

« Non les gens (ils) sont pas dangereux » : variation et redoublement du sujet

Géraldine Zumwald Küster¹

(1) Université de Fribourg
geraldine.zumwaldkuester@unifr.ch

Résumé Ce papier s'interroge sur différents aspects de la variation touchant les propositions canoniques du français construites à l'aide d'un sujet SN. Après une brève description de la notion de « variation syntaxique » et de ses diverses interprétations, nous donnons un aperçu des propriétés de l'alternance entre sujet SN simple et sujet SN redoublé, pour souligner finalement les difficultés qui apparaissent lorsqu'il s'agit d'étudier cette variation, notamment à travers des approches quantitatives.

Abstract This paper raises some issues about the variation observed within the French canonical clauses composed of DP subjects. First, the notion of “syntactic variation” is looked at, along with its several different interpretations. Then, I expose some properties of the alternation between simple and doubled DP subjects and finally, I show that many methodological problems can arise when it comes to studying this variation phenomenon, especially when using quantitative measures.

Mots-clés : double marquage, dislocation, accord sujet-verbe, variation syntaxique, sociolinguistique variationniste

Keywords: subject doubling, dislocation, subject-verbe agreement, syntactic variation, variationist sociolinguistics

En français contemporain, les séquences [SNsujet + SV] font l'objet de deux réalisations concurrentes. Soit le SN sujet est directement suivi par le verbe (on parlera de « simple marquage » du sujet) :

- (1) simplement euh + pendant la mise en scène pendant après que enfin tout le t- **tout le monde était** là parce que euh ce qui se passe c'est que + on peut pas euh [oral, CRFP]

Soit le SN sujet est doublé d'un indice personnel (désormais IP) sujet de la série [*je, tu, il, nous,...*] (on parlera alors de « double marquage » ou « redoublement » du sujet) :

- (2) je suis comédien sur un autre spectacle + **tout le monde il est** comme ça + dispatché par rapport à sa profession [oral, CRFP]

La coexistence d'exemples comme (1) et (2), énoncés par un même locuteur lors du même entretien, suggère que la présence vs l'absence d'un IP subséquent au SN sujet peut être envisagée comme un cas de variation : on aurait ainsi, pour la séquence [SNsujet + SV], deux variantes en concurrence :

V1 = SN SV[Ø, V,...] = simple marquage

V2 = SN SV[IP, V,...] = double marquage

La reconnaissance d'un phénomène de variation impliquant V1/V2 et l'étude des modalités de cette alternance posent un certain nombre de questions que ce papier vise à mettre en lumière.

1 La variation syntaxique¹ : problématique générale

L'expression « variation syntaxique » renvoie souvent aux travaux variationnistes d'inspiration labovienne, qui ont pour principal objectif de quantifier l'influence de facteurs sociaux ou stylistiques sur l'usage de constructions dites « variantes ». Le cadre théorique de ces travaux a été élaboré dans le domaine de la phonologie, et son extension à des données qui ne concernent pas seulement des phonèmes mais aussi des morphèmes porteurs de sens a causé de nombreux débats (cf. Lavandera 1978, Romaine 1984, Gadet 1992). Au cœur de ces débats, on trouve la notion d'*équivalence* des variantes, posée par Labov (1972) comme condition nécessaire à l'exploitation sociale de la variation, mais critiquée par ceux qui considèrent « qu'à toute différence de forme correspon[d] une différence de sens » (Blanche-Benveniste, 1997, 20).

Suite à ces débats, plusieurs idées ont été proposées pour adapter le cadre théorique labovien à l'étude d'alternances en morphologie ou en syntaxe : Labov a suggéré de faire reposer l'équivalence des variantes sur une identité de sens descriptif uniquement, tandis que d'autres (Lavandera ou Romaine) ont proposé de prendre en compte le sens pragmatique des variantes. Face à ces désaccords, des chercheurs comme Jacobson (1989) ou Winford (1996) ont suggéré que c'était au chercheur de choisir, suivant les propriétés de l'alternance observée, le type de sens (descriptif, pragmatique, connotatif ou autre) sur lequel se base l'équivalence.

¹ Les IP sujets étant assimilés ici à des affixes, l'alternance entre V1 et V2 concerne plutôt la morphologie : nous parlons donc de « variation syntaxique » dans un sens large, pour désigner des alternances qui impliquent des morphèmes grammaticaux et/ou des constructions syntaxiques.

On voit donc que la notion de « variation syntaxique » héritée des travaux de Labov en phonologie connaît des interprétations et des applications diverses. Toutefois, la plupart des études qui traitent de variation syntaxique ont comme intérêt commun l'identification et parfois la quantification des différents facteurs qui peuvent influencer l'usage de l'une ou l'autre des constructions variantes. Suivant les propriétés des variantes et la constante de sens choisie, plusieurs types de facteurs peuvent être pris en compte : si les premières études de Labov se concentraient sur l'influence du contexte linguistique et des caractéristiques sociales des locuteurs, d'autres travaux ont cherché à mettre au jour des facteurs de nature pragmatique liés notamment à la fonction communicative des énoncés ou à l'organisation de l'information. Comme le soulignent Jacobson et Winford, le recours à des outils formels destinés à quantifier l'influence de tel ou tel facteur dans un phénomène de variation donné peut être plus ou moins approprié selon les cas.

2 Description des variantes

L'existence même des constructions à sujet redoublé est controversée : pour certains (par ex. Rizzi 1986, De Cat 2007), la présence d'un IP coréférent et cooccurrent avec un SN sujet indique que l'on a affaire à une construction disloquée. Selon cette analyse, lorsqu'un SN sujet est doublé d'un IP, il s'interprète toujours comme un élément syntaxiquement détaché, c'est-à-dire non incorporé à la proposition qui le suit.

Les constructions à SN initial disloqué repris par un IP coréférent sont généralement assimilées à des *thématisations*, c'est-à-dire, à des dispositifs qui signalent explicitement le statut de thème du SN sujet initial (voir par exemple Lambrecht 1981, Creissels 2006). Or, certaines unités comme les SN indéfinis du type *personne, quelqu'un, tout le monde, etc.*, sont incompatibles avec un statut thématique, de sorte qu'elles ne peuvent pas apparaître en position disloquée initiale. Les cas comme (2) *supra* ou (5) *infra*, où un sujet du type *tout le monde* ou *chacun* est doublé de l'indice *il* forcent donc l'analyse des SN indéfinis en termes de sujets intra-propositionnels, attestant de ce fait l'existence des phénomènes de redoublement du sujet (voir Berrendonner 2008 pour une argumentation plus détaillée).

On voit donc que les constructions V1 et V2 peuvent être envisagées comme deux variantes d'une seule et même séquence syntaxique, à savoir l'enchaînement tout à fait ordinaire d'un SN sujet et d'un SV constituant ensemble une proposition (ou l'initiale d'une proposition) canonique. L'identité de contenu descriptif et pragmatique est en principe assurée entre V1 et V2 : en effet, l'IP est redondant par rapport au SN et sa présence ne modifie pas les propriétés référentielles ou informationnelles de la proposition (voir toutefois les exceptions mentionnées plus bas). Nous sommes donc apparemment face à un phénomène de variation au sens le plus « strict » du terme, et il est dès lors pertinent d'envisager son traitement à travers des outils quantitatifs tels que ceux mis en place par les variationnistes laboviens.

L'alternance entre V1 et V2 est susceptible d'être conditionnée par différents critères, les uns ayant trait au système linguistique même et les autres concernant le domaine extralinguistique. Les propriétés oppositives de V1 et V2 sont nombreuses, mais la brièveté de ce papier ne nous permettra d'en exposer qu'un nombre très restreint.

Bien que l'information transmise par l'IP soit en principe redondante par rapport à celle qu'apporte le sujet SN, la variante V2 peut, dans certains cas, être plus informative que V1. Par exemple, lorsque le genre d'un SN sujet n'est pas marqué formellement, la variante V1 (3) peut présenter une ambiguïté que ne partage pas la variante V2 (3') :

- (3) L'artiste utilise tous ses sens. (3') L'artiste elle utilise tous ses sens.

La variante doublement marquée peut également apporter des informations sur le caractère spécifique ou générique du SN sujet ; ainsi en (4), le SN peut être assimilé à *un camembert* en particulier ou à *du camembert* en général. En (4'), la présence de *il* ou de *ça* permet d'orienter l'interprétation du SN vers un sens plutôt spécifique ou plutôt générique (exemple tiré de Blasco-Dulbecco, 1999) :

- (4) Le camembert sent très fort. (4') Le camembert il/ça sent très fort.

Parallèlement à ces conditionnements internes au système, il semble que les occurrences de V1 et V2 puissent aussi être influencées par des facteurs de nature extra-linguistique. Ainsi, Nadasdi (2000) a observé, dans un corpus de français parlé ontarien, que la répartition de V1 et V2 était sensible à l'origine sociale des locuteurs : il note 54% d'usage du redoublement chez les locuteurs de la classe ouvrière, contre 30% pour la petite bourgeoisie et 25% pour la moyenne bourgeoisie. Coveney (2003) observe une tendance similaire dans un corpus de français parlé de Picardie.

Le caractère socialement marqué du redoublement du sujet se vérifie aussi dans les textes littéraires. Souvent, les doubles marquages apparaissent dans des contextes bien spécifiques, notamment pour représenter le langage oral spontané des personnages (cf, en (5), la présence d'autres marques d'oral familier telles que *eh ben*, *pis*, *s'rait*, etc.) :

- (5) [Sarah] s'efforçait d'imposer, consensuel pourtant, un point de vue selon lequel eh ben **chacun il** avait qu'à faire comme y voulait et pis comme ça **tout l'monde y** s'rait content, y s'rait content, y s'rait content, y s'rait cont... [Benoziglio, *La voix des mauvais jours et des chagrins rentrés*, 2004]

3 Difficultés dans le traitement de la variation

Pour calculer la répartition des variantes V1 et V2, Nadasdi (2000) et Coveney (2003) s'inspirent de la méthodologie labovienne. Cette démarche a l'avantage de mesurer l'influence de divers types de facteurs sur l'alternance V1/V2, mais elle comporte un certain nombre de simplifications et d'approximations qui nuisent à la fiabilité des résultats.

Une première difficulté réside dans la distinction entre les constructions à SN disloqué à gauche et celles à SN redoublé : en effet, si cette distinction s'établit clairement en théorie, elle n'est pas toujours perceptible dans la pratique. Ainsi, de nombreuses séquences qui font apparaître un SN sujet susceptible d'être thématiqué et un IP sujet coréférent peuvent se prêter aux deux analyses. Pour exclure les dislocations et ne garder que les redoublements, Nadasdi se base principalement sur des critères d'ordre prosodique.² Or, Avanzi (2012) a montré que la prosodie ne permettait pas d'établir une distinction systématique entre dislocations et doubles marquages.³ Coveney considère quant à lui que la présence de l'IP sujet est variable

² Il classe les suites SN+IP+V comme des dislocations si l'un des trois points suivants apparaît : 1) pause entre le SN et l'IP 2) absence d'enchaînement entre le SN et l'IP 3) accent tonique sur la dernière syllabe du SN.

³ Son étude montre notamment qu'un enchaînement n'est pas un signe d'intégration syntaxique puisqu'il peut s'observer entre un SN clairement disloqué et un IP subséquent (par ex., entre l'objet *les livres* et le sujet *elle* dans *les livres elle les avait envoyés*). On voit aussi que la présence d'un accent tonique sur la dernière syllabe du SN ne dépend pas seulement de son statut disloqué vs intégré mais aussi du nombre de syllabes dont il est composé.

tant dans le cas des dislocations que dans le cas des doubles marquages, et il prend le parti de ne pas distinguer les deux constructions dans son étude. On voit donc que la difficulté de différenciation entre les structures à SN sujet intégré et celles à SN sujet disloqué pose problème pour la récolte des données servant de base à l'analyse quantitative.

Un autre point difficilement conciliable avec une étude quantitative concerne les occurrences de V2 faisant apparaître un IP *il/elle* et celles faisant apparaître un IP *ce/ça*. Conscients que l'usage de *il/elle* vs *ce/ça* est susceptible de transmettre une opposition du type *spécifique* vs *générique*, et en vertu du principe d'équivalence, Nadasdi et Coveney écartent de l'analyse quantitative tous les cas de SN doublés par *ce* ou *ça*. Cette mesure ne résout cependant pas le problème des alternances entre *il/elle* et \emptyset , qui peuvent faire l'objet d'une distinction *spécifique* vs *générique*, comme en (4) vs (4'). De plus, les propositions composées d'un SN sujet simple qui, s'il était doublé, le serait probablement par *ce* ou *ça*, posent problème :

- (6) Perso, je ne suis pas un grand pratiquant de planche à voile. Je débute et ne demande qu'à en apprendre davantage. Mais **la planche à voile me plaît bien** et quand on m'a proposé de bosser la dessus, j'ai préféré ça à l'aérodynamique et structure des ailes d'avions. [www.directwind.com]

Si les redoublements par *ce* ou *ça* sont exclus de l'analyse, on devrait s'attendre, parallèlement, à ce que les occurrences du type de (6) le soient aussi. Or, aucune mesure de ce type n'est mentionnée dans les études quantitatives que nous avons consultées. Du reste, même si une telle mesure apparaît comme une contrepartie nécessaire à l'exclusion des redoublements en *ce/ça*, sa mise en œuvre nécessiterait un travail de reconstruction et d'imagination (qu'est-ce qui *aurait pu être dit* mais qui ne l'a pas été ?) qui pourrait affaiblir la valeur scientifique des résultats.

La complexité de la question des redoublements en *il/elle* vs *ce/ça* est accentuée encore par l'existence de cas comme (7), où les deux types d'IP ne semblent pas s'opposer :

- (7) vous savez il nous manquait de tout pendant la guerre et alors je me laissais un peu courtiser, je crois que j'avais treize-quatorze ans, par deux garçons, **un il était** le fils du boucher, donc j'étais sûre toujours de d'avoir de la charcuterie à la maison, et **l'autre c'était** le fils du boulanger, car nous avions des tickets de pain [oral - radio]

Dans cet exemple, *il* et *ce* sont employés dans deux énoncés construits de manière parallèle, et ils accompagnent deux pro-SN de nature identique, qui renvoient à deux référents humains envisagés dans leur individualité. Il semble donc que dans certains contextes tels que (7), *il* et *ce* peuvent alterner sans être associés à des distinctions du type *spécifique* vs *générique*.

Ce bref survol des différents aspects de l'alternance entre V1 et V2 laisse entrevoir la complexité d'un traitement quantitatif de la variation touchant les séquences [SNsujet + SV] et les coûts théoriques qui peuvent en résulter. Cela ne veut pas dire qu'il faille abandonner tout essai de mesure quantitative dans le domaine de la variation entre V1 et V2. Pour mieux comprendre cette variation, il nous semble important de répertorier toutes les propriétés oppositives des variantes à l'intérieur du système linguistique, pour pouvoir ensuite interpréter au mieux les données relatives à leur répartition extra-linguistique. Ainsi, on pourra déterminer dans quelles mesures les variantes veulent *dire la même chose* et dans quelles mesures leur répartition inégale parmi les différents locuteurs et/ou les différents registres est associée à des différences au niveau du sens.

Références

- AVANZI M. (2012). *L'interface prosodie/syntaxe en français : dislocations, incises et asyndètes*. Bruxelles ; Bern : Peter Lang
- BERRENDONNER A. (2008). « Dislocation et conjugaison en français ». *Cahiers de praxématique* 28, 85-110
- BLANCHE-BENVENISTE C. (1997). « La notion de variation syntaxique dans la langue parlée ». *La variation en syntaxe*. F. Gadet (ed.). Paris. *Langue française* 115, 19-29
- BLASCO-DULBECCO M. (1999). *Les dislocations en français contemporain : étude syntaxique*. Paris : H. Champion ; Genève : Diff. Slatkine
- COVENEY A. (2003). « Le redoublement du sujet en français parlé. Une approche variationniste ». *Structures linguistiques et interactionnelles dans le français parlé*. A.B. Hansen & M.-B. Mosegaard-Hansen (eds.). Copenhague : Museum Tusulanum Press, 110-143
- CREISSELS D. (2006). *Syntaxe générale, une introduction typologique* (vol.2). Paris : Lavoisier
- DE CAT C. (2007). *French Dislocation, Interpretation, Syntax, Acquisition*. Oxford : Oxford University Press
- GADET F. (1992). « Hétérogénéité et variation : Labov, un bilan ». *Langages* 108, 127 p.
- JACOBSON S. (1989). « Some approaches to syntactic variation ». *Language Change and Variation*. R.W. Fasold and D. Schiffrin (eds.). Amsterdam/Philadelphia : John Benjamins Publishing Company, 381-393
- LABOV W. (1972). *Sociolinguistic Patterns*. Philadelphia : University of Pennsylvania Press
- LAMBRECHT K. (1981). *Topic, antitopic and verb agreement in non-standard French*. Amsterdam : J. Benjamins
- LAVANDERA B. (1978). « Where does the sociolinguistic variable stop? ». *Language in society* 7, 171-182
- NADASDI T. (2000). *Variation grammaticale et langue minoritaire : le cas des pronoms clitiques en français ontarien*. Munich : Lincom Europa
- RIZZI L. (1986). « On the Status of Subject Clitics in Romance ». *Studies in Romance Linguistics*. Jaegli O. & C. Silva-Corvalán (eds.). Dordrecht : Foris Publications, 391-419
- ROMAINE S. (1984). « On the Problem of Syntactic Variation and Pragmatic Meaning in Sociolinguistic Theory ». *Folia Linguistica* Vol. XVIII, n° 3-4, 409-438
- WINFORD D. (1996). « The problem of syntactic variation ». *Sociolinguistic variation: data, theory, and analysis*. J. Arnold & al (ed.). Stanford : Center for the Study of Language and Information, 177-192