec accord intrapositionnel, comme (15) y sont bien attestées :

- (15) et pis la *plupart* des gens *veulent* pas (unine08-fja, unine15a83d)

mais une partie importante des exemples des Ncoll apparaissent dans des constructions syntaxiques plus complexes. Dans beaucoup de cas, il s'agit d'une construction à extraposition à l'instar de (4), (6), (14) ou (16-17) où le nom collectif est placé en périphérie de la proposition qui contient le SV, soit par dislocation à gauche :

- (16) le l/ l'*équipe* des squatteurs je *les* rencontre encore maintenant et *ils* étaient heureux parce que je *leur* maintenais | \_ | le courant (unine11-jfz, unine11d10m)

ou, plus exceptionnellement, par dislocation à droite :

- (17) mais alors eh je te te | \_ | je te mens pas tu peux *leur* demander l'*équipe* | \_ | je crois qu'*ils* ont jamais autant ri de *leur* vie (unine15-934, unine15z59d)

et avec reprise pronominale au sein de la proposition contenant le SV. Les 'questions d'accord' se posent alors non pas pour l'accord verbal mais par rapport à la relation entre le Ncoll extraposé et le pronom personnel de relais, sans qu'il soit impératif que le Ncoll se trouve en position sujet, comme le montrent les ex. (5), (16) et (17). Ce pronom peut être de type défini, comme dans les exemples qu'on vient d'énumérer, mais si la pronominalisation porte sur la position sujet, on trouve aussi le pronom indéfini *ce/ça*, comme dans (7) ou (18) :

- (18) je suis plus à l'aise avec l'accent américain mais la *majorité* par exemple des séries télé | moi j'adore les séries télé euh | \_ | c'est en américain donc ça va (unine15-104, unine15b56m)

[27] Le tableau 5 donne un aperçu des Ncoll dans *OFROM* pour lesquels des emplois en dislocation avec reprise pronominale ont été identifiés :

Ncoll	Position périphérique avec reprise par			Total
	pronom personnel défini	pronom personnel indéfini	autre	
<i>école</i>	7	10	0	17
<i>famille</i>	11	5	0	16
<i>université</i>	1	3	0	4
<i>classe</i>	2	3	0	5
<i>équipe</i>	6	0	0	6
<i>groupe</i>	5	1	0	6
<i>majorité</i>	2	2	0	4
<i>(la) moitié</i>	1	1	0	2
<i>nombre</i>	0	1	0	1
<i>partie</i>	0	1	0	1
<i>plupart</i>	15	3	1	19
<i>tout le monde</i>	3	0	0	3
Total	53	30	1	84

Tableau 5: Ncoll disloqués et types de reprise pronominale dans *OFROM* (chiffres absolus)

Les chiffres illustrent la récurrence avec laquelle les Ncoll d'*OFROM* sont intégrés dans des constructions à extraposition, mais on n'est pas en mesure de proposer une explication de la raison pour laquelle certains Ncoll sont plus enclins à apparaître dans ce type de constellation syntaxique que d'autres. Si l'on compare la distribution des types de reprise pronominale, on constate – sans surprise – que le pronom personnel indéfini *ce/ça* apparaît de manière privilégiée avec des Ncoll déterminés qui ne portent le trait sémantique [+humain] que par métonymie ; cela concerne avant tout *école* et, dans une moindre mesure, *université* (et *classe*, qui dans notre échantillon a cependant été rangé du côté des Ncoll sous-déterminés).

C'est d'ailleurs pour *classe* qu'*OFROM* propose un exemple remarquable, où le nom disloqué est repris deux fois, d'abord par un pronom personnel au singulier et ensuite par un pronom au pluriel (avec les accords verbaux correspondants) :

- (19) si tu as | \_ | deux classes une classe de forte une classe de moins forte | \_ | la classe de forte *elle* va être boostée et *ils* vont bosser à fond (unine14-pra, unine14a06m)

La seule occurrence qu'il fallait ranger dans la catégorie *Autres* est celle de l'ex. (20), où le Ncoll n'est pas repris par un pronom mais par le SN *les gens* ; or, Cappeau & Schnedecker (2014) montrent que *les gens* est en voie de grammaticalisation vers un statut quasi-pronominal :

- (20) la *plupart* des | \_ | des morts sur la route s/ euh *les gens* ont entre dix-huit et vingt-cinq ans (unine08-ada, unine08a15m)

[28] Ayant relevé les cas d'extraposition avec les Ncoll, il se pose la question d'une relation éventuelle entre ce phénomène syntaxique et l'accord variable. La dislocation, en positionnant l'élément disloqué en périphérie, le retire de la proposition et de sa structure argumentale. Il y a donc une frontière syntaxique – faible, certes, car non marquée d'un morphème jonctif – entre le Ncoll et la proposition qui contient la/les cible(s) de l'accord. En plus, il peut y avoir une distance syntagmatique plus ou moins grande entre le Ncoll et la/les cible(s) en fonction de la structure syntaxique de la proposition. L'éloignement du Ncoll et la présence de frontières syntaxiques se sont avérés être des facteurs qui affaiblissent l'accord morphologique et favorisent l'accord sémantique (Levin 2001 : 92-102). Pour vérifier si un tel effet existe dans les données *OFROM*, il faut comparer les chiffres des tableaux 5 et 1. Or, les résultats sont mitigés : avec *famille*, *équipe*, *groupe* et *majorité* on trouve des Ncoll de tous les types (sauf ceux formés sur *monde*) où une corrélation entre la fréquence d'apparition en position périphérique et l'accord au pluriel semble exister. Le Ncoll *équipe* montre cet effet de la façon la plus palpable : 50% (6 sur 12) de ses occurrences se trouvent disloqués et pour 66% des occurrences on note un accord au pluriel. D'autre part, il y a des items qui ne confirment pas du tout un tel effet, comme les Ncoll dérivés d'institutions (*école* et *université*) ou le Ncoll non déterminé *nombre* qui malgré un accord sémantique avec 60% (6 sur 10) de ses occurrences n'apparaît en extraposition qu'une seule fois. Les faibles corrélations associées aux nombres d'occurrences modestes pour certains Ncoll concernés ne permettent donc pas de dire avec certitude si une relation entre l'extraposition et l'accord variable existe.

[29] Ce que le tableau 5 démontre avec certitude, cependant, c'est qu'il est absolument nécessaire, pour saisir le phénomène de l'accord variable avec les Ncoll dans une étude sur corpus, de prendre en compte l'accord pronominal. Mais

comme les chaînes anaphoriques peuvent être longues et que la distance syntagmatique entre un Ncoll et son pronom de relais est parfois grande, il va sans dire qu'il existe un risque considérable que des relations d'accord restent inaperçues et que l'objectif d'une exhaustivité dans l'analyse s'avère vain<sup>12</sup>.

### 5.5 Les Ncoll dans les clivées et dans d'autres constructions à relative

[30] Un dernier aspect de la distribution des Ncoll dans *OFROM* qu'il reste à approfondir concerne, lui aussi, l'environnement syntaxique dans lequel ces noms apparaissent. Il s'agit d'exemples comme (8), (12) ou les suivants (à noter, dans (21), l'accord grammatical au singulier suivi d'un accord sémantique au pluriel et, dans (22), avec un accord verbal en nombre indécidable, l'accord en genre de l'attribut on ne peut plus sémantique) :

- (21) tout est bien y a *tout le monde* qui est | \_ | super respectueux | \_ | enfin ils ont ouais enfin ils sont respectueux entre eux et pis | \_ | je trouve que ça change par rapport à ici (unine15-111, unine15b65m)
- (22) tout d'un coup y a *tout le monde* qui tombait enceinte (unine16-009, unine16a15d)

Dans tous ces exemples, la relation d'accord s'établit au sein d'une construction biclausale du type clivé avec un premier segment en *il y a* et un deuxième segment sous forme de relative, construction plurifonctionnelle mais connue surtout dans sa fonction présentative (Karssenbergh 2018). Tristram (2014 : 79) remarque, à propos de la portée de l'accord des Ncoll dans son corpus : « Another striking aspect of the interview data is the number of tokens where the target [de l'accord ; C.P.] occurs in a subordinate clause after some form of the presentative *il y a* [...] ». Les données d'*OFROM* pointent dans le même sens, mais il faut souligner qu'à côté des clivées en *il y a* – effectivement les plus nombreuses – on trouve d'autres constructions présentatives biclausales, comme les clivées en *c'est* (23-24, ce dernier ex. avec effacement du morphème de subordination de la [pseudo-]relative) ou à segment présentatif elliptique (25 ; l'ellipse est marquée par le signe Ø) :

- (23) mais après *c'est c'est c'est un groupe* qui sont tout le temps tout le temps tout le temps tout le temps ensemble (unine12-adb, unine12a06m)
- (24) *c'est une famille* ils ont trois enfants (unine16-001, unine16a01m)
- (25) parce qu'en plus ben comme *c'est le dimanche soir* | \_ | Ø jamais *tout le monde* qu'est là (unine18-001, unine18a03d)

12 D'autre part, comme le rappelle un évaluateur anonyme à propos de l'exemple (9), il peut y avoir un autre nom dans un contexte gauche plus large auquel le pronom de relais pourrait être lié. Si ce n'est pas le cas pour l'exemple en question, cette possibilité est tout à fait envisageable pour d'autres exemples du corpus de travail, ce qui rend une analyse exhaustive encore moins aisée.

ou encore d'autres constructions de segmentation propositionnelle faisant appel à des relatives. Parmi les Ncoll d'*OFROM*, un total de 137 occurrences se trouve dans de tels environnements syntaxiques, distribué comme suit :

Ncoll	Constructions biclauses présentatives				Autres constructions relatives	Total
	en c'est + Ncoll + relative	en il y a + Ncoll + relative	en Ø + Ncoll + relative	autre + Ncoll + relative		
<i>école</i>	4	0	0	0	0	4
<i>famille</i>	5	5	7	3	7	27
<i>police</i>	0	1	0	0	0	1
<i>université</i>	1	0	1	0	0	2
<i>classe</i>	2	1	1	1	2	7
<i>entourage</i>	0	0	0	0	1	1
<i>équipe</i>	3	0	0	1	0	4
<i>groupe</i>	5	2	0	0	9	16
<i>minorité</i>	1	0	0	0	0	1
<i>(la) moitié</i>	1	4	0	1	0	6
<i>nombre</i>	0	1	0	1	7	9
<i>partie</i>	0	7	1	2	0	10
<i>plupart</i>	1	2	1	0	6	10
<i>série</i>	2	0	0	0	2	4
<i>tout le monde</i>	0	18	2	0	0	20
autre sur monde	0	11	0	1	3	15
Total	25	52	13	10	37	137

Tableau 6: Ncoll dans des constructions syntaxiques à subordination relative (chiffres absolus)

[31] En premier lieu, il vaut la peine de souligner que 32% des Ncoll attestés dans *OFROM* apparaissent dans des constructions à relatives, dont la plus fréquente est le clivage qui concerne presque un quart des occurrences du corpus. On a vu supra, lors de la discussion des Ncoll extraposés, que la dislocation introduit une frontière syntaxique (faible) entre le Ncoll et les éléments concordants tout en les éloignant sur le plan syntagmatique ; la même observation vaut pour les constructions impliquant la relativisation, à la différence près que la frontière syntaxique introduite est forte car marquée par un morphème jonctif démarcatif (le pronom relatif / la particule relative). Aussi la question d'une relation éventuelle entre l'intégration des Ncoll dans les constructions relatives et l'accord variable se pose-t-elle à nouveau. L'image qui émerge de la comparaison des tableaux 6 et 1 est comparable à celle qui résultait de l'analyse de l'extraposition : les données *OFROM* confirment en partie une telle relation, mais ne sont pas probantes en raison des divergences considérables entre les Ncoll individuels et le nombre parfois

très réduit de leurs occurrences. Avec *famille* et *groupe*, on a deux items pour lesquels l'extraposition comme l'intégration dans des constructions à relatives semblent avoir un effet favorisant l'accord sémantique. Hormis *plupart* qui, comme on a vu, a pratiquement grammaticalisé l'accord au pluriel, c'est avec *nombre* (qui semblait insensible aux effets de l'extraposition) et les Ncoll formés avec *monde* (sauf *tout le monde*) que la corrélation sous examen est la plus prononcée<sup>13</sup>. D'autre part, on a *classe* et *partie* dont plus de la moitié des occurrences est intégrée dans des constructions à relative mais qui ne montrent qu'une très faible tendance vers l'accord au pluriel. On peut donc conclure que l'effet d'éloignement, selon lequel le placement du Ncoll en périphérie de la proposition, dans le segment présentatif d'une construction biclausale ou une autre construction relative, favorise l'accord sémantique, effet que Mougeon & Mougeon (2017a : 148, 2017b : 11) voient confirmé par leurs corpus, et l'effet (souvent concomitant) de l'introduction d'une frontière syntaxique entre le Ncoll et l'élément en relation d'accord, ne sont pas contredits par les données d'*OFROM* ; d'autre part, le comportement très divergent des différents Ncoll de l'échantillon par rapport à ces facteurs ne permet pas non plus de parler d'une tendance nette.

## 6 Conclusion

[32] L'objectif de cet article était de dresser un panorama descriptif des Ncoll dans les données d'*OFROM* et de mettre en évidence leur comportement par rapport à l'accord morphologique verbal et pronominal qui, pour ce groupe de mots, est connu comme étant variable, notamment à l'oral. La base de l'analyse est un corpus de travail comprenant 21 Ncoll pour lesquels le dépouillement d'*OFROM* a fourni un total de 432 occurrences. L'analyse a montré que les locuteurs et locutrices romands se conforment *grosso modo* à la norme : d'une façon globale, ils privilégient clairement l'accord grammatical au singulier pour tous les Ncoll examinés sauf *plupart*, qu'ils emploient majoritairement avec un accord pluriel, usage lui aussi en accord avec la norme. Cependant, pour une partie non négligeable des occurrences (13,5% à 26,5% selon le type de Ncoll), on trouve quand même un accord divergent qui, pour une partie des Ncoll, s'explique par la présence d'un complément pluriel après le Ncoll, menant à une forme verbale au pluriel. Si l'influence de cet effet d'accord par proximité est probante, les résultats sont moins nets quant à l'effet de l'éloignement entre le Ncoll et les éléments verbaux ou pronominaux avec lesquels il est censé être en relation d'accord. Cet éloignement syntagmatique peut s'opérer par une dislocation du Ncoll vers la périphérie de l'énoncé (19,5% des occurrences des Ncoll étudiés d'*OFROM* se rencontrent dans une telle construction) ou par des constructions biclausales à relative notam-

13 Cela vaudrait aussi pour *série* dont toutes les occurrences apparaissent intégrées dans des constructions à relative et dont le taux d'accord pluriel est de 50% – hélas, *OFROM* ne fournit que 4 exemples de ce Ncoll.

ment du type clivé (ce qui concerne une proportion remarquable de 32% des Ncoll analysés). Pour certains items du corpus de travail, cet éloignement entre le Ncoll et l'élément semble favoriser l'accord sémantique, pour d'autres cet effet s'avère inefficace.

[33] Parallèlement à la description des Ncoll en français suisse romand, j'ai entamé une comparaison avec deux autres variétés de la langue pour lesquelles des analyses sur corpus sont disponibles. Malheureusement, le corpus conversationnel de Tristram (2014), de par sa faible taille, ne se prête pas vraiment à une lecture comparative. Par contre, les travaux de Mougeon & Mougeon (2017a, 2017b), basés sur un corpus de travail quantitativement impressionnant, se sont avérés propices à une comparaison, laquelle a fait ressortir de fortes différences dans l'accord avec les Ncoll entre les francophones suisses et les francophones ontariens. Il serait intéressant de répliquer notre analyse sur les Ncoll avec un corpus conversationnel du français de France plus grand et un corpus du français laurentien du Québec afin d'obtenir une image plus complète et plus fiable de la variation dans ce domaine de la langue en francophonie.

## Abréviations

CN = Count nouns

NCN = Non-count nouns

Ncoll = Noms collectifs

Ncomp = Nom en position de complément

## Bibliographie

- Avanzi, Mathieu, Marie-José Béguelin, Federica Diémoz 2012-2019. *Présentation du corpus OFROM – corpus oral de français de Suisse romande*. Neuchâtel : Université de Neuchâtel. <http://www.unine.ch/ofrom>.
- Berrendonner, Alain, Marie-José Reichler-Béguelin 1995. Accords 'associatifs'. *Cahiers de praxématique* 24, 21-43. <https://journals.openedition.org/praxématique/3025>.
- Blinkenberg, Andreas 1950. Le problème de l'accord en français moderne. Essai d'une typologie. København : Munksgaard.
- Bußmann, Hadumod 2002. *Lexikon der Sprachwissenschaft*. 3e édition. Stuttgart : Kröner.
- Cabezas Holgado, Emilio 2017. *Las expresiones colectivas en español*. Madrid : Arco Libros.
- Cappeau, Paul, Catherine Schnedecker 2014. Des *gens* bien différents à l'écrit et à l'oral. Hypothèses sur la pronominalisation des SN en *gens*. *Verbum* 26, 55-74.
- Corbett, Greville G. 2006. *Agreement*. Cambridge : Cambridge University Press.
- Flaux, Nelly 1999. À propos des noms collectifs. *Revue de Linguistique romane* 63, 471-502.
- Frantext = Laboratoire Analyse et traitement informatique de la langue française (ATILF) (éd.) 1998-2019. *Base textuelle Frantext*. <http://www.frantext.fr>.
- Høybye, Poul 1944. L'accord en français contemporain. Essai de grammaire descriptive. København : Høst.
- Joosten, Frank 2006. Why *club* and *lingerie* do not belong together. A plea for redefining collective nouns. Georges Kleiber, Catherine Schnedecker, Anne Theissen (éds.). *La relation partie-tout*. Leuven : Peeters, 73-88.
- Karssenber, Lena 2018. *Non-prototypical clefts in French. A corpus analysis of il y a clefts*. Berlin : De Gruyter.
- Kleiber, Georges, Catherine Schnedecker, Anne Theissen (éds.) 2006. *La relation partie-tout*. Leuven : Peeters.
- Koch, Peter, Wulf Oesterreicher 1990. *Gesprochene Sprache in der Romania: Französisch, Italienisch, Spanisch*. Tübingen: Niemeyer.
- Lammert, Marie 2010. *Sémantique et cognition. Les noms collectifs*. Genève : Droz.
- Lammert, Marie, Michelle Lecolle 2014. Les noms collectifs en français : une vue d'ensemble. *Cahiers de lexicologie* 105, 203-222.
- Laycock, Henry 2006. Mass nouns, count nouns, and non-count nouns: philosophical aspects. Keith Brown (éd.). *Encyclopedia of Language and Linguistics. Vol. 7*. 2e édition. Amsterdam : Elsevier, 534-538.
- Lecolle, Michelle 2019. *Les noms collectifs humains en français. Enjeux sémantiques, lexicaux et discursifs*. Limoges : Lambert-Lucas.
- Levin, Magnus 2001. *Agreement with collective nouns in English*. Thèse de doctorat, Lunds Universitet.
- Michaux, Christine 1992. The collectives in French: A sociolinguistic investigation. *Linguisticae Investigationes* 16, 99-124.
- Mougeon, Françoise, Raymond Mougeon 2017a. L'accord verbal en nombre avec les collectifs dans le français parlé en Ontario. Henry Tyne et al. (éds.). *La variation en question(s)*.



- Hommages à Françoise Gadet*. Bruxelles : Lang, 137-156.
- Mougeon, Françoise, Raymond Mougeon 2017b. Accord verbal de nombre dans le français parlé en Ontario. *Le français moderne* 2017, 4-21.
- OFROM = Mathieu Avanzi, Marie-José Béguelin, Federica Diémoz (éds.) 2012-2019. *Corpus oral de français de Suisse romande*. <http://www.unine.ch/ofrom>.
- Riegel, Martin, Jean-Christophe Pellat, René Rioul 2018 [1994]. *Grammaire méthodique du français*. 7e édition. Paris : Presses universitaires de France.
- Tristram, Anna 2010. L'accord sujet-verbe en français contemporain : une étude de variation socio-linguistique. Franck Neveu et al. (éds.). *CMLF 2010 - 2ème Congrès mondial de Linguistique française, 1987-1999*. [https://www.linguistiquefrancaise.org/articles/cmlf/abs/2010/01/cmlf2010\\_000115/cmlf2010\\_000115.html](https://www.linguistiquefrancaise.org/articles/cmlf/abs/2010/01/cmlf2010_000115/cmlf2010_000115.html).
- Tristram, Anna 2014. *Variation and change in French morphosyntax. The case of collective nouns*. London: Modern Humanities Research Association, Routledge.
- Tristram, Anna, Wendy Ayres-Bennett 2012. From negation to agreement: Revisiting the problem of sources for socio-historical linguistics. *Neuphilologische Mitteilungen* 113, 365-393. <https://www.jstor.org/stable/43344667>.