

La créativité et la conventionnalité de groupes nominaux atypiques déterminant + pronom indéfini et leurs contextes communicatifs

1. Introduction

L'objectif de cet article est de dégager des critères permettant de reconnaître un emploi créatif d'un emploi conventionnel. La question a un intérêt heuristique et théorique. Son importance heuristique est la définition du statut d'usage courant, conventionnel mais atypique, créatif ou erroné d'une forme. Le statut d'usage n'est pas a priori marqué : la marque étant le point de départ de la linguistique de corpus qui constitue la perspective adoptée dans ce travail, cette approche se trouve dans ses versions actuelles (Teubert et Krishnamurthy 2007 : 8-9 par exemple) à affiner ses outils par la recherche sur des phénomènes atypiques (Hathout *et al.* 2008). Cette recherche est pertinente pour distinguer l'atypique et l'erroné dans l'analyse de l'oral spontané ou de textes de langues exotiques ou anciennes. L'intérêt théorique est le statut de l'usage des items dans une langue. Les connaissances constitutives des langues incluent des représentations abstraites comme les schémas syntaxiques. Ces schémas n'en sont pas moins actualisés par des items concrets qui peuvent imposer des collocations. Par exemple, ce n'est qu'à la première et non à la deuxième personne qu'on retrouve la séquence *défier qui que ce soit* ; en outre, la seule suite *je défie qui que ce soit* représentait plus de la moitié des attestations de *qui que ce soit* avec une lecture paraphrasable par *toute personne* dans le corpus de Larrivée (2007 : 73). De tels faits ne peuvent être accidentels, et suggèrent que l'usage fait partie des connaissances de la langue, comme force à le conclure l'existence de sélection de prépositions, des figements, des idiomes, des proverbes. La langue n'est donc pas un pur appareil de schémas générateurs de séquences. Cette réalité pose la question de savoir ce qui est conventionnel et ce qui est créatif dans une langue.

1 Aston University

Nous nous proposons de répondre à la question des caractéristiques des emplois créatifs en examinant un phénomène grammatical particulier caractérisant les pronoms indéfinis. Les formes retenues sont pour le français *qu(o)i que ce soit, quiconque, n'importe qu(o)i, quelqu'un / quelques-uns / quelque chose, je ne sais qu(o)i, Dieu sait qu(o)i* précédées d'un déterminant indéfini, défini ou démonstratif. C'est une construction atypique dans la mesure où un pronom qui fonctionne habituellement seul comme groupe nominal est accompagné par un déterminant qui introduit normalement une tête nominale (Larrivée 2009). La construction est considérée en anglais à travers les formes *somebody / someone / something, nobody / no one / nothing, anybody / anyone / anything* et *everybody / everyone / everything* introduites par les mêmes déterminants. Le choix d'un ensemble représentatif dans chaque langue vise non pas la grammaire contrastive des séquences, mais l'identification des caractéristiques formelles et linguistiques des emplois créatifs. L'hypothèse est faite que ces emplois s'associent à des environnements textuels discernables, ce qui nous amène à recourir aux corpus comparables de la banque Frantext et de la Bank of English, qui sont opposés aux données internet. Cette comparaison permettra de tester le rapport entre créativité linguistique et environnements textuels.

2. Données du français

Les pronoms indéfinis retenus (*qu(o)i que ce soit, quiconque, n'importe qu(o)i, quelqu'un / quelques-uns / quelque chose, je ne sais qu(o)i, Dieu sait qu(o)i*) connaissent des emplois atypiques où il sont précédés d'un déterminant (indéfini, défini et démonstratif). Le recensement du nombre d'emploi du déterminant immédiatement suivi de l'indéfini a été fait en octobre 2006 dans la banque littéraire Frantext.

	un / une / des	le / la / les	ce / cette / ces	
<i>qui que ce soit</i> 168	0	0	0	0
<i>quoi que ce soit</i> 712	0	0	1	1
<i>quiconque</i> 830	3	0	1	4
<i>n'importe qui</i> 678	0	0	1	1
<i>n'importe quoi</i> 1 234	3	3	3	9

<i>quelqu'un</i> 8 890	6	3	57	66
<i>quelques-uns</i> 2 290	4	5	3	12
<i>quelque chose</i> 19 720	15	8	158	181
<i>je ne sais qui</i> 75	1	1	0	2
<i>je ne sais quoi</i> 809	151	35	160	346
<i>Dieu sait qui</i> 6	0	0	0	0
<i>Dieu sait quoi</i> 53	0	0	0	0
	183	55	384	622

Tableau 1. Pronoms indéfinis précédés immédiatement par un déterminant dans les textes de Frantext depuis 1900

Les séquences brutes n'illustrent certes pas toutes un groupe nominal, qui n'est pas concerné par la suite du pronom sujet inversé «ce» et de l'attribut «quelqu'un» dans l'exemple suivant :

1. Baslèvre reprend : -était-ce quelqu'un de chez nous ? (E. Estaunié, 1919, *L'Ascension de M. Baslèvre*. Frantext)

Cet autre donne une forme qui représente l'ancien usage du *quelqu'un* d'approximation précédant le *un* partitif :

2. S'il arrivait malheur à un quelqu'un des leurs, ils mettaient le pays à feu et à sang. (A. France, 1918, *Le Petit Pierre*. Frantext)

Les emplois nominaux de *personne* et *tout*, adjectivaux de *nul*, *certain*, *divers*, *différent*, *nul* poseraient les mêmes interférences. Tous les groupes nominaux pertinents ne sont pas recensés par la recherche limitée à une proximité immédiate, qui ne donnerait pas l'exemple suivant s'il figurait dans Frantext :

3. Ceci étant, vous avez aussi raté **un vrai quelque chose**. Hier avait lieu le tout premier match de tennis nu diffusé sur Internet. (J. Dion. 22.5.03. En avoir ou pas. *Le Devoir* B6)

Ces interférences n'entament pas la pertinence des résultats relatifs (Church et Hanks 1990). Le rapport du nombre de combinaisons au nombre d'indéfinis montre qu'il s'agit bien d'emplois atypiques : aucune combinaison n'atteint le 1 % des emplois de l'indéfini, sauf *je ne sais qui* dont les 2 combinaisons représentent 2.67% des 75 occurrences, et *je ne sais quoi*, dont près de 43% des occurrences instancient la combinaison (au point où elle est empruntée par l'anglais). L'emploi connaît des disparités selon les formes qui laisse à croire à un statut d'usage différent. Le diagnostic de ce statut ne saurait se faire à partir d'un nombre d'occurrences absolu : un seuil de 10 occurrences par exemple ferait de *quelques-uns* une séquence conventionnelle alors que *n'importe quoi* n'en serait pas une, ce qui est contre-intuitif – *du n'importe quoi* semblant plus reconnaissable en français continental informel que *des quelques-uns*. Un nombre relatif est plus susceptible de tester le statut considéré, qui n'est pas établi par le rapport entre combinaisons et indéfinis seuls, comme le montre la faible proportion de la séquence *ce quelqu'un* qui semble intuitivement conventionnelle. Le rapport pertinent est celui des combinaisons entre elles. La majorité des séquences sont constituées par les trois formes *je ne sais quoi*, *quelque chose* et *quelqu'un*, qui ensemble livrent 95% des occurrences. Ces combinaisons se trouvent confirmées dans leur statut conventionnel par le fait qu'elles servent de modèle à des combinaisons hautement créatives, avec récursion du déterminant :

4. a. Ces filles qui font du folk ont **ce un je-ne-sais-quoi** qui les rendent [*sic*] sexys. (<http://www.mrchristopher.fr/bafouille/?p7>)
- b. Si l'âme est **un «ce quelque chose»**, elle est dès lors un individu. Or tout individu est d'une espèce et d'un genre déterminés. (<http://www.aquin.com/Pages/Traductions/QDAnimaHtm/QDAnimaQ1.htm>)

ainsi qu'en atteste le métadiscours d'une note du traducteur à la suite de (4b) :

5. **«Ce quelque chose»** est la traduction littérale de l'expression «hoc aliquid» qui elle-même est la traduction littérale de l'expression technique aristotélicienne *tode ti*. Les oreilles souffrent mais pourquoi faudrait-il rendre littéraire ce qui est littéral? Le démonstratif «ce» indique qu'il s'agit d'un individu, «quelque chose» indique qu'il s'agit d'une substance.

Certaines combinaisons ne sont d'autre part pas attestées. Cette absence d'attestations peut provenir de l'impossibilité des séquences, comme l'agrammaticalité de la suite *déterminant singulier + indéfini pluriel*. Elle peut relever des limites du corpus. Une recherche sur le corpus plus large des pages accessibles à Google en février 2008 donne les résultats suivants.

	<i>un / une / des</i>	<i>le / la / les</i>	<i>ce / cette / ces</i>
<i>qui que ce soit</i> 1 420 000	123 000 / 0 / 4	5 / 2 380 / 455	871 000 / 0 / 1
<i>quoi que ce soit</i> 4 260 000	01 370 / 7 / 7	1 060 / 1 170 / 5	1 670 / 1 / 0
<i>quiconque</i> 3 580 000	748 / 151 / 53	248 / 388 / 54	441 / 4 / 10
<i>n'importe qui</i> 1 810 000	13 700 / 4 910 / 2 030	15 700 / 8 290 / 2 570	1 220 000 / 2 / 398
<i>n'importe quoi</i> 6 940 000	3 820 / 4 390 / 6 320	82 600 / 27 700 / 542	332 000 / 4 / 1 770
<i>quelqu'un</i> 33 800 000	19 800 / 2 710 / 775	6 970 / 5 760 / 677	234 000 / 125 / 601
<i>quelques-uns</i> 2 920 000	738 / 254 / 585	909 / 2 130 / 19 500	603 / 2 / 65 800
<i>quelque chose</i> 21 700 000	48 000 / 3 930 / 12 800	9 640 / 112 000 / 26 000	159 000 / 682 / 4 500
<i>je ne sais qui</i> 105 000	2 450 / 8 / 881	836 / 3 560 / 9	45 000 / 4 / 4
<i>je ne sais quoi</i> 1 200 000	115 000 / 1 620 / 5 460	12 600 / 14 800 / 1 150	296 000 / 89 / 605
<i>Dieu sait qui</i> 7 040	0 / 0 / 0	1 / 2 / 0	0 / 0 / 0
<i>Dieu sait quoi</i> 39 300	2 / 0 / 0	4 / 7 / 0	1 / 0 / 0

Tableau 2. Pronoms indéfinis précédés immédiatement par un déterminant d'après Google

Un examen manuel des combinaisons avec moins de 10 occurrences brutes révèle des interférences. Hormis la simple erreur (*Donc je ne peux pas vous des quoi que ce soit au sujet de la capitale.* <http://www.un.org/icty/transf34/020614FE.htm>), des homonymes du déterminant interviennent comme l'adverbe *là* rattaché à un item qui le précède, expliquant les taux élevés de *la + pronom indéfini* ; de même pour le pronom *un* qui peut figurer en fin de proposition avant une autre qui commence par un indéfini. L'absence de sensibilité du moteur de recherche aux frontières syntagmatiques est encore illustrée par 2 occurrences de *le qui que ce soit* (p.ex.»Ignore-le, qui que ce soit»). À cela s'ajoutent le statut de la séquence (*Dieu sait quoi* servant de titre à une œuvre d'art contemporain d'un J.-P. Pollet) ; la mention plutôt que l'usage effectif de formes (comme dans l'exemple *Le qui que ce soit* montre que ce peut aussi être un païen, comme cet Egyptien du côté de son

père ; deux cas de *cette n'importe qui* reprennent la même occurrence de Frantext dans un article scientifique), des fragments émanant de traductions automatiques (2 *cette quiconque* figurant dans des listes de termes de pages pornographiques ; des fragments en français mal traduit d'une page en anglais concernent 2 *cette quelques-uns* et 2 *une je ne sais qui*). Est donné ci-dessous le nombre d'occurrences qui illustrent les groupes nominaux atypiques créatifs :

	une	des	le	les	ce	cette	ces	
qui que ce soit		4	1				1	6
quoi que ce soit	7	3		2	1			13
quiconque					1			1
n'importe quoi						4		4
je ne sais qui	4			5		2	3	14
Dieu sait qui			1					1
Dieu sait quoi	2							2
	13	7	2	7	2	6	4	41

Tableau 3. Répartition des emplois *déterminant + indéfini* représentés moins de dix fois dans les résultats bruts de Google

Les emplois créatifs sont marqués par différents traits qui permettent leur identification. Ces traits incluent des adaptations orthographiques comme les traits d'union ou les guillemets :

6. a. Seulement, en apprenant que ma maison avait été brûlée, que **ces je-ne-sais-qui** s'en étaient pris à Ylvin, je n'avais pas réfléchi... (<http://ozaline.over-blog.com/article-1361812.html>)
 - b. **Cette «je ne sais qui»** s'appelle en fait Clémence et est de mon département. elle était accompagnée d'Elsa, elle aussi de l'Ain. (<http://gutte-comics.over-blog.com/>)

Les points de suspension miment l'hésitation de l'oral, le fait de chercher ses mots dans une production non préparée pouvant amener à produire une suite atypique :

7. Si vous sentez que **cette... quoi que ce soit...** puisse être la clé d'une partie essentielle de votre histoire personnelle, (<http://www.jagfashion.net/viewtopic.php?p=104896&highlight=&sid=17ad2592abf0e545d36883cd256832df>)

Le métadiscours rend explicite la création néologique, comme en (5).

Les principales motivations linguistiques de l'emploi sont la coordination de groupes nominaux avec nom commun, que clôt un indéfini avec article indiquant la difficulté d'identifier plus précisément l'élément qui fermerait la série.

8. l'on me reproche, amèrement, d'avoir changé d'opinion sur les Juifs, les patriotes, les militaires, **les je ne sais qui, les je ne sais quoi.** (<http://fr.wikisource.org/wiki/Palinodies>)

Ces coordinations se retrouvent dans 12 des 41 occurrences. L'emploi anaphorique intervient cinq fois, à quoi s'ajoute une cataphore, avec le défini dans un cas et le démonstratif dans les autres. Un exemple est sans motivation attendue, et semble en effet peu acceptable :

9. En revanche, si vous possédez un pare-brise athermique, il lui sera plus difficile de repérer **les quoi que ce soit...** idem si vous êtes flashés par l'arrière, l'investissement aura été vain. Dommage, car à 1000 euros le détecteur, la contravention fait encore plus mal ! (<http://news.caradisiac.com/L-arme-absolue-anti-radar-existe-883>)

Les cas acceptables renvoient à un nom indisponible au moment de la construction du message auquel supplée un indéfini.

Particulièrement éloquent est l'environnement textuel. Les interventions dans des blogs et des forums constituent près de 66% des occurrences rapportées dans le tableau 3. Les autres emplois sont principalement des créations littéraires : 3 *des qui que ce soit*, 3 *une quoi que ce soit*, 1 *les quoi que ce soit*, 2 *une je ne sais qui* (un exemple dans une chanson, un autre de *une je ne sais qui de Ninon* dans un texte littéraire du 18^e siècle), 1 *les je ne sais qui*, 1 *cette je ne sais qui*, 2 *ces je ne sais qui* (un dans un texte de promotion d'un théâtre expérimental, un dans une traduction de Saint-Augustin), 1 *le Dieu sait qui*. On trouve dans des textes de ton informel à cause de leur public ou de leur sujet une occurrence de *une quoi que ce soit* dans une page s'adressant à des jeunes d'une école secondaire québécoise, et un emploi dans une chronique électronique sur les détecteurs de radar (en (9)).

On conçoit que les interventions non préparées des forums et blogs amènent des créations selon les besoins d'une expression construite en ligne. Qu'en est-il cependant des productions préparées des créations littéraires ? Comme la publicité ou l'humour (Munat 2007), les créations littéraires démontrent la virtuosité de l'auteur au-delà des formes consacrées par la norme afin de retenir l'attention du lecteur. Ce qui réunit les deux pratiques est la licence qui est accordée à une parole personnelle, à cause des conditions de production ou pour l'affirmation de sa virtuosité.

Cette section a établi les déterminismes textuels des emplois créatifs de la construction envisagée, ainsi que ses caractéristiques linguistiques et formelles. Le statut d'usage a été diagnostiqué par le rapport des nombres d'occurrences des combinaisons entre elles. La démarche est reproduite dans la section suivante pour la même construction en anglais.

3. Données de l'anglais

L'anglais connaît également les suites atypiques des pronoms indéfinis – *anybody / anyone / anything, everybody / everyone / everything, nobody / no one / nothing*, et *somebody / someone / something* – précédés d'un déterminant indéfini (*a, an*), défini (*the*) et démonstratif (*this, those, these*) (Quirk *et al.* 1985 : 376). Les combinaisons ont été recensées en mai 2008 dans le corpus de la *Bank of English* (par accès via l'Université de Birmingham, sous droit de HarperCollins). Composée de plus de 448 millions de mots, la BoE réunit des textes des années 1980 à nos jours de plusieurs variétés d'anglais – britannique, américain, australien, canadien notamment – et reflètent différents genres textuels, journalistiques, littéraires, techniques, universitaires, pour l'oral et l'écrit. Cette diversité a rendu nécessaire l'examen manuel des occurrences, dont moins d'un tiers relevaient du groupe nominal. Les résultats bruts n'illustrant pas la construction recherchée relevaient de problèmes de transcription des données orales, de polyfonctionnalité et de l'interférence de structures grammaticales autres. L'oral pose le problème de la fiabilité de la transcription, qu'illustre la difficulté de savoir si *a* renvoie à la préposition *of* ou au déterminant indéfini, si *the* correspond au pronom de troisième personne pluriel *they*, au déictique *there* ou au déterminant. La polyfonctionnalité de la forme *that* entre déterminant, relatif et conjonctif a amené à l'exclusion de la recherche initiale. Au-delà des quelques cas de franchissement de frontière syntagmatique, le fait qu'un déterminant précède immédiatement le pronom peut être dû au fait que ce pronom est un prédéterminant d'un nom, comme c'est le cas dans *the 'anyone for tennis' image*, et dans cet autre exemple :

10. Hailed by regular users as «the everything» cream.

Le nombre de groupes nominaux figure entre parenthèses à côté du nombre d'occurrences brutes des combinaisons dans le tableau suivant :

Pronoms indéfinis	Déterminants						
	<i>the</i> 24780121	<i>a</i> 9925232	<i>an</i> 1458277	<i>this</i> 1734383	<i>these</i> 400181	<i>those</i> 326168	
<i>anybody</i> 19451	3 (0)	2 (0)	2 (0)	6 (1)	0 (0)	0 (0)	13 (1)
<i>anyone</i> 68357	14 (3)	6 (0)	5 (1)	16 (0)	0 (0)	0 (0)	41 (4)
<i>anything</i> 112990	48 (5)	18 (0)	30 (3)	29 (0)	3 (0)	0 (0)	128 (8)
<i>everybody</i> 29120	10 (1)	8 (0)	3 (0)	25 (0)	1 (0)	0 (0)	47 (1)
<i>everyone</i> 71352	34 (1)	4 (0)	15 (0)	65 (1)	2 (0)	0 (0)	120 (2)
<i>everything</i> 85574	39 (7)	12 (0)	19 (2)	40 (1)	2 (0)	2 (0)	114 (10)
<i>nobody</i> 30310	18 (6)	214 (200)	4 (0)	55 (1)	2 (0)	0 (0)	293 (207)
<i>no one*</i> 48497	6 (0)	23 (4)	0 (0)	60 (0)	2 (0)	2 (0)	93 (4)
<i>nothing</i> 125795	74 (27)	132 (42)	0 (0)	52 (6)	2 (0)	2 (0)	262 (75)
<i>somebody</i> 25449	20 (12)	40 (29)	0 (0)	27 (5)	1 (0)	0 (0)	88 (46)
<i>someone</i> 80570	42 (32)	26 (16?)	0 (0)	46 (13)	0 (0)	2 (0)	116 (61)
<i>something</i> 196538	172 (93)	155 (51?)	3 (1?)	209 (30)	2 (0)	1 (0)	542 (175)
	480 (187)	640 (342)	81 (7)	630 (58)	17 (0)	9 (0)	1857 (594)

Tableau 4. Pronoms indéfinis précédés immédiatement par un déterminant dans la *Bank of English*

Ces chiffres montrent qu'on a affaire à une séquence atypique. Même la combinaison la plus fréquente de *a nobody* avec ses 200 occurrences ne représente qu'une portion infime (0.66%) de l'emploi du pronom qui apparaît 30 310 fois. De même, le pronom le plus fréquent (*something*, 196 538 emplois) ne livre que 175 exemples de la suite recherchée (0.089%). Comment de ces suites diagnostiquer les conventionnelles ? Ce test ne se trouve pas dans le rapport entre le nombre de combinaisons et de pronoms seuls : les suites également conventionnelles (*a nobody* et *this something*) entretiennent des rapports quantitatifs disproportionnés (0.66% de *a nobody* face à *nobody*, 0.015% de *this something* face à *something*). Comme pour le français, c'est le rapport entre les occurrences de la combinaison qui est révélateur : 84% des emplois sont donnés par les huit suites *a nobody*, *the nothing*, *a nothing*, *a somebody*, *the someone*, *the something*, *a something* et *this something*. Une indication du caractère conventionnel de ces suites est donnée par un fait propre à l'anglais qui est la morphologie plurielle s'ajoutant à l'indéfini : en effet, les indéfinis pluriels précédés d'un article sont *the everybodies* (1), *the everythings* (1), *the nobodys* (1), *the nobodies* (8), *three nothings* (1), *the somebodies* (1), *the somebodies* (5), *the someones* (1) et *the somethings* (1). La variation orthographique *ys* face à *ies* montre que ces emplois ne sont pas eux-mêmes conventionnels, mais qu'ils supposent l'existence d'un emploi conventionnel au singulier pour la majorité d'entre eux (sauf pour *everybodies*, *everythings*), et la dominance de *the nobodies* à cet égard est parallèle à celle de *a nobody*.

Ces propositions impliquent que les emplois atypiques des formes en *any-* et *every-*, de *no one*, de *nothing* avec démonstratifs, de *the + this + these + those nobody*, de *somebody* avec l'indéfini et le démonstratif, de *someone* avec le défini et le démonstratif ont un caractère créatif. Ce statut est confirmé par le ratio souvent défavorable entre les nombres d'emplois effectifs de la séquence et d'occurrences brutes, de 1 à 65 pour *this everyone* par exemple. Les marques qualitatives des guillemets, de la répétition et des hésitations identifiées en français se retrouvent pour l'anglais :

11. a. Then we give or, preferably, sell much of what we buy to state, universities, other conservation groups -- any responsible organization that can care for and protect the land from anyone. Unless **the "anyone"** builds nests or eats acorns. [USA: *ephemera*]
- b. It presented "a little bit of everything," says Eric Edwards, who is the venue's new, enthusiastic publicist. Among **the "everything"** were stripper Gypsy Rose Lee (born here), big-band star Duke Ellington and singer Al Jolson. [USA: *newspapers: 23 May 1996*]
- c. Between 1980 and 1988 **the "nobody"** was one of the finest bodyguards in the Spanish security company... [UK: *newspapers: The Times/Sunday Times: 31 Jul 1999*]

12. a. ...rather like Coleridge's Kubla Khan isn't a work that follows any prescriptive rules of writing. So that it isn't or it isn't **an anything** in a sense his Kubla Khan he says it is a fragment. A great deal of romantic writing and perception is by its nature fragmentary. [UK: spoken: college lecture: 'Coleridge as Critic']
- b. It was as much you know, when I was on land, the wind blowing and the sand getting in the paint and the bugs and **the everything, the everything**. [USA: radio: 8 Mar 1993]
- b'. Because people love the night so much. They love the night and the team and the beer and the ... **the everything**. [AUS: newspapers: 30 May 1999]
- c. Mrs Joseph agrees, even managing to summon some pity for him. "When I saw this boy in court, I couldn't believe that **this nothing**, this nonentity of a boy, had killed my gem of a child," she says. [UK: newspapers: The Times/Sunday Times: 27 Jun 2001]

La coordination intervient, pour expliciter l'interprétation de l'emploi inattendu en (12c) ci-dessus, ou pour amener l'indéfini dans l'élément final d'une liste à laquelle échappe au locuteur le lexème nominal (voir aussi (b),(b')):

13. He is loath to call himself a philosopher, an empiricist, a Buddhist or **an anything**, but in an era of autobiographical art it comes as something of a relief that he considers the big questions to be more important than the contents of his knicker drawer. [UK: newspapers: The Times/Sunday Times: 18 Jun 2001]

Un examen préliminaire montre que les textes dont proviennent les combinaisons créatives sont principalement littéraires, avec une bonne proportion de dialogues, des chroniques journalistiques, et quelques ouvrages philosophiques. On ne les retrouve pas dans les textes éditoriaux par exemple, où il s'agit d'affirmer une prise de position collective plutôt qu'un style personnel. Il serait évidemment intéressant de procéder à la considération des types de textes pour les séquences attestées par Google. Cette considération est malheureusement difficile, puisqu'une recherche en mai 2008 montre que les combinaisons sous les 10 000 occurrences sont rares (1 760 *the anything*, 2 910 *the everybodys*, 9 370 *the everyone*s, 5 870 *these everybody*, 6 020 *those everybody*, 4 760 *the everything*s, 8 710 *these nobody*, 7 880 *those nobody*, 531 *the noones*, 656 *these noone*, 499 *those noone*, 1 130 *the somebody*s, 3 870 *these somebody*, 4 170 *those somebody*, 1 130 *the somebody*s, 6 940 *the someone*s). Même les séquences réputées créatives sont associés à des nombres de pages qui dépassent les ressources raisonnables de traitement. On se contentera de noter que les occurrences au pluriel pourraient permettre de poursuivre les liens entre créativité et environnement textuel.

Les données de l'anglais nous ont permis de vérifier la robustesse d'une mesure relative pour le diagnostic des emplois créatifs, celle entre les occurrences de la construction elle-même. Ont été également confirmés les déterminismes textuels des emplois créatifs de la construction envisagée, ainsi que ses caractéristiques linguistiques et formelles.

4. Remarques finales

Le présent travail a montré qu'une langue ne correspond pas uniquement à la connaissance de mécanismes générateurs de séquences. Elle comprend également la connaissance de séquences effectives. Des séquences conventionnelles incluent des constructions fréquentes et des constructions atypiques, qui se mesurent par le rapport quantitatif entre une suite et ses unités constitutives. Les constructions atypiques ont aussi leurs instanciations conventionnelles, dont la mesure est donnée par le rapport majoritaire que ces instanciations représentent dans l'ensemble des combinaisons attestées. Comportant un ratio souvent très faible aux séquences brutes non pertinentes, les combinaisons créatives se reconnaissent par leurs marques formelles, linguistiques et textuelles. La variation orthographique des pluriels anglais, les marques graphiques du trait d'union, des guillemets, de l'hésitation, et le métadiscours caractérisent maintes attestations d'emplois créatifs. L'intervention créative des pronoms indéfinis se manifeste là où un nom n'est pas disponible, à la fin de séries de coordination notamment. La construction en ligne du message est une motivation textuelle majeure des emplois créatifs de la structure. Les autres environnements textuels où elle se retrouve sont les chroniques de journaux et la création littéraire. Ils démontrent le déploiement de la virtuosité d'une expression personnelle. La visibilité de l'énonciateur caractérise donc la créativité de l'expression spontanée et de l'expression préparée.

Une mesure novatrice des emplois créatifs des séquences atypiques par relation quantitative entre les combinaisons attestées est donc fournie par ce travail. Il pose la question théorique de savoir d'où viennent les créations, et montre qu'elles ne relèvent pas de la seule analogie. Leur source peut se trouver dans les contraintes de production en ligne. Le montrent la coordination et les hésitations, où le pronom indéfini intervient après un déterminant plutôt qu'un nom que le locuteur ne peut préciser. Le travail sur l'axe vertical, pour reprendre le terme de l'analyse en grille de Claire Blanche-Benveniste, explique ainsi certaines structures créatives. Les contraintes de la production sont donc également une source de créativité linguistique, qui appartient tout autant aux langues que les liens de la convention.

5. Bibliographie

- Bank of English*. <http://www.titania.bham.ac.uk/docs/about.htm>
- Church, K. W. et Hanks, P. 1990 : Word association Norms, Mutual Information, and Lexicography. *Computational Linguistics* 16, 1. 22–29.
- DeFrancq, B. 2006 : Étudier une évolution linguistique «en ligne» : *n'importe et peu importe*. *Français moderne* 74, 2. 159–182.
- Frantext*. <http://atilf.atilf.fr/frantext.htm>
- Hathout, N., Montermini, F. et Tanguy, L. 2008 : Extensive data for morphology : using the World Wide Web. *Journal of French Language Studies* 18, 1. 67–85.
- Hundt, M., Nesselhauf, N. et Biewers, C. (Éds) 2007 : *Corpus Linguistics and the Web*. Amsterdam : Rodopi.
- Kilgarriff, A. et Grefenstette, G. : 2003. Introduction to the special issue on the web as corpus. *Computational Linguistics* 29, 3. 335–347.
- Larrivé, P. 2007 : *Du tout au rien : libre-choix et polarité négative*. Paris : Champion.
- Larrivé, P. 2009 : What Nominal Phrases are all about: The Atypical Case of Indefinite Pronouns with a Determiner. *Romanische Forschungen* 121, 1. 3–19.
- Lauwers, P. 2008 : The nominalization of adjectives in French: From morphological conversion to categorial mismatch. *Folia Linguistica* 42, 1–2. 135–176.
- Malrieu, D. et Rastier, F. 2001 : Genres et variations morphosyntaxiques. *Traitements automatiques du langage* 42, 2. 547–577.
- Munat, J. 2007 : *Lexical Creativity, Texts and Contexts Studies*. Amsterdam : Benjamins.
- Quirk, R., Greenbaum, S., Leech, G. et Svartvik, J. 1985 : *A Comprehensive Grammar of the English Language*. London : Longman.
- Teubert W. et Krishnamurthy, R. 2007 : General Introduction. Éd. W. Teubert et R. Krishnamurthy. *Corpus Linguistics*. London : Routledge. 1–37.

