

Faits déviants et tri des observables

Marie-José Reichler-Béguelin
Universités de Neuchâtel et de Fribourg

La linguistique actuelle entretient une attitude ambiguë par rapport aux écarts langagiers.

D'un côté, elle se conçoit comme non normative, ce qui semblerait indiquer que les faits déviants, fautes, particularismes ou anomalies, font *ipso facto* partie des observables, du matériau à modéliser. Rompant avec une tradition prescriptive longtemps dominante en grammaire, cette visée descriptive se traduit dans des déclarations de principe bien connues:

"La linguistique est l'étude scientifique du langage humain. Une étude est dite scientifique lorsqu'elle se fonde sur l'observation des faits et s'abstient de proposer un choix parmi ces faits au nom de certains principes esthétiques ou moraux. "Scientifique" s'oppose donc à "prescriptif". Dans le cas de la linguistique, il est particulièrement important d'insister sur le caractère scientifique et non prescriptif de l'étude: l'objet de cette science étant une activité humaine, la tentation est grande de quitter le domaine de l'observation impartiale pour recommander un certain comportement, de ne plus noter ce qu'on dit réellement, mais d'édicter ce qu'il faut dire. (...) le linguiste contemporain, en face de la *lettre que j'ai écrit, occasion à profiter, la femme que je lui ai parlé*, se refuse aussi bien à la vertueuse indignation du puriste qu'à l'exultation de l'iconoclaste. Il voit là simplement des faits qu'il lui faut noter et expliquer dans le cadre des usages où ils apparaissent. Il ne sortira pas de son rôle s'il relève les protestations ou les railleries de certains auditeurs et l'indifférence des autres; mais il s'abstiendra, pour sa part, de prendre parti." (A. Martinet, *Eléments de linguistique générale*, Paris, 1960: 6-7)

Mais d'un autre côté, quand même elle se veut impartiale, il arrive que la linguistique réintroduise subrepticement, dans sa pratique, la notion de faute de langue, et surtout la démarche d'exclusion qui lui est as-

sociée¹. La mise à l'écart de certaines productions langagières emprunte au moins deux voies.

D'abord, elle peut consister à disqualifier globalement, de manière plus ou moins affichée, le parler de toute une catégorie de locuteurs. Dans son compte rendu de la *Grammaire des fautes* d'Henri Frei, le célèbre comparatiste Antoine Meillet s'est ainsi fait, sans détour, le porte-parole d'une conception élitiste de la "langue":

"....des lettres d'illettrés adressées à des illettrés ne fournissent au linguiste que des données médiocres"... (BSL 30, 1929-30: 147)²

Une telle évaluation a de quoi surprendre sous la plume d'un diachronicien, habitué par métier à exploiter des documents langagiers de sources extrêmement diverses, qui sont loin d'être toujours attribuables à des "lettrés". Dans un de ses ouvrages les plus connus, Meillet avait pourtant lui-même appelé de ses vœux une meilleure connaissance des langues vivantes et de leurs conditions d'utilisation par les locuteurs; il y voyait le seul moyen de conférer une validation scientifique aux reconstructions opérées par les comparatistes:

"Ce qui importe au linguiste, c'est de connaître comment les gens qui parlent français se comportent vis-à-vis des règles. Or, on n'a là-dessus que des idées vagues; il n'a été fait aucune enquête méthodique, à peine quelques sondages partiels." (A. Meillet, *La méthode comparative en linguistique historique*, Oslo et Paris, 1925: 112)

Meillet ne pouvait en effet pas ignorer que toute réflexion à caractère historique sur le statut de la "faute de langue" aboutit, tôt ou tard, à observer que la norme d'aujourd'hui est largement constituée par des inflexions aux règles d'hier.

¹ Cf. Berrendonner, 1982.

² Linguiste genevois, élève de Charles Bally, Henri Frei s'était donné pour but d'observer et d'expliquer sur des critères fonctionnels les tendances grammaticales à l'oeuvre en français populaire. Il exploitait notamment, dans son livre, les lettres de soldats collectées dès 1914 par *L'Agence des Prisonniers de Guerre* rattachée au Comité International de la Croix-Rouge. Je remercie Annette Fryba d'avoir porté à ma connaissance, il y a un certain temps déjà, le compte rendu de Meillet.

Comment expliquer, dès lors, la sévérité de Meillet à l'égard de la démarche de Frei? Sans doute existe-t-il chez tout locuteur une ambiguïté à laquelle n'échappe vraisemblablement pas le linguiste lui-même: face au parler d'autrui, il n'est pas toujours facile d'accommoder certains réflexes épidermiques de censure avec une réflexion approfondie sur la "valeur" des variantes langagières, considérée indépendamment de leur degré de prestige social. Le passage de Proust reproduit ci-dessous exprime à merveille ce dilemme interne qui consiste à censurer occasionnellement le langage de l'autre, en vertu de dissensions socio-culturelles ou interpersonnelles, tout en admettant, à tête reposée, la relativité de la notion même de "faute" de langue:

<Le narrateur, contrarié pour d'autres raisons, vient de reprocher à Françoise ses cuirs et sa mauvaise façon de prononcer le français.>

"Et ce reproche était particulièrement stupide, car ces mots français que nous sommes si fiers de prononcer exactement ne sont eux-mêmes que des "cuirs" faits par des bouches gauloises qui prononçaient de travers le latin ou le saxon, notre langue n'étant que la prononciation défectueuse de quelques autres. Le génie linguistique à l'état vivant, l'avenir et le passé du français, voilà ce qui eût dû m'intéresser dans les fautes de Françoise. L'"estoppeuse" pour la "stoppeuse" n'était-il pas aussi curieux que ces animaux survivants des époques lointaines, comme la baleine ou la girafe, et qui nous montrent les états que la vie animale a traversés?" (Marcel Proust, *Sodome et Gomorrhe*, 736-7)

Une autre manière plus subtile et aussi plus répandue de proscrire les faits déviants se concrétise, en linguistique, dans la pratique du jugement d'acceptabilité, qui autorise à trier entre données "pertinentes" et données "non pertinentes" sur une base qui, bien souvent, reste purement intuitive. Le problème, insidieux, naît chaque fois qu'est déclarée mal formée, et donc indigne d'être prise en compte dans la description, telle ou telle structure pourtant utilisée, parfois couramment, par les sujets parlants. Le procédé a ceci de dangereux qu'il aboutit à confondre sous le même astérisque disqualificateur certains artefacts proprement irréalistes, n'ayant aucune chance d'être un jour performés, avec des énoncés qui sont au

contraire accessibles à l'observation empirique, même si leur forme ne correspond pas en tout point à la norme dominante.

Sans pouvoir mener, dans le cadre de cette étude, une réflexion de fond sur le statut de ce que la linguistique appelle "accident de performance", je voudrais tout de même poser, à ce propos, deux questions qui ont une portée à la fois méthodologique et épistémologique:

1. Comment s'y prennent les linguistes pour poser des jugements d'acceptabilité, comment utilisent-ils ces jugements?
2. Quelles sont les conséquences qui découlent de leur pratique pour la théorie linguistique elle-même?

Le domaine de faits qui sera envisagé ici semble particulièrement favorable à une réflexion méta-théorique. Il concerne la linguistique du discours, c'est-à-dire l'étude des contraintes qui régissent l'agencement des énonciations à l'échelle "transphrastique", ou macro-syntaxique. Je m'intéresserai plus particulièrement au problème, amplement débattu ces dernières années, de la construction d'une grammaire (ou, plus modestement, d'un mode d'emploi) pour les expressions référentielles, ou *pointeurs*. La notion de *pointeur* s'applique à des expressions dont le signifié véhicule un présupposé existentiel: le référent qu'elles désignent est considéré comme "déjà présent", ou déjà validé³, dans les représentations communes des partenaires de l'interlocution. Répondent notamment à cette définition les syntagmes nominaux définis et les pronoms non liés⁴. Je m'interrogerai sur le problème des critères d'acceptabilité de ce qu'il est convenu d'appeler les anaphores pronominales et lexicales, sur l'usage qui est fait, quand on les étudie, des exemples forgés et des jugements de grammaticalité.

³ Cette validation peut avoir lieu en gros de trois manières: soit situationnellement (en cas d'usage déictique des pointeurs), soit linguistiquement, quand le pointeur rappelle un référent précédemment introduit par des moyens langagiers, soit sur une base logique, quand il rappelle un référent inférable contextuellement.

⁴ Les pronoms non liés sont ceux dont l'apparition n'est pas structurellement contrainte, et qui peuvent, à conditions référentielles égales, commuter avec des SN définis: cf. Berrendonner et Reichler-Béguelin, 1989; Berrendonner *et alii*, à paraître.

1. Enchaînements discursifs et critères d'acceptabilité: le cas des anaphores démonstratives

1.1. Inconvénients du travail sur des séquences forgées

La sémantique référentielle a emprunté à la linguistique "phrastique" la méthode consistant à travailler, de manière privilégiée, sur des exemples fabriqués, volontairement simplifiés et raccourcis. Pour évaluer ces séquences, elle recourt à l'opposition *grammatical* vs *agrammatical*, ou encore *correct* vs *déviant*, héritée des pratiques d'analyse qui ont traditionnellement cours en morpho-syntaxe. Appliquée à l'emploi des pointeurs, cette procédure de description permet de neutraliser des paramètres contextuels considérés comme perturbateurs. Elle a toutefois le défaut de privilégier les cas d'enchaînement les plus typiques⁵, et de ne travailler, en principe, que sur des séquences brèves, souvent des couples de phrases où l'introducteur et le pointeur occupent des positions argumentales et thématiques extrêmement conventionnelles (cf. *infra* les ex. donnés dans l'extrait 1).

A force de neutraliser les paramètres énonciatifs et contextuels, l'approche classique accrédite aussi une conception étroitement textualiste et segmentaliste des phénomènes anaphoriques, fréquemment conçus, dans la littérature sur la question, comme reflétant des relations linguistiquement réglées entre deux fragments de discours clairement identifiés et autonomisables (antécédent ou forme-source actualisé dans la chaîne, ou encore introducteur, et forme de reprise). Or, une telle façon de voir rencontre assez rapidement ses limites, dès que l'on tente d'appliquer ses résultats aux productions linguistiques attestées.

Je vais tenter de montrer les difficultés à laquelle se heurte cette conception à propos d'un exemple précis: celui des règles d'emploi les plus couramment formulées à propos du SN démonstratif anaphorique.

Dans les travaux qui approchent l'anaphore démonstrative en se servant du critère de grammaticalité, l'abstraction du contexte à laquelle il vient d'être fait allusion transparait de manière caractéristique dans le fait

⁵ Ainsi que le fait observer Schnedecker, 1992.

que les jugements d'acceptabilité sont souvent formulés à propos d'énoncés *tronqués* ou *incomplets*. L'extrait suivant, tiré d'un ouvrage à finalité didactique, est représentatif d'habitudes qui ont également cours en recherche fondamentale:

1. "Pour se rendre compte des contraintes liées à l'emploi de telles reprises <i.e. les anaphores lexicales démonstratives>, il suffit d'opposer, aux enchaînements corrects de A, ceux de B qui, eux, semblent inacceptables:

A Il acheta un château. Cette propriété...
Elle rencontra une vache. Cet animal...

B Il acheta une propriété. Ce château...
Elle rencontra un animal. Cette vache..."

(E. Genevay et alii, *Notes méthodologiques de grammaire 8e*, Lausanne, Editions L.E.P. Loisirs et pédagogie S.A., 1986, p. 34)

La règle d'enchaînement suggérée ci-dessus voudrait que l'anaphore démonstrative standard ne fonctionne que par reprise d'implications lexicales ou de présupposés déjà contenus dans un "antécédent" linguistique, ou encore que la tête lexicale du SN démonstratif n'apporte aucune information nouvelle par rapport à ce qui a déjà été prédiqué du référent. En conséquence, cette description grammaticale confère un statut normatif privilégié aux cas d'anaphore "fidèle", ou "par hyperonyme" ou nom générique (séquences types: *un soldat... ce soldat*, ou *un soldat... cet homme*). Elle s'accorde, en cela, aux travaux théoriques qui invoquent la nécessité d'une relation de type "être-X", ou encore "être un N ou du N (ou N)", entre un segment source et le substantif N de la description démonstrative: l'existence d'une telle relation permettrait en effet que la forme de l'anaphore démonstrative soit, autant que faire se peut, légitimée par anticipation⁶. Dans une telle optique, l'opération anaphorique se réduirait bel et bien à une relation de chaînage entre deux segments de texte isolables de leur contexte d'occurrence, et les contraintes pesant sur

le contenu lexical du SN démonstratif pourraient être formulées en termes de contrôle sémantique exercé par le premier sur le second.

1.2. Rôle des facteurs contextuels dans le choix du N-tête du SN démonstratif: deux situations concrètes

Or, si l'on se limite aux relations lexicales d'hyponymie / hyponymie mises en jeu dans les exemples invoqués, et sans même faire appel aux SN démonstratifs comprenant des expansions à caractère pictif (étudiées par D. Apothéloz, 1993), il est relativement facile de trouver dans la littérature des séquences, parfaitement acceptables, correspondant au type B ci-dessus, où le contenu lexical du SN démonstratif est à la fois *plus informatif* que l'introducteur verbal qui précède, et *non prédictible* à partir de celui-ci (mon exemple 3 mériterait tout à fait le qualificatif de reprise par hyponyme):

2. Ils <des trous> contenaient chacun une pierre ronde, obscure, et qui paraissait très lourde. Les gens d'un esprit supérieur, seuls, honoraient *ces abbadirs tombés du ciel*. (Flaubert, *Salammbô*, p. 121)
3. C'est ainsi qu'au Rivier (domaine de la présidente) de magnifiques oiseaux blancs ont pu être observés pendant 15 jours. Ils étaient très grands, nombreux, dotés d'un long et gros bec jaune-rouge qu'ils glissaient dans l'eau pour pêcher, en avançant à pied dans la vasière. *Ces "spatules"* ont pu être vues par tous ceux qui ont emprunté les promenades piétonnières organisées pour le public; ce sont des espèces rares sur nos côtes. (*Bulletin de l'Association des Amis des Sites de la région de Mesquer*, juin 1989)

De tels exemples, il faut le souligner, invitent à s'interroger sur la pertinence scientifique des verdicts de bonne ou mauvaise formation textuelle opérant sur des séquences tronquées. Car dans 2 et 3, les facteurs qui conduisent à la production d'un certain pointeur, c'est-à-dire à celle d'un SN démonstratif doté d'un contenu lexical donné, ne doivent pas être recherchés dans les caractéristiques sémantiques internes des SN corréférénts, mais bel et bien dans des paramètres contextuels, qui favorisent ou au contraire inhibent l'utilisation de tel ou tel descripteur dans l'expression anaphorique. En l'occurrence, il est nécessaire de prendre en

⁶ Voir références et discussion in Reichler-Béguelin, 1989:306.

compte les *objectifs communicationnels* et les *caractéristiques de planification textuelle* qui président à la sélection des expressions concernées.

En effet, les séquences 2 et 3 présentent un type bien attesté d'anaphore lexicale démonstrative à fonction que j'appellerai *didactique*, où l'introduction d'un terme technique présumé inconnu du lecteur se fait par le biais de l'opération référentielle, l'"antécédent" qui précède jouant, *grosso modo*, le rôle de *definiens*. Dans 2, pour un décodeur qui n'a jamais rencontré le N-tête du SN anaphorique, c'est l'instruction interprétative véhiculée par le déterminant démonstratif qui garantit, à elle seule, l'interprétation coréférentielle; pour celui qui, d'aventure, le connaîtrait déjà, ce lexème à caractère technique participerait, au contraire, à la fonction identificatoire. Dans 3, l'apport de la dénomination spécifique (*ces "spatules"*) s'entoure d'ailleurs de guillemets, indice d'hétérogénéité énonciative et de distanciation (Authier, 1981, Cheong, 1988) qui rend encore plus évidente la fonction de reformulation didactique assumée par le SN référentiel (l'expression équivaut à quelque chose comme: *ces [oiseaux] à qui l'on donne le nom de, ou qui sont appelés, "spatules"*).

La séquence inverse, où le *definiendum* est fourni dans l'introducteur, le *definiens* dans le SN démonstratif, se rencontre elle aussi, et donne lieu à une anaphore qui peut être qualifiée de *définitionnelle*: plus banalement peut-être, le scripteur y met à profit l'opération référentielle pour gloser un terme technique qui vient d'être introduit, dont le lecteur potentiel peut aussi bien ignorer que connaître l'exacte signification. L'avantage le plus immédiat de la paraphrase définitionnelle est bien entendu d'éviter une répétition lexicale, tout en prévenant une éventuelle lacune dans l'information lexicale et/ou encyclopédique du lecteur:

4. vitamine C: elle est apportée par le cynorrhodon. *Ce fruit de l'églantier* est une source importante de vitamine C, bien connue comme vitamine de l'effort. (Notice accompagnant un médicament)

De tels enchaînements ne s'expliquent qu'à condition d'envisager les disparités qui sont susceptibles d'affecter les connaissances d'arrière-plan des interlocuteurs. Conscients de la possibilité de telles disparités, les sujets parlants cherchent souvent à les réduire de manière inaffichée, sans

exhiber la manoeuvre de "mise à jour" du savoir partagé à laquelle ils procèdent. Ainsi, les anaphores démonstratives à fonction didactique ou définitionnelle permettent-elles de faire utilement l'économie d'une parenthèse à caractère métalinguistique, qui serait destinée soit à *justifier une dénomination* ("*ces pierres étaient appelées des abbadirs*"), soit à *expliciter une signification* ("*le cynorrhodon est le fruit de l'églantier*"). Outre qu'elles interrompraient le fil du discours, de telles parenthèses courraient en effet le risque de paraître oiseuses ou redondantes à une partie des destinataires visés.

A ce point de la démonstration, on pourrait objecter que les anaphores démonstratives jugées anormales dans l'extrait de grammaire cité sous 1:

B Il acheta une propriété. Ce château...

Elle rencontra un animal. Cette vache..."

ne mettent pas en jeu de terminologie technique, et ne sauraient, par conséquent, tomber sous le cas de la stratégie textuelle visant à introduire une dénomination savante ou technique par le truchement du rappel anaphorique. Cependant, il existe une autre catégorie d'exemples, tout aussi bien représentée, où l'anaphore lexicale est spécifiante, quoiqu'elle n'ait rien de technique:

5. Au crépuscule d'une belle journée de printemps, ce cultivateur de la région de Madarounfa, au sud du Niger, examine une dernière fois sa future récolte avant de rentrer au village. Vendu un bon prix, *ce coton* devrait lui permettre d'acheter le mil qui manquera, cette année encore, pour la soudure. (*L'Événement du Jeudi*, 2-8.6.88)
6. C'est aussi le cas de Behrens, qui, après avoir eu la douleur de perdre un petit enfant, prit le risque, à 46 ans, d'en avoir un autre. Six ans plus tard, la présence de *cette petite fille* fait de Behrens une jeune mère à l'épanouissement physique spectaculaire. (*Le Point*, 25.4.88)

7 5 et 6 figurent déjà, avec d'autres exemples du même ordre, in Reichler-Béguelin 1989: 314 sqq.

Les séquences de ce type sont, elles aussi, produites en présence de facteurs contextuels précis et répertoriables. En 5, la description démonstrative "rappelle" bien, d'une certaine manière, *sa future récolte*, en tout cas au décodage; mais elle est en même temps interprétable comme un déictique de pensée représentée, assumé par un sujet de conscience instancié par le texte (le *cultivateur*), dans un passage assimilable à une sorte de discours indirect libre. Le lecteur admet fort bien que l'anaphore soit spécifiante, dans la mesure, précisément, où il est invité à *endosser le point de vue, qui peut être évolutif, d'un personnage sur le référent*. Un effet de dénomination rapportée est également latent dans l'exemple de Flaubert (2), si l'on veut bien admettre que le prédicat *honorer (x,y)*, dans lequel *x* est marqué [+animé], est un contexte opaque qui induit une ambiguïté sur la responsabilité énonciative de la dénomination choisie pour *y*: elle peut être le fait soit de l'énonciateur, soit le fait de *x*. Ces exemples permettent de considérer comme plausible la réalisation d'enchaînements comme 5', réalisés en exploitant les séquences réprouvées dans l'extrait de grammaire cité sous 1:

5'. Le directeur s'intéressa à une propriété en vente dans le voisinage. Ce château lui permettrait d'accueillir luxueusement tous ses clients. (Séquence forgée)

Elle compte les bêtes qui rentrent dans l'écurie. Ces quelques vaches lui fourniront au moins de quoi nourrir ses enfants. (Séquence forgée)

L'exemple 6, quant à lui, illustre encore un cas où l'information est introduite d'une manière cataphorique, procédé bien toléré dès que l'établissement d'un lien de coréférence est garanti par des moyens autres que strictement linguistiques, notamment quand il s'effectue en quelque sorte par défaut, en accord avec les normes plus générales d'un contrat narratif. Ici, l'anaphore démonstrative *cette petite fille* ne saurait, en effet, viser un autre référent que celui qui a été validé en mémoire discursive grâce à la séquence *prit le risque... d'en avoir un autre*. D'autre part, il est probable que la spécification assurée par la description démonstrative (on apprend par le biais de l'anaphore que l'enfant dont il s'agit est une fille) accompagne, en même temps qu'elle était, un changement de

perspective sur le référent, associé à la rupture temporelle opérée par "Six ans plus tard". Un phénomène du même genre peut d'ailleurs être décelé dans 3, où le suspens qui consiste à ne pas donner, dès la première mention du référent, son nom spécifique, permet la catalyse d'un observateur qui constate l'existence d'oiseaux inconnus, *avant* de les identifier plus précisément: la plausibilité de la séquence événementielle accroît dès lors très directement l'acceptabilité de la séquence prétendument pros-crite: "de magnifiques oiseaux blancs.... *Ces 'spatules'*".

Ces quelques cas suffiront à montrer, d'ores et déjà, à quel point les phénomènes de planification discursive, de polyphonie énonciative et de focalisation, au sens que Genette a donné à terme, sont déterminants pour expliquer la forme sous laquelle se concrétise linguistiquement une opération anaphorique. En abstrayant les pointeurs de leurs conditions d'occurrence, comme on a tendance à le faire dans les exemples forgés soumis aux tests d'acceptabilité, on court donc le risque de manquer les paramètres énonciatifs et interactionnels qui contribuent à déterminer la nature du matériau linguistique mis en jeu dans les opérations de pointage. Certes, les contre-exemples que j'ai apportés peuvent être considérés comme des "cas marqués", dans la mesure où les séquences du type *une vache... cet animal* seraient également utilisables dans les contextes particuliers qui admettent celles du type *un animal... cette vache*. On s'abstiendra néanmoins de faire le pas de trop qui consiste à caractériser cette dernière séquence comme déviante, alors qu'un inventaire des paramètres contextuels qui la rendent plausible, voire stylistiquement habile, semble une entreprise parfaitement réalisable. Quant à la relation "être un N ou du N (ou N)" qui, selon une formule de Corblin amendée par Kleiber, devrait être vérifiée entre l'introduit et le substantif N de la description démonstrative (cf. *supra*), plutôt que de la concevoir dans un absolu purement lexical, il convient de la faire dépendre d'une instance d'énonciation construite par le discours, ou inférable du discours: "être un N ou du N (ou N) pour X". La construction de cette instance d'énonciation par le décodeur fait aussi pleinement partie du processus interprétatif que l'exploitation des relations lexicales entretenues par les N-têtes du SN introducteur d'une part, et du SN anaphorique de l'autre.

2. Le cas du *ils* collectif

Pour évaluer sur un autre exemple l'utilisation qui est faite des jugements d'acceptabilité dans l'approche des phénomènes référentiels, nous examinerons un raisonnement proposé par Georges Kleiber, 1992, dans un intéressant article consacré à l'emploi du *ils* collectif dépourvu d'"antécédent" textuel ("Ils ont encore augmenté les impôts"). L'auteur observe d'abord que cet emploi particulier du *ils* concerne spécialement les référents animés et humains, avant de chercher à cerner quels types de N collectifs sont plus particulièrement aptes à être anaphorisés par un pronom au pluriel.

7. "Des auteurs comme D. Cruse (1986) et R. Mayer (1980) distinguent selon la relation qu'elles entretiennent avec les individus les composant trois types d'entités singulières collectives:

(i) celles comme *foule, troupe, groupe, poignée, tas, troupeau, galerie*, etc., appelées *Mengenquantitative* par R. Mayer et *collection* par D. Cruse, qui ne concernent pas typiquement les humains et qui, du point de vue métréologique, correspondent plus à des agrégats qu'à des entités organisées en systèmes;

(ii) celles comme *bourgeoisie, aristocratie*, etc., appelées *Klassenquantitative* (R. Mayer), ou *Classes* (D. Cruse), qui forment un ensemble d'humains, dont la réunion est "justified more by the possession of commun attributes than a common purpose" (D. Cruse);

(iii) celles comme *famille, jury, comité, ville, équipe*, etc., appelées *abgeschlossene Mengenquantitative* (R. Mayer) ou *group* (D. Cruse), dont la particularité est de regrouper des humains en *systèmes*: plus que des attributs en commun, les membres d'un groupe ont un but ou une fonction en commun, ce qui assure à l'ensemble un facteur cohésif plus grand que celui que possèdent les membres d'une classe (D. Cruse).

Cette distinction suffit pour mettre en relief le type d'entité référentielle associé au pronom *ils* collectif. Comme le montrent les énoncés ci-dessous:

*J'étais pris dans la foule. Ils ont failli m'étouffer

*La troupe progresse doucement. Ils sont fatigués

*Bien que l'agriculture n'a jamais connu de chômeurs, à l'heure actuelle ils doivent émigrer (exemple relevé par M.-J. Reichler-Béguelin)

ils collectif n'est compatible qu'avec les entités collectives de la troisième catégorie, celles qui, métréologiquement, constituent des groupes. Il s'agit bien d'une contrainte linguistique, puisque d'un point de vue cognitif les énoncés inappropriés ci-dessus ne posent pas de problèmes d'interprétation."

Dans ce passage, Kleiber affirme donc que seul un certain sous-ensemble de collectifs, ceux qu'il appelle à la suite de Cruse "noms de groupes", peuvent légitimement donner lieu à un rappel par un *ils* collectif. Selon un procédé usuel en linguistique, l'administration de la preuve est explicitement confiée aux jugements d'agrammaticalité qui affectent les trois séquences astérisquées par l'auteur: c'est leur caractère déviant qui, en fait, lui permet de "montrer" le fonctionnement de *ils*, et de manifester l'existence supposée d'une contrainte linguistique.

Parmi ces séquences, il faut observer que si les deux premières sont forgées, la troisième ne l'est pas: il s'agit d'un exemple tiré d'une copie d'étudiant de français langue seconde. C'est donc une production langagière authentique qui se trouve, par là même, exclue des observables à modéliser.

On pourrait certes justifier cette exclusion en estimant que les productions de locuteurs non natifs ne méritent pas plus de crédit que Meillet n'en accordait aux textes d'"illettrés" étudiés par Frei: d'ailleurs, l'auteur de l'exemple fait preuve d'une compétence inaboutie en utilisant incorrectement l'indicatif avec *bien que*. Cette erreur de morpho-syntaxe fait toutefois surgir un nouveau problème: n'est-ce pas elle qui joue un rôle déterminant dans l'impression générale de déviance ou de maladresse suscitée par l'énoncé en cause, ne contribue-t-elle pas, décisivement peut-être, à en motiver la disqualification? La forme prise par l'anaphore pronominale, en tout cas, ne semble pas caractéristique d'un apprenant non

francophone, et pourrait tout aussi bien se rencontrer dans la bouche ou sous la plume d'un natif: les textes français en fournissent de nombreux exemples (cf. *infra* dès 8). Et l'on hésitera sans doute à juger agrammaticaux des emplois similaires du *ils* collectif, quand ils apparaissent sous la plume d'écrivains comme Modiano ou Mme de Sévigné.

Par ailleurs, on peut se demander si les enchaînements que présentent les deux premiers exemples fournis par Kleiber sont vraiment moins "grammaticaux" que ne le seraient des séquences identiques, où le "nom d'agrégat" serait remplacé par un des substantifs catégorisés par l'auteur comme "noms de groupe", théoriquement aptes à être rappelés par un *ils* collectif:

7'. J'étais retenu dans la ville. Ils ont failli m'étouffer. (Séquence forgée)

L'équipe progresse doucement. Ils sont fatigués. (Séquence forgée)

Il paraît délicat, voire arbitraire, de juger corrects les exemples 7' au contraire de ceux donnés sous 7; et l'on voit à quel point l'intuition linguistique cesse rapidement d'être un recours solide quand il s'agit d'évaluer, hors de tout contexte, des séquences de ce genre.

En fait, une conception restrictive de la grammaticalité conduirait à rejeter 7' au même titre que les exemples contenant *foule* ou *troupe*. En effet, la norme puriste propage une certaine idée de la textualité qui hypertrophie les relations de rectio et d'implication réciproque entre unités linguistiques, extrapolant à l'échelle macro-syntaxique les contraintes d'accord spécifiques de la micro-syntaxe⁸. Sous l'influence d'une telle norme, un enseignant de langue ne laisserait probablement passer, dans une rédaction d'élève, ni les unes, ni les autres des séquences discutées ici. Car dans la pratique scolaire, les interventions du maître sur les copies sont massivement guidées par la règle -en partie artificielle et associée à des normes de genre- qui veut qu'un pronom, même non lié, tienne en principe ses marques d'accord d'un "antécédent" textuel explicite.

⁸ Même si cette démarche, qui revient à confondre *liage* et *pointage*, est indéfendable d'un point de vue scientifique; cf. Berrendonner et Reichler-Béguelin, 1989, et Berrendonner et alii, à paraître.

En contraste, on opposera une conception plus empiriste de la grammaticalité, tournée vers les occurrences effectives, et soucieuse de prendre en compte les spécificités de la communication orale, où les contraintes liées à la présence d'un contexte verbal ne sauraient être surinvesties comme elles le sont à l'écrit. Dans l'oral spontané, la forme prise par un pronom non lié n'est pas directement déterminée par l'état de ce contexte verbal, dont la "lettre" n'est pas mémorisable à long terme par les interlocuteurs, mais elle est déterminée par un certain *état de leur savoir partagé*, d'ailleurs également alimenté par de l'information non verbale. En matière de référence pronominale, ce sont alors les contenus mnésiques des sujets qui jouent un rôle déterminant, le "cotexte" étant pour ainsi dire inaccessible en tant que tel, car constamment retraité par la mémoire. Ailleurs que dans l'écrit étroitement normé, on hésitera donc à proscrire comme agrammaticaux non seulement les exemples figurant sous 7', mais aussi ceux qui sont astérisqués par Kleiber dans 7. La communication orale s'accommode couramment de *ils* collectifs désignant des référents absents de la situation d'énonciation et non mentionnés préalablement, qu'ils soient inférés du contexte ou utilisés, si je puis dire, *ex nihilo*, en "deixis mémorielle" (comme dans l'exemple: "Ils ont encore augmenté les impôts"). *A priori* et dans les conditions qui sont celles de la communication orale, on voit mal pourquoi une agrammaticalité naîtrait précisément et uniquement dans les cas où le référent visé par *ils* aurait été déjà introduit par un SN contenant un collectif "non conforme", nom d'agrégat ou nom de classe.

Au vu des problèmes, difficilement solubles, rencontrés dans la pratique du jugement d'acceptabilité, il est tentant de se rabattre sur les données, et de remplacer l'opposition *grammatical vs agrammatical* par une opposition *attesté vs inattesté*. Que devient la grammaire du *ils* collectif si l'on quitte le terrain de l'intuition du linguiste pour se confronter aux emplois réels? Voici, par exemple, deux rappels par *ils* survenant après des expressions collectives, mentionnés par Grevisse au compte de la "syllepse du nombre".

8. Je me trouvais [...] au premier étage de l'ancien hôtel de Zaharoff. Beaucoup de monde. Comme d'habitude, *ils* ne quittaient pas leurs pardessus. (Modiano, cité par Grevisse¹², p. 998)

9. Mon fils n'aura pas le chagrin de commander la noblesse de la vicomté de Rennes et de la baronnie de Vitré: *ils* l'ont élu malgré lui pour être à leur tête. (Mme de Sévigné, cité par Grevisse¹², p. 999)

A ces exemples littéraires, on peut comparer 10-14, tirés de différents types textuels, qui enchaînent semblablement un *ils* collectif sur un nom d'agrégat ou de classe, dans la catégorisation retenue par Kleiber. Le fonctionnement de la référence pronominale et son appropriation au contexte n'y diffèrent pas fondamentalement, semble-t-il, des cas comme 15 et 16, où c'est un "nom de groupe" qui est concerné par le rappel anaphorique au pluriel:

10. Je m'approchai du petit groupe afin de me retrouver parmi *eux*. (Copie)
11. Le groupe Cartier espère donc que *leurs* collègues et concurrents reviendront sur leur décision l'an prochain (...) (*La Suisse*, 16.4.91; l'accord au pluriel se trouve même ici dans un contexte intra-clausal.)
12. Le troupeau avance paisiblement. Des cailloux roulent sous *leurs* pieds. (Copie d'élève)
13. J'ai reçu une délégation des hôpitaux conduite par Skonieczny (administrateur). *Ils* ont déclaré que le Dr Hagen n'est pas en mesure de les aider (...). (*Journal du ghetto de Varsovie* d'Adam Czerniaków, *Temps modernes* 550, mai 1992: 42)
14. Le procès attire aussi la classe politique et syndicale. *Ceux-là* sont arrivés en autocar ou en avion. (*Courrier*, 13.2.90)
15. Les familles ne m'ont pas donné beaucoup d'amour maternel. *Ils* me ressentaient comme un étranger. (Oral, A2, 7.11.90)
16. Suite aux élections, je suis atterré de voir nos autorités dans leur léthargie et le peu de courage qu'*ils* ont à traiter ces problèmes. (Presse)

L'exemple peut-être le plus couramment attesté dans cette série concerne le nom *peuple*, difficile à classer dans la typologie inspirée de Cruse et Mayer:

17. Aussi le peuple d'Angleterre, qui se trouva le plus fort contre un de *leurs* rois, déclara-t-il ... (Montesquieu, cité par Grevisse¹² p. 954)⁹

Les polysémies à caractère métonymique, si fréquentes dans cette catégorie de substantifs, ont pour effet que des SN comme *le Bureau de contrôle pour la lutte contre le marché noir, le courant de JM Le Pen, ce journal, l'entreprise...* peuvent également donner lieu à des rappels par un pronom collectif *ils*¹⁰.

Au début de son article, Kleiber soutient, après Goudet, que les *ils* collectifs sans "antécédents" sont toujours au masculin et restreints aux référents humains. Si cette observation recouvre bon nombre de cas, il ne faudrait pas y voir un trait de signification des *ils* en question. Il existe en effet des exemples, relativement proches, où un pronom anaphorique au pluriel vise une *classe* ou une collection d'individus après mention, au singulier, d'un de ses membres¹¹, pas forcément marqué [+ humain] comme le montrent 19 et 20:

18. Si, après avoir fait ce test, vous avez encore des doutes, prenez rendez-vous chez un dermatologue. *Ils* ne s'entretenant plus sur le sujet comme avant, et le vôtre fera rapidement le point. (*Cosmopolitan*, mars 1992)

⁹ Les cas examinés précédemment mériteraient aussi d'être mis en rapport avec d'autres discordances d'accord, par exemple entre un sujet et le verbe qu'il régit:

Tout le reste de la maisonnée cet imbécile de Turandot compris *iront* au Mont-de-Piété (Queneau, in Grevisse¹² p. 711)

Grevisse souligne à plusieurs reprises la fréquence des accords qu'il appelle sylleptiques en ancien et moyen français, ainsi qu'à l'époque classique où ils étaient fort bien tolérés. Cf. Reichler-Béguelin, 1993.

¹⁰ Pour des raisons d'espace, les exemples (authentiques) auxquels je fais allusion ne seront pas donnés ici.

¹¹ Souvent introduit dans le texte par un SN indéfini à valeur non spécifiée ou cursive, ou encore par un SN dont le déterminant est un opérateur de sériation comme *chaque*.

19. Jamais il n'eût tourmenté un chat inutilement. Il *les* respectait. (Troyat, *in* Grevisse¹² p. 998)
20. Il articulait chaque syllabe et *leur* donnait une valeur musicale très sensible. (Valéry, *in* Grevisse¹² p. 998)

On notera, dans le même ordre d'idées, que le féminin n'est pas exclu en emploi collectif, si l'on en croit l'expression courante:

21. *Elles* sont bien toutes les mêmes.

L'anthropocentrisme des discours ambiants explique sans doute, à lui seul, la dominance des emplois où *ils* collectif désigne des animés humains. Mais les exemples 19 et 20 permettent de penser que, quand le prédicat incluant le pronom au pluriel est sous-catégorisé pour un certain type d'argument, et/ou quand le référent visé se trouve déjà, de quelque manière, activé en mémoire discursive, rien ne s'oppose théoriquement à ce que le pronom au pluriel désigne d'autres référents que les animés humains. En fonction de l'état des connaissances partagées et en fonction des thèmes de discours sélectionnés par le contexte, il paraît plausible de prédire des emplois comme:

22. A Berlin, *ils* sont fermés le samedi après-midi. <les magasins> (Séquence forgée)

Au zoo, on *leur* donne de la viande crue. <aux fauves> (Séquence forgée)

Dans ce tiroir, *ils* sont tous rouges. <les boutons> (Séquence forgée)

Il faut en tout cas souligner qu'un *ils* collectif est loin d'être toujours interprétable de manière purement locale, sans recours à des informations contextuelles parfois fort larges:

23. L'important, c'est d'être heureux. Si je ne me sens pas bien à Berlin, je changerai, ou *ils* me videront. (24 Heures, 23.3.1993)

Le pronom *ils* désigne, en l'occurrence, "les responsables de la Philharmonie de Berlin". La construction du référent requiert d'abord une connaissance du thème général de l'article dont 23 est extrait (un jeune flûtiste vient d'être engagé à la Philharmonie, prestigieux orchestre berlinois). Vient se greffer sur cette information le calcul interprétatif permis par les propriétés de sous-catégorisation du verbe *vider* dans son acception de "licencier". En fonction de ces propriétés, le référent ne peut viser ni les Berlinoises dans leur ensemble, ni même les membres de la Philharmonie: il s'agit forcément d'un sous-ensemble autorisé de ceux-ci.

C'est donc au-delà des contraintes locales liées à la forme et même à la catégorie sémantique d'un quelconque "antécédent" textuel que les exemples authentiques nous invitent à chercher la réponse au problème de l'interprétation des pronoms non liés, qu'ils aient d'ailleurs ou non valeur collective. On mettra un terme à ce trop bref tour d'horizon en suggérant que l'emploi d'un *ils* collectif dépourvu d'antécédent linguistique explicite pose, selon les cas et parfois simultanément, deux grandes questions (cf. Berrendonner *et alii*, à paraître): 1) Celle des règles logiques ou pseudo-logiques qui permettent d'inférer, en mémoire discursive, un référent pluriel ou une classe à partir d'un collectif ou d'un individu singulier (et inversement). 2) Celle de définir les circonstances de communication où un locuteur peut se permettre de pointer pronominalement sur un référent non explicitement introduit: le pronom étant par définition dépourvu de tête lexicale, le choix de ce marqueur implique en effet que le locuteur renonce à dénommer le référent à l'occasion du pointage.

Pour conclure sur le statut des données non standard, on relèvera qu'en sciences du langage, l'argumentation fondée sur la grammaticalité ou l'agrammaticalité d'exemples *attestés* est fragile, et ne tient qu'à un fil: son sort dépend exclusivement de la foi que tel ou tel lecteur sera prêt à accorder aux jugements de déviance émis par l'auteur de l'argumentation. Toute dissension à ce sujet ne peut que déboucher sur un conflit théorique, ce qui est regrettable. L'approche empiriste permet peut-être d'éviter cet écueil. Mais en collectant des observables, et en refusant de les trier avant description, le linguiste est, sans nul doute, astreint à réviser en permanence ses conceptions personnelles de la langue et du fonctionnement langagier.

Références

- Apothéloz, Denis (1993): "Aspects pragmatiques de l'opération de référence: les syntagmes nominaux surdéterminés", à paraître in Berrendonner et alii.
- Authier, Jacqueline (1981): "Paroles tenues à distance", in *Matérialités discursives*, Presses universitaires de Lille, 127-142.
- Berrendonner, Alain (1982): *L'éternel grammairien. Etude du discours normatif*, Berne, Peter Lang.
- Berrendonner, Alain et Reichler-Béguelin, Marie-José (1989): "Décalages. Les niveaux de l'analyse linguistique", *Langue française* 81, 110-135.
- Berrendonner, Alain, Reichler-Béguelin, Marie-José, Apothéloz, Denis, Rouault, Jacques (à paraître): *Stratégies référentielles et calcul des noms*.
- Cheong, Kye-Seop (1988): "Les guillemets, marqueurs de reformulation modulée", *Modèles linguistiques X*, fasc. 1, 71-81.
- Grevisse, Maurice (1986): *Le bon usage*, 12^e éd. refondue par A. Goosse, Paris-Gembloux, Duculot.
- Kleiber, Georges (1992): "Ils ont encore augmenté les impôts ou Sur le ils collectif", in L. Tasmowski et A. Zribi-Hertz, éd., *Hommages à Nicolas Ruwet*, 327-344.
- Reichler-Béguelin, Marie-José (1989): "Anaphores, connecteurs, et processus inférentiels", in *Modèles du discours. Recherches actuelles en Suisse romande*, C. Rubattel (éd.), Berne, Peter Lang, coll. "Sciences pour la communication", 303-336.

- Reichler-Béguelin, Marie-José (1993): "Anaphores associatives non lexicales: incomplétude macrosyntaxique?", in Stanislaw Karolak et Teresa Muryn (éds), *Complétude et incomplétude dans les langues romanes et slaves. Actes du 6e Colloque international de Linguistique romane et slave* (Cracovie, 29 sept.-3 oct. 1991), Cracovie, 327-379.
- Schnedecker, Catherine (1992): *Référence et discours: chaînes de référence et redénomination (Essai sur l'emploi en seconde mention du nom propre)*, Thèse de doctorat, Université des Sciences humaines de Strasbourg, Département de linguistique générale.