Séquences et bifurcations
Éléments de syntagmatique

Didier SAMAIN

UMR CNRS « Histoire des théories linguistiques » / Sorbonne Université

Résumé : Une approche syntagmatique des productions langagières effectives, normales ou pathologiques, met en évidence une alternance de séquences stables, au déroulement prévisible car elles sont proches d'une phraséologie, et de zones de bifurcation, caractérisées par une baisse brutale de la prévisibilité. Nous proposons ici une interprétation de ce phénomène, en l'articulant sur l'histoire des modèles syntaxiques contemporains.

Mots-clés : bifurcation, didactique, langage ordinaire, prévisibilité, séquence, syntaxe

Abstract: A syntagmatic analysis of concrete flowing speech, either normal or pathological, shows an alternation of stable chunks, whose patterns are predictable for they are close to phraseology, and bifurcation zones, characterized by a sudden drop in predictability. The paper offers an interpretation of this phenomena, with a hinge to the history of contemporary syntactic models.

Keywords: bifurcation, didactics, everyday language, predictability, chunk, syntax

Introduction

Nous avons intériorisé des notions comme celle de mot ou de phrase, sur lesquelles nous avons greffé, pendant notre scolarité, une bonne partie des catégories métalinguistiques qui se sont progressivement mises en place à partir de la grammaire alexandrine. Comme toute production scientifique ou intellectuelle, ces notions ont évolué au cours du temps ; malgré l'extrême stabilité de certaines d'entre elles, elles n'ont pas toujours eu l'apparence que nous leur connaissons aujourd'hui et auraient pu conduire à un édifice différent. La question de leur adéquation descriptive est complexe et leur assimilation directe à des catégories cognitives relèverait d'un scientisme naïf. Ce constat vaut aussi pour les formes modernes, savantes, de ces savoirs. En d'autres termes, il s'agit de *modèles* dans l'acception technique du terme (Blanckaert ; Léon ; Samain 2016), que leur forte prégnance institutionnelle ne suffit aucunement à natu-

raliser. Même des notions que nous percevons comme intuitives n'en sont pas moins, historiquement, des artefacts. On songera à celles de signe et de signification par exemple, à laquelle une philosophie du langage ordinaire sera en droit de demander comment il se fait qu'on puisse éprouver le sentiment de comprendre peu ou prou « la même chose » en entendant ou lisant un énoncé sans jamais éprouver le besoin, ni n'avoir jamais le moyen, de définir ni les mots, ni les choses. De son côté, le linguiste de terrain, qui parvient en principe à obtenir de ses informateurs la désignation de tel ou tel objet, susciterait presque à coup sûr leur perplexité s'il s'avisait de leur demander le « sens » de tel ou tel mot. – « Mais qu'est-ce que tu veux savoir au juste ? C'est ainsi qu'on dit dans telle situation. » – Rien d'étonnant à cela : la notion de signification est, elle aussi, une production culturelle, qui a trouvé sa quintessence dans l'artefact lexicologique, mais n'est aucunement un universel cognitif. Les mots des locuteurs sont normalement associés à des usages, non à des significations.

Des remarques similaires s'appliquent *mutatis mutandi* à toutes les catégories du métalangage grammatical et / ou linguistique. S'agissant de la syntaxe, c'est naturellement le cas de la phrase supposée canonique et de ses enchâssements. Il ne faut pas oublier que cette notion est du reste apparue tardivement dans le métalangage grammatical européen. Jusqu'au XVII^{e} siècle, les modèles étaient plutôt syntagmatiques que syntaxiques *stricto sensu*, et accordaient une place centrale à la dépendance, qu'il s'agisse de la *rection* dans le discours grammatical, ou de la *connotation* chez les logiciens. Et si on se penche sur les faits langagiers eux-mêmes, que la perspective soit diachronique ou qu'on s'intéresse sans préjugés normatifs aux productions réelles des locuteurs, force est de constater que les énoncés observés ne correspondent guère au schéma de la phrase dite canonique. Ce que l'observation met en évidence, ce sont plutôt des suites de syntagmes concaténés sans clôture programmée. J'insiste sur cette horizontalité des énoncés car elle va nous conduire à la question de la prévisibilité[1]. En bref, il en résulte qu'aucun des deux paliers traditionnels de l'analyse, le mot entendu comme signe, et l'énoncé interprété comme phrase, ne sont des donnés immédiats. Un constat que tout enseignant de langue en collège pourrait à juste titre juger banal[2].

1 Une réponse empiriste à l'idéalisme chomskyen

Le propre d'une modélisation est d'insérer une grille interprétative entre l'objet et sa description, mais l'écart qui en résulte se trouve parfois dédoublé par un paramètre propre aux sciences sociales, en l'occurrence la dimension normative. Ce rapport à la norme n'est pas anodin. Il a eu en l'occurrence des conséquences non négligeables sur la théorisation linguistique, comme en témoigne, au siècle dernier, le débat outre-Atlantique entre distributionnalisme et grammaire générative naissante. Tandis que Chomsky reprochait au distributionnalisme son incapacité à émettre des prédictions,

[1] S'agissant du français, les travaux de J.P. Seguin ont depuis longtemps montré à quel point la notion de phrase est peu pertinente pour rendre compte du « sentiment linguistique français » jusqu'au XVII^{e} siècle. Ils montrent notamment qu'on percevait plutôt l'unité d'une *période* dans ce que la tradition ultérieure nous a accoutumés à découper en « phrases » (Séguin 1998).

[2] Les enseignants connaissent en effet les difficultés des élèves à segmenter leurs énoncés en « phrases », et plus encore, à appréhender les signes dans un cadre lexicologique. Le fait est d'autant plus significatif qu'à l'inverse, et contrairement à l'opinion courante, ils maîtrisent généralement les particularités cognitives associées à l'écrit – quant à l'ordonnancement et la temporalisation des constituants d'un récit par exemple. (Sur ce point : Samain 2021.)

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

Hockett et les linguistes qui s'en sont inspirés pouvaient à bon droit lui répondre que ce qui allait bientôt s'appeler la *compétence* ne portait que des énoncés abstraits et normés, des *types* donc, et non des occurrences réelles. Or les énoncés effectifs ne se limitent pas, loin s'en faut, à des phrases déclarées grammaticales. Il s'agissait là d'une objection sérieuse aux prétentions « scientifiques » de la grammaire générative[3]. Dans les années qui ont suivi, des propositions alternatives sont apparues, qui ont visé à répondre à l'objection de Chomsky, tout en conservant les acquis descriptifs du distributionnalisme. Elles diffèrent profondément de la grammaire générative, en particulier par leur recours à des modèles syntagmatiques – horizontaux – très différents des enchâssements génératives. C'est le cas par exemple des travaux de Benoit Mandelbrot, qui a tenté d'appliquer les processus de Markov aux entités linguistiques (Mandelbrot 1961, Léon 2016). Pour mémoire, on appelle « chaîne de Markov » une séquence d'événements stochastiques dont la probabilité dépend des événements immédiatement antérieurs. Considérer un énoncé comme une chaîne de Markov implique donc qu'on prenne en considération l'environnement immédiat de chaque unité, et non sa structure putative globale. Cette méthode permet de faire apparaître des régularités et donc d'effectuer des prévisions, mais, contrairement à la grammaire générative, un tel « modèle » ne vise pas à dégager des macrostructures, puisque seules les relations de voisinage immédiat sont prises en compte. Il n'a pas non plus de prétentions explicatives, puisqu'il est purement statistique. Enfin et surtout, les énoncés abordés ne sont donc pas des *types*, mais bel et bien des *occurrences*, tandis que les unités qui les composent sont traitées comme des entités matérielles.

Cette présentation très sommaire suffit à montrer à quel point ces deux conceptions de la description linguistique étaient antagonistes, et probablement incompatibles. Du côté générativiste, nous avons affaire à un modèle visant explicitement à l'idéalisation de son objet, de l'autre, à un empirisme affiché, qui ne recourt pas aux catégories consacrées, et se contente d'évaluer la prévisibilité statistique d'un élément *e* à un endroit de la chaîne. D'autres théories ont témoigné d'une volonté similaire d'en finir avec les présupposés inhérents au métalangage hérité, au profit d'une démarche plus empiriste quoique moins radicale que celle de Mandelbrot[4]. C'est le cas des notions de *prefered argument structure* (Du Bois 1986) ou de *prefered clause unit* (Lambrecht 1987), dont l'argument est en substance que les « structures » grammaticales résultent de *régularités statistiques* sédimentées, et non d'une grammaire *a priori*. Il n'y est donc pas fait appel au critère d'acceptabilité au sens usuel du terme : sont jugées « acceptables » des séquences qui se sont progressivement stabilisées au cours du temps, généralement sous l'effet de choix discursifs[5].

[3] Assimilé à un critère non plus normatif mais objectif, équivalent d'une expérience cruciale dans les sciences de la nature, le jugement d'acceptabilité échappe difficilement à la circularité : la grammaire générative permet de générer tous les types de phrases définis comme grammaticales par la grammaire générative.

[4] Je me borne ici à des travaux syntaxiques « classiques ». L'approche dynamique des comportements langagiers inhérente à la sociolinguistique mériterait également d'être mentionnée.

[5] *« Preferred argument structure is that grammatical structuring of arguments which is statistically 'preferred' in clause tokens in discourse. […] It should be emphasized that the phrase « preferred argument structure » is meant to imply, not a discourse structure, but a discourse preference for a syntactic structure »*(Du Bois 1986 : 349).

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

Ce ne sont là que des exemples. Retenons-en simplement que ces analyses s'appuient sur de véritables corpus, sans préjuger de la nature profonde des structures mises en évidence et que, retrouvant en cela une tradition ancienne, leur perspective est syntagmatique et ne fait pas appel à la notion de constituance. Comme nous venons de le voir, la question que pose une chaîne de Markov est technique : *qu'est-ce qui est prévisible à un endroit donné de la chaîne ?* S'agissant d'un énoncé, la réponse semble intuitivement évidente. On s'attend à ce que l'observation révèle, d'une part, des microstructures phonétiques et morphosyntaxiques fortement contraintes, prévisibles en proportion, et, d'autre part, des structures textuelles plus larges, réglées de manière souple par des facteurs sémantico-pragmatiques. En fait ce n'est pas si simple.

On observe effectivement des zones stabilisées, associées au caractère fortement prescrit de la morphosyntaxe[6], et des zones de baisse brutale de la prévisibilité, qui sont autant de points de bifurcation, mais, c'est du moins ce que je tenterai de montrer rapidement, ces variations ne correspondent pas mécaniquement à la répartition consacrée en mot, phrase et texte. Mon ambition est par ailleurs d'illustrer plus généralement les facteurs de stabilisation des séquences. À cette fin, je m'appuierai en partie sur des énoncés déviants. En effet, qu'il s'agisse de pathologies ou, plus banalement, de productions dues à des locuteurs ordinaires « non experts », la plupart des énoncés anormaux se caractérisent par la faiblesse (voire l'absence dans les cas pathologiques) de certaines composantes du langage normé. Ils se révèlent donc heuristiquement intéressants puisqu'on peut, avec les précautions d'usage, les utiliser comme des miroirs grossissants[7].

2 Les aphasies dites de Wernicke

Je commence par un groupe d'aphasies, associées au nom de Carl Wernicke (1848-1905), qui, dans sa thèse, *Das aphasische Symptomen-Complex* (1874), montre que la destruction d'une aire corticale située derrière le sillon de Rolando, entraîne un type d'aphasie particulier, en quelque sorte inverse de l'aphasie dite Broca : le discours reste le plus souvent phonétiquement, voire grammaticalement correct, mais il est sémantiquement incohérent. Pour fixer les idées, voici un schéma tiré de l'*Encyclopedia Britannica*, qui indique les principales aires cérébrales sollicitées dans la pratique du langage oral et écrit.

Cette classification en *aphasies dites « de Broca »* (avec perte de la morpho-syntaxe) et *aphasies dites de Wernicke* (avec perte du sémantisme) a depuis été critiquée, mais elle reste un outil taxinomique commode, encore utilisé de nos jours. On a parfois qualifié le « Broca » de mutique intelligent et le « Wernicke » d'imbécile prolixe, ce qui est grossièrement caricatural, mais il est vrai que la perte plus ou moins prononcée de guidage sémantique peut conduire l'aphasique de Wernicke à des productions qui s'apparenteraient chez le locuteur normal à un jeu verbal. Après un énoncé du type

[6] Afin de simplifier au maximum la démonstration, je n'aborderai que très incidemment la prévisibilité phonétique.

[7] La documentation sur les aphasies de Wernicke provient d'une collaboration ancienne avec des aphasiologues (Samain 2000), la seconde est constituée d'exercices donnés à des élèves de collège en 2017 et 2018 par des professeurs stagiaires de l'Institut National Supérieur du Professorat et de l'Éducation de Paris.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

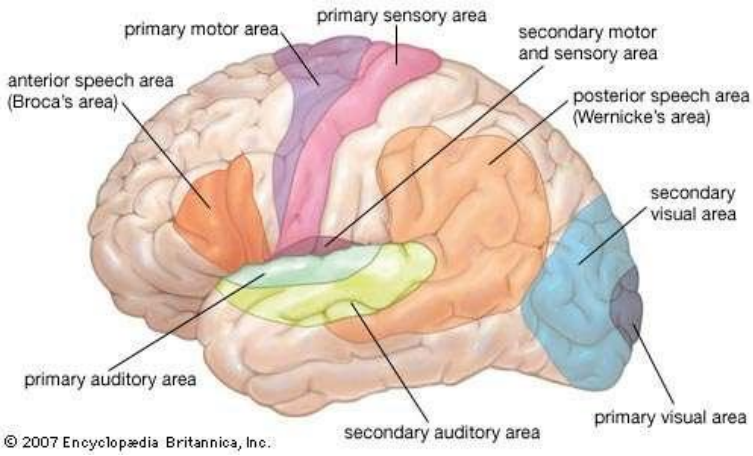


Fig. 1 : Aires cérébrales du langage

c'est en forgeant qu'on devient forgeron, ce type de patient (pour autant du moins que son lexique se soit maintenu) produira sans sourciller *… en lisant qu'on devient liseron*, *en mentant qu'on devient menton*, ou encore *en dinant qu'on devient dindon*[8]. Plus exactement, la plupart de ces patients ne perdent pas toute représentation sémantique, car ils sont généralement capables d'identifier une forme correcte si on la leur fournit. Il est plus juste de dire que leur syntaxe est uniquement constituée de séquences courtes, fortement mécanisées, associées à des réancrages sémantiques ponctuels mais dépourvues de stabilité topique. En voici quelques illustrations.

Alors qu'on lui a demandé de fournir le féminin de *dindon*, un patient répond *dinette*, ce que nous pouvons comprendre comme la production d'un dérivé approximatif, stabilisé par un lexème réellement existant. Selon le même mécanisme, *cordon* est lu *cornette*, immédiatement commenté par : *la cornette c'est les sœurs !* Tout se passe donc comme si l'identification d'une partie du lexème avait fonctionné comme un stimulus phonologique, lequel a déclenché un automatisme lexical, suivi d'un éventuel réancrage sémantique. – Est-ce fondamentalement différent de nos lapsus, voire de l'activité langagière normale ? Ce n'est pas sûr ainsi que nous le verrons plus loin.

On observe des automatismes semblables au niveau syntaxique. Sabouraud (1995 : 110) mentionne par exemple le cas d'un patient qui, au lieu de lire *courants maritimes*, lit *courants d'air maritimes*. – Que s'est-il passé ? *Courant d'air* est un syntagme figé, contrairement à *courant maritime*. Ces deux programmes syntaxiques sont entrés en concurrence, et la forme lexicalisée, automatisée, a prévalu, sans effacer toutefois celle qui impliquait un choix : un « courant » peut être *maritime, marin, aérien, électrique*, etc. Cette incapacité de bifurquer au bon moment (après *courants*), comme de choisir entre les deux formes, mérite également d'être soulignée. Quoi qu'il en soit, la grammaire distingue depuis longtemps entre structures figées et véritables constructions syntaxiques, mais des phénomènes de ce genre nous invitent à accorder une portée

[8]Les exemples qui suivent sont pour la plupart tirés de Sabouraud (1995). Ils ne visent qu'à fournir au non spécialiste quelques informations indispensables.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

bien plus générale à cette distinction, qui se révèle n'être qu'une manifestation particulière de la *différence entre ce qui est mécaniquement prévisible et ce qui ne l'est pas*. Ceci conduit à comprendre qu'une récupération partielle permette à ces patients un emploi globalement correct des relations dépendancielles (sujet-verbe, article-nom, …) – car il s'agit dans ce cas de *séquences* –, mais non des conjonctions, c'est-à-dire de mots qui se trouvent *entre* deux séquences et sont donc des points de bifurcation, imposant des choix.

La prégnance de la *forme matérielle* est constamment perceptible. La littérature spécialisée a souvent souligné que, même après récupération, ces patients ont des difficultés à interpréter correctement les métaphores. – On rapporte ainsi le cas d'un sujet (cultivé) à qui on demande comment il comprend le proverbe *Passé la fête, adieu le saint !* et qui le glose par *Je célèbre la Fête-Dieu*[9]. La seule explication possible de cette réponse est que la *contiguïté matérielle* des deux repères langagiers *fête* et *–dieu* a déclenché, à la façon d'un stimulus, la production de cet énoncé. Nous ignorons ce que le patient a éventuellement identifié entre ces repères, et nous ne savons pas davantage quelle représentation sémantique il a associé à cette séquence, mais il semble exclu qu'elle corresponde à des significations lexicologiques déterminées. Il s'agit selon toute vraisemblance d'un halo sémantique diffus. Quoi qu'il en soit, c'est donc moins la pertinence pragmatique d'un énoncé qui stabilise les réponses de ces patients que son existence comme formule dans la langue. – *Est-ce que ça se dit ?* demandent-ils volontiers.

Les points de bifurcation, de basse prévisibilité, qui débouchent sur de vraies discontinuités, allant du zeugme au changement d'univers sémantique, sont en fait des phénomènes courants dans les échanges ordinaires. Si ce n'est qu'ils se produisent sur des séquences moins brèves et qu'ils répondent en principe à des stratégies d'adaptation[10], tandis que notre aphasique semble en revanche dépourvu de cette capacité d'orienter correctement ses choix. Limiter ce phénomène à des bifurcations syntaxiques serait par contre réducteur et, sur ce point également, la sémiologie aphasique fournit une documentation utile. Pour les mêmes raisons que précédemment, la dénomination d'un objet ou l'écriture d'un mot isolé pose problème, alors que cette incapacité peut souvent être levée si les mêmes mots sont inclus dans des phrases familières. Dans le même ordre d'idée, un patient pourra dire quel jour nous sommes, à condition de réciter les jours de la semaine : lundi, mardi, mercredi… Il parviendra à indiquer le nombre de ses enfants, non pas en les dénombrant, mais par une récitation similaire : un, deux, trois…. Nous avons ici affaire à une syntagmatique élémentaire, qui passe

[9] Un locuteur normal qui ignore ce proverbe reste en mesure d'en inférer approximativement le sens : on oublie rapidement les services rendus. (La « Fête-Dieu » est le nom d'une fête dans l'Église catholique.)

[10] Sur ce point précis, les linguistes qui analysent les discours de Donald Trump disposent d'un intéressant corpus d'énoncés intermédiaires entre pathologie et normalité. Des études ont par ailleurs été consacrées aux discontinuités surgissant dans l'interaction ordinaire. Dans ses travaux sur la grammaire émergente, Hopper (qui rejette, tout comme les auteurs déjà mentionnés, l'idée d'une grammaire préexistante) montre qu'un énoncé se développe selon de constantes renégociations. – Par exemple, un locuteur qui change d'interlocuteur tout en parlant modifie son énoncé en conséquence. Exemple cité par Morling (2016) : *I gave up smoking cigarettes / … one, one week ago / actually.* Résumé dans le langage utilisé ici, cela signifie qu'il ne s'agit pas de structures syntaxiques, mais plutôt de séquences courtes, ponctuées de bifurcations aboutissant à des réorientations sous l'effet de mécanismes interactionnels (Hopper 1988, Hopper ; Traugott 1993).

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

par la récitation de paradigmes figés, contournant la difficulté à générer directement un élément non déterminé.

Ces phénomènes suggèrent que, loin de se limiter à une syntagmatique dans l'acception usuelle du terme et / ou à une phraséologie, les mécanismes séquentiels correspondent à des processus associatifs universels, intégrant nombre de phénomènes traditionnellement considérés comme paradigmatiques.

Autres exemples de contiguïté paradigmatique : le cas d'un patient qui pour *septembre* lit *octobre*, ou encore quand *cheval* est lu *jument*, ou quand *tâter le pouls* est lu *prendre la température*. Comme on le voit, le caractère *matériel* de cette contiguïté n'implique pas qu'elle soit phonétique ou syntagmatique. Ce système réticulé se montre au contraire insensible à la distinction {lexique / morphologie / syntaxe}, comme à celle entre langue et culture ou encyclopédie. Toute contiguïté stabilisée peut y donner naissance à une séquence, laquelle sera par définition pourvue de valences multiples et hétérogènes. Ces points de bifurcation sont des lieux de prévisibilité quasi nulle chez l'aphasique de Wernicke. Ce sont des lieux de choix sémantiques et de réglage social chez le sujet normal.

3 Le langage ordinaire

Nous venons de voir que la mise en séquence est un mécanisme transversal et multimodal, dont le fonctionnement n'est pas substantiellement modifié selon qu'il opère à l'un quelconque des « niveaux » de l'analyse linguistique, qu'il concerne des phénomènes « horizontaux » ou paradigmatiques, des faits langagiers ou des assemblages culturels. Dans les exemples analysés ci-dessus, les séquences, une fois amorcées, semblent fonctionner par saccades ponctuées de réancrages sémantiques. Lorsque les réglages sémantico-pragmatiques qui caractérisent le langage « normal » sont défaillants, cette multimodalité rend le couplage entre deux séquences quasi imprévisible. Les exemples qui vont suivre vont nous fournir d'autres illustrations de ce fonctionnement séquentiel. Je les prends dans un contexte très différent, en l'occurrence dans des productions d'élèves de collèges parisiens.

Les premiers documents proviennent d'une séance de grammaire assurée dans deux classes de 5e[11]. L'enseignante, qui a déjà parlé à ses élèves du *Supplément au voyage de Bougainville*, les a répartis en groupes et leur propose de donner un titre à une illustration tirée d'une édition d'époque. La tâche suivante consiste à lister tous les éléments qu'ils ont sous les yeux, *en les caractérisant* (couleurs, formes, dimensions, etc.). Il s'agit d'un cours sur le syntagme nominal, et elle espère éliciter ainsi des constructions de type {D+N+expansion}, que les élèves pourront ensuite classer et analyser.

Voici la gravure :

[11] En France, cette classe correspond à la deuxième année de collège, vers 12-13 ans. Les élèves sont censés avoir déjà acquis quelques notions métalinguistiques élémentaires (parties du discours et syntaxe de la phrase simple notamment). Dans tous les exemples qui suivent, les graphies ont été normalisées.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574



Fig. 2 : Illustration tirée d'une édition d'époque

Puis les titres proposés par les élèves :

5e A	5e C
L'arrivée de Bougainville à Tahiti	La nouvelle vie de Tahiti
La carte postale	Le marché des Tahitiens
L'amitié entre le monde	Partage de cultures
Le peuple tahitien	Le pillage de Christophe Colomb chez les Indiens
L'arrivée des peaux blanches à Tahiti	Les Tahitiens au plein milieu d'une île déserte
Des gens à Tahiti	La nouvelle terre
Bougainville enseigne les Tahitiens	Une nouvelle île
Petites affaires sur une île	La paix avec les Tahitiens
Le marché des Cythériens	Voyage merveilleux dans un pays étranger
L'arrivée à une nouvelle île habitée	Île tropicale commerciale
Une nouvelle terre	Nouveau troc
La commercialisation d'un monde	La colonisation

Et enfin les éléments observés, accompagnés de leurs caractérisations :

Les femmes à pied	Le journal qui est posé sur le sol
La mouette qui vole	Le peintre avec son tableau
La montagne marron et verte	Une plante verte
La loupe de l'homme	La mer bleue
Les hauts cocotiers	Le grand galion marron
Le cochon qui marche sur l'herbe	Le monsieur avec sa loupe
La statue à coté du cochon	Le tahitien qui fait à manger
Le grand nuage rond et blanc	L'eau bleue
Le monsieur qui cueille des fruits	La statue en pierre
Les noix de coco dans l'arbre	La mouette aux ailes grises
Les fleurs rouges	Le cochon qui marche sur l'herbe

Le résultat est sans appel.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

Que s'est-il passé ? S'agissant du titre, les enfants ont réinvesti des clichés culturels ou idéologiques disponibles, qui ont déterminé leur perception. L'adulte ne s'attendait sans doute pas à ce que cette image tahitienne génère le syntagme « Christophe Colomb »… Celui-ci a charrié avec lui un récit préfabriqué, *prévisible*. Mais ce constat vaut pour la plupart des autres titres, y compris lorsque ce dernier délivre un contenu approprié.

Quelles sont les similitudes avec les cas précédents ? Elles résident dans le fait que, contrairement à ce que supposait l'enseignante, les éléments qui composent la gravure n'ont pas été réellement vus, et encore moins observés. Ils ont bien plutôt été « reconnus » (correctement ou non, peu nous importe ici) et associés à des clichés préexistants, ce qui a généré des courts-circuits cognitifs. Cette impression de stéréotypie se confirme dans le deuxième corpus, quasi dépourvu d'indications de situation ou d'interaction. On note au passage la très faible *sensorialité* de ces énoncés, un trait également caractéristique des productions des aphasiques de Wernicke et que nous allons retrouver dans un instant. Sous une forme certes atténuée, cette prégnance de la formule, avec ce qu'elle charrie de halos représentationnels, rappelle par plusieurs traits ce qu'on observe en pathologie. Des mécanismes similaires[12] peuvent être observés dans la production écrite spontanée, où prédominent des blocs syntagmatiques plutôt que de véritables constructions syntaxiques, à l'intérieur desquels les accords grammaticaux s'opèrent souvent par simple contiguïté. Il en va de même, *mutatis mutandi*, de la lecture ordinaire, qui s'effectue le plus souvent, y compris chez le lecteur expert, par saccades, sans que ni la valeur lexicale des unités, ni la structure syntaxique des énoncés ne soient réellement perçues ou analysées. Il suffit pour une compréhension approximative d'identifier des repères dans la chaîne ; ces derniers fonctionnent alors comme des amorces ou des indices, complétés par les schémas syntaxiques et stéréotypes culturels disponibles. Une disjonction entre *l'exprimé* et *le compris* est par conséquent inévitable, dont l'importance variera selon que le sujet est cultivé, lecteur apprenant, dislexique, etc.

Dans une autre classe, l'enseignante a demandé aux élèves de décrire une scène qu'ils ont personnellement vue et mémorisée. L'exercice s'appuyant sur le vécu individuel des enfants, on pouvait donc s'attendre à des productions plus originales que celles de l'exercice précédent. Voici un bref échantillon des textes fournis.

> – Dans ma cité : la cité blanche, il y a des habitats [*sic*], un jardin où on cultive des légumes. Il y a aussi des fleurs et il y a beaucoup d'espace. Il y a aussi un garage avec beaucoup de voitures. Il y a aussi une cave reliée au garage.
>
> – Ce paysage représente un petit toboggan bleu qui est isolé. Ce parc est tout seul sous le ciel bleu et le soleil tapant sur le toit de l'attraction. Les oiseaux volent en liberté. Des passants se promènent. Ils s'assoient sur l'herbe pour pique-niquer. L'enfant veut faire du toboggan. Il mange, après il va vers l'attraction. Après la journée est finie. Il rentre chez lui, dans sa petite maison loin du parc.

[12] Les observations qui suivent résultent elles aussi d'un travail mené avec des enseignants de terrain. Les documenter en détail excèderait les dimensions prévues pour cet article.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

La pauvreté descriptive peut s'expliquer par des limites expressives, normales chez des enfants de cet âge, mais les similitudes formelles avec les énoncés précédents n'en sont pas moins significatives, dès lors que l'exercice demandé était très différent. Ces « descriptions » se limitent en effet pour l'essentiel à des dénominations. Nous relevons la même pauvreté sensorielle, la même absence de véritables indications spatiales : ni sons, ni odeurs, ni spatialité. Les rares notations sensorielles se bornent à des indications stéréotypées de couleur : *cité blanche*, *ciel bleu, soleil tapant*. Nul doute que les enseignants de ces enfants auront fort à faire pour *sensorialiser* leur langage. L'absence de mots de liaison, et plus généralement d'articulations logiques, est également frappante.

Au regard de ce que nous avons observé jusqu'à présent, tout cela n'en est pas moins instructif. Ici encore il ne faut sans doute pas attribuer à cette série de clichés un contenu sémantique dans l'acception lexicologique du terme. Ce sont juste des supports lexicaux, peut-être suffisants pour déclencher la mémoire associative des enfants qui les ont utilisés, mais naturellement insuffisants par eux-mêmes à reconstituer le paysage avec précision. Le paragraphe précédent s'est achevé sur le hiatus entre *exprimé* et *compris*. Nous observons ici le même mécanisme, mais du point de vue de l'émetteur, c'est-à-dire entre *l'intentionné* et *l'exprimé*. Comme précédemment, la largeur de ce hiatus dépend des compétences langagières et culturelles des locuteurs (c'est précisément la tâche de l'enseignant que de le réduire !), et le succès de la communication *ordinaire* repose donc en bonne partie sur la prévisibilité des séquences culturelles que les repères présents dans le message seront susceptibles de déclencher chez le récepteur. Dans les exemples précédents, l'information véhiculée tendra donc à se limiter à l'espace préalablement partagé entre les enfants. *Stricto sensu*, il n'y a donc guère d'apport informatif, puisque pour cela il faut de l'imprévisibilité[13], laquelle ne peut survenir que sur une bifurcation. Dans ce cours de français, comme dans une rééducation aphasique, c'est ce qui se passera, lorsque nos jeunes locuteurs parviendront par exemple à employer des adjectifs épithètes, et plus généralement des expansions du nom, de façon non prescrite. On comprend inversement pourquoi (en dehors naturellement de leur emploi stéréotypé), le recours à des épithètes, tout comme le recours aux mots de liaison, sont moins spontanés chez les élèves que la nomination elle-même. On peut émettre l'hypothèse que la gestion d'une bifurcation demande davantage d'énergie cognitive que l'acquisition, et *a fortiori* l'emploi, d'une séquence.

Nous pourrions sans doute décrire cela à l'aide d'une grammaire dépendancielle améliorée. Les linguistes connaissent le fameux « stemma » tesniérien[14] et, quand on évoque ces grammaires, on songe immédiatement à des choses du genre ci-dessous :

[13] C'est le principe de la théorie de l'information !

[14] Tesnière (1959) considère que le sujet n'est qu'un complément parmi d'autres, et que le verbe constitue le noyau de la phrase, « complété » par des groupes nominaux. Il appelle *jonction* une relation hiérarchisée entre deux éléments (ci-dessous, entre *petits* et *ruisseaux* par exemple) et *connexion*, une relation non hiérachisée (telle qu'une coordination).

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

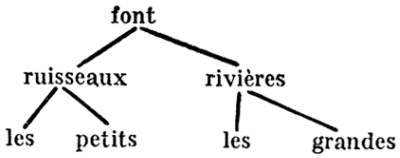


Fig. 3 : Grammaire dépendancielle

Cette représentation, extrêmement schématique, a le défaut de subsumer sous le nom de *jonction* deux relations dépendancielles distinctes, la dépendance simple et la corrélation. Disons, pour faire bref et en reprenant les exemples précédents, que la présence d'un adjectif implique celle d'un substantif, mais non l'inverse. C'est ce que les logiciens médiévaux appelaient une connotation. En revanche, la relation de l'article et du nom est d'implication réciproque en français, puisque la présence de l'un de ces éléments entraîne automatiquement celle de l'autre. C'est donc une corrélation. Cette assimilation indue entre dépendance simple et corrélation a sans aucun doute été favorisée par la confusion, systématique chez Tesnière, entre dépendance sémantique et dépendance syntaxique. Quoi qu'il en soit, si l'on voulait adapter le modèle dépendanciel à ce que nous venons d'observer, il faudrait donc le réorganiser de manière à distinguer entre les relations obligatoires (qui génèrent des séquences, le plus souvent constituées de corrélations) et les relations non obligatoires (regroupant les « connexions » et certaines jonctions). Il faudrait bien entendu prendre également en compte le fait que les relations syntaxiques ne sont qu'un mode parmi d'autres de ces mécanismes dépendanciels, dont fait également partie la phraséologie, les usages qui règlent les interactions, les clichés culturels, etc.

Il n'entre nullement dans les ambitions du présent article d'exposer, même rapidement, un tel modèle. Il importe simplement de souligner que, quel que soit le domaine concerné, on retrouve toujours cette différence entre deux types de relations. Dans le cadre qui nous intéresse ici, elle correspond à une différence entre deux régimes hétérogènes de prévisibilité ou d'imprévisibilité relative. Si nous en revenons donc à cette évidence que l'acquisition d'automatismes ne procède pas des mêmes mécanismes que la capacité à choisir, alors on ne peut au passage que regretter que ce facteur ne soit pas davantage pris en compte dans l'enseignement des langues, voire dans l'enseignement en général.

Conclusion

Pour le linguiste et l'historien des sciences, la thématique du présent volume invite donc à redisposer différemment les catégories léguées par la tradition. Il résulte en effet de ce qui précède que, dès lors qu'on se propose de décrire, non des langues, mais le fonctionnement langagier, les catégories usuelles sont peu utiles. Il est plus réaliste de leur substituer un modèle horizontal et statistique, cherchant à identifier des séquences stabilisées et les moments de baisse drastique de prévisiblité, que j'appelle des points de bifurcation. La notion de prévisibilité permet à ce titre une approche plus empiriste du langage, affranchie du poids de la tradition grammaticale occidentale.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

Cette tradition s'est bâtie sur l'opposition entre lexique et syntaxe. C'est commode pour grammatiser[15] des langues, mais il suffit d'analyser des productions d'enfants pour constater que *la coupure réelle se situe ailleurs*, puisque ces enfants n'emploient pas des mots mais plutôt des phrasèmes, et qu'*elle est également plus abstraite* et générique, puisque les zones de prévisibilité ne se limitent pas à des enchaînements morphosyntaxiques, et que les points de bifurcation sont de même tout aussi bien lexicaux (un *item* non disponible par exemple), ou culturels, que syntaxiques (tels des épithètes ou des connecteurs).

Il résulte aussi de ce qui précède qu'il n'existe pas plus de structures syntaxiques abstraites que de catégories *a priori* de l'espace et du temps. Le mérite, tant des grammaires dépendancielles que de quelqu'un comme Mandelbrot, est d'avoir au contraire mis en évidence le rôle central des unités concrètes. Du point de vue cognitif, espace et temps n'existent que par les objets ou les souvenirs que nous y plaçons. De même, l'« espace » syntaxique, et *a fortiori* l'espace langagier, n'existent que par les points de repères qui soutiennent les séquences.

Cela nous rappelle enfin que tout sujet intègre des rites discursifs, plus ou moins contraints, pour s'exprimer et communiquer. En voici pour finir deux illustrations volontairement choisies dans des champs disciplinaires très différents. D'abord l'activité de récitation. On décline les lettres de l'alphabet, comme on chante une comptine, comme on décline *rosa, rosa, rosae*, etc. Ces « paradigmes » reposent, non sur la liberté des commutations, mais sur le caractère contraint de la série. J'ai indiqué l'importance de ces récitations en aphasiologie. Il en va de même à mon sens en ontogenèse, pour l'apprentissage du langage. Ensuite, dans un tout autre domaine, la socialisation de l'expression des affects. On songera à la notion d'*idiome de détresse*, qui a donné lieu à de riches analyses anthropologiques depuis son introduction par Nichter à la fin du siècle dernier (Nichter 1982). L'intérêt de cette notion n'est pas, trivialement, de souligner la traductibilité somatique de la détresse, mais de suggérer qu'il s'agit d'*idiomes*, au sens anglais : *idioms of distress*, c'est-à-dire de dispositifs socialement codifiés qui lui fournissent une charpente expressive[16].

Peut-être le lecteur parvenu jusqu'ici sera-t-il tenté de se récrier : est-il donc impossible d'échapper à cela ? Faut-il en déduire que la prétention au discours singulier n'est qu'une illusion ? Sur le plan des principes, la réponse va de soi : *Que pourrait-on formuler en dehors du code* et des prescriptions qui le sous-tendent ? En ce qui concerne les stratégies individuelles, les observations qui précèdent laissent la question ouverte. Mais ce topos philosophique importe assez peu pour la question qui nous occupe. Ces observations témoignent en revanche, et cela m'importe davantage, du rôle proprement cognitif de la formule.

[15] La *grammatisation*, à ne pas confondre avec la grammaticalisation, désigne le processus historique par lequel les langues ont été décrites et normalisées (Auroux 1995).

[16] Cela ne signifie pas non plus, à l'inverse, que les sujets seraient irrémédiablement enfermés dans des structures discursives. Comme dans le langage ordinaire, la santé mentale se mesure à la capacité à choisir l'expression, et donc à négocier le figement. C'est cette dialectique que les études sur les idiomes de détresse ont mise en évidence.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

Bibliographie

Auroux, Sylvain, *La révolution technologique de la grammatisation*, Mardaga, Liège, 1994.

Blanckaert, Claude ; Léon, Jacqueline ; Samain, Didier (dir.) (2016), *Modèles et modélisations dans les sciences du langage, de l'homme et de la société*, L'Harmattan, Paris.

Du Bois, John (1986), « Competing Motivations », *in* Haiman, John (dir.), *Iconicity in Syntax*, John Benjamins Publishing Company, Amsterdam / Philadelphie, p. 343-365.

Hopper, Paul (1988), « Emergent Grammar and the A Priori Grammar Postulate », *in* Tannen, Deborah (dir.), *Linguistics in Context : Connecting Observation and Understanding*, Ablex Publishing Corporation, Norwood, p. 117-134.

Hopper, Paul ; Traugott, Elizabeth, *Grammaticalization*, Cambridge University Press, Cambridge, 1993.

Lambrecht, Knud (1987), « On the status of SVO sentences in French Discourse », *in* Tomlin, Russell (dir.), *Coherence and Grounding in Discourse*, John Benjamins Publishing Company, Amsterdam / Philadelphie, p. 217-261.

Léon, Jacqueline (2016), « Les chaînes de Markov. Parcours d'un "modèle" au fondement de la mathématisation et de l'automatisation de la linguistique », *in* Blanckaert, Claude ; Léon, Jacqueline ; Samain, Didier (dir.), *Modèles et modélisations dans les sciences du langage, de l'homme et de la société*, L'Harmattan, Paris, p. 121-147.

Mandelbrot, Benoît (1961), « On the theory of word frequencies and on related Markovian Models of Discourse », *in* Jakobson, Roman (dir.), *Structure of Language and its Mathematical Aspects*, American Mathematical Society, *Proceedings of Symposia in Applied Mathematics*, vol. 12, p. 190-219.

Morling, Charles-Henry (2016), « Quels modèles pour la grammaire émergente ? », *in* Blanckaert, Claude ; Léon, Jacqueline ; Samain, Didier (dir.), *Modèles et modélisations dans les sciences du langage, de l'homme et de la société*, L'Harmattan, Paris, p. 149-164.

Nichter, Mark (1982), « Idioms of Distress : Alternatives in the Expression of Psychosocial Distress : A Case Study from South India », *Culture Medicine and Psychiatry*, vol. 5, No. 4, p. 379-408.

Sabouraud, Olivier, *Le langage et ses maux*, Odile Jacob, Paris, 1995.

Samain, Didier (2000), « Le langage et l'idiome : les partitions sur l'espace grammatical au vu de quelques pathologies », *Sémiotiques*, vol. 18 / 19, *Incidences de l'impossible dans le langage*, p. 31-63.

Samain, Didier (2022), « Métalangage vs. stratégies langagières. Remarques sur le modèle grammatical et ses fictions didactiques », *in* Lewi, Olivia ; Longhi, Blandine, *Segmenter, connecter, ponctuer à l'écrit : deux défis pour l'enseignement*, Peter Lang, Berne p. 105-127.

Séguin, Jean-Pierre, *L'invention de la phrase au XVIII^e siècle : contribution à l'histoire du sentiment linguistique français*, Peeters, Louvain / Paris, 1993.

Tesnière, Lucien, *Éléments de syntaxe structurale*, Klincksieck, Paris, 1959.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574