

# Méthodologies pour la description de quelques phénomènes syntaxiques d'un recueil de langue parlée. Application à l'hébreu moderne

II-II YATZIV- MALIBERT

Thèse dirigée par Claire Blanche-Benveniste

La thèse présente une étude syntaxique et lexicale d'un recueil de langue parlée. Le recueil est composé des transcriptions de douze monologues enregistrés de locuteurs israéliens natifs. La durée de chaque enregistrement varie de 10 minutes à 40 minutes. Les enregistrements ont été effectués par Madame Yatziv-Malibert.

C'est dans l'intention d'utiliser le corpus de langue parlée dans un but de description grammaticale de la langue hébraïque que cette recherche a été entamée. Il s'agit d'affiner la description de faits déjà connus, présents chez tous les locuteurs.

Pour mener cette recherche sur l'hébreu parlé, domaine rarement étudié dans la linguistique hébraïque, il a fallu s'inspirer de méthodes d'analyse appliquées à d'autres langues, notamment au français parlé (les méthodes pratiquées par le Groupe Aixois de Recherche en Syntaxe) et à l'anglais.

La première partie de la thèse aborde la question de savoir pourquoi l'hébreu parlé est si peu étudié. Le premier facteur d'explication est lié à l'histoire particulière de l'hébreu, et notamment à la longue rupture dans l'usage parlé de cette langue alors que le deuxième facteur regroupe des éléments explicatifs plus généraux qui concernent également d'autres langues à savoir la mise à l'écart de l'oral par rapport à la langue écrite et la perception de l'oral comme fautif. Cette partie rassemble des travaux antérieurs sur l'hébreu parlé et quelques recueils existants (Bar-Adon, Blanc et Shatil). De plus elle présente le cadre méthodologique ainsi que les principales unités d'analyse dégagées: les constructions verbales, les constructions non verbales, les divers dispositifs et notamment le dispositif pseudo-clivé étudié plus en détail.

Pour ce faire deux grandes lignes de recherche ont été croisées: nous nous sommes inspirés d'une part des travaux du Groupe Aixois de

Recherche en Syntaxe et d'autre part de la tradition linguistique hébraïque et particulièrement de la réflexion sur les schèmes de phrases de Haïm Rosen et de Maya Fruchtman.

La deuxième partie présente les phénomènes syntaxiques étudiés : le mécanisme paradigmatique, les insertions lexicales et syntaxiques qui perturbent le déroulement syntagmatique des énoncés et les répétitions du lexique.

Dans le chapitre traitant le mécanisme paradigmatique nous avons montré que des phénomènes appartenant à des domaines de nature différente : syntaxe, lexique ou énonciation et des effets tels que erreurs, répétitions intentionnelles ou non, recherche de dénomination ou anticipation peuvent être décrits à l'aide du même instrument de travail qu'est « le listing ». Nous avons pu trouver de fortes ressemblances avec la description de ces phénomènes dans d'autres langues.

Le procédé d'insertion consiste à intercaler une ou plusieurs constructions dans le cours d'une séquence d'énoncés. Sur le plan syntaxique, ce sont des « parasites » qui viennent couper des constructions grammaticales. Sur le plan énonciatif, ce sont des interventions métalinguistiques et notamment des éléments formant des commentaires sur l'énonciation. Ici également la description des faits trouvait son équivalent dans la description des insertions en français parlé.

Le chapitre sur les répétitions du lexique présente trois phénomènes : des reprises de termes à l'identique sans qu'elles provoquent de changement syntaxique, des glissements de lexique qui sont accompagnés de changement de fonction syntaxique et des reformulations de la **racine** autour de la distinction verbe et nom ainsi qu'autour de la notion de **schème**. C'est ce dernier phénomène qui est propre à l'hébreu parlé et probablement à d'autres langues sémitiques.

La troisième partie présente une synthèse des analyses sous forme de grilles, mode de présentation des discours oraux déjà appliqué à d'autres langues. Après avoir défini les unités syntaxiques d'analyse et les structures récurrentes apparaissant dans le corpus de l'hébreu oral, il est proposé de démontrer que, malgré les idées reçues sur le désordre de la langue parlée et les difficultés qui en découleraient pour l'étudier, le discours oral peut être organisé et structuré.

À l'examen des extraits de textes oraux sont apparues des configurations syntaxiques qui se présentent comme un ensemble structuré par plusieurs énoncés et délimité par des répétitions et des modifications de constructions. Ces configurations peuvent servir d'esquisses de classification de genres de discours parlé.

# Étude comparative du fonctionnement des nombres singulier et pluriel dans les langues romanes : portugais, espagnol, catalan et français

Eulàlia VILAGINES SERRA

Thèse dirigée par Claire Blanche-Benveniste

Le phénomène de la « syllepse » a intéressé les grammairiens depuis l'Antiquité. Il s'agit d'un phénomène d'accord apparemment non régulier qui montre un décalage entre l'élément qui donne l'accord et celui qui le reçoit. L'accord de syllepse en nombre<sup>1</sup> est entre *un groupe d'hommes*, considéré comme syntagme nominal singulier, et le verbe *partent*, au pluriel. Diverses attitudes normatives se sont développées pour accepter ou refuser l'accord de syllepse.

On a souvent parlé d'« accord par le sens » plutôt que par la morphologie. Pour certains il s'agit d'un accord avec le complément de nom ou avec un « complément de nom non exprimé ». D'autres posent la notion de perspective selon qu'on prend en compte la forme ou le signifié du collectif. Des analyses plus récentes utilisent les notions de « collectif » et de « quantifieur » et définissent l'accord en fonction d'une analyse interne du syntagme : notion de tête nominale, de tête de syntagme, etc. Ces analyses intègrent des points de vue sémantique et grammatical. J'ai choisi de les utiliser dans une optique lexico-grammaticale pour l'étude de la syllepse.

Les langues romanes considérées, portugais, espagnol, catalan et français, diffèrent par leur lexique et leur grammaire mais elles présentent suffisamment de similitudes pour qu'on puisse les comparer. J'ai constitué un corpus de 4000 exemples attestés écrits et oraux qui se répartissent de la façon suivante : 741 en portugais, 923 en espagnol, 979 en catalan et 1342 en français.

Pour qu'il y ait possibilité de choix entre le singulier et le pluriel, il faut que le « syntagme nominal source » contienne un « quantifieur-

---

<sup>1</sup> La syllepse de genre, la syllepse de personne et la syllepse selon le distributif ne font pas l'objet de cette étude, qui se consacre uniquement à la syllepse de nombre.

collectif». J'ai réuni sous ce nom ce que la grammaire traditionnelle a nommé collectif (type *le personnel*) et quantifieur (type *la plupart*), mais j'ai aussi retenu certains lexèmes qui ne fonctionnent comme des « quantifieurs-collectifs » que dans certaines conditions (type *un chapelet de*).

Les données attestées montrent que ces lexèmes ont différentes réalisations lexico-grammaticales. Parfois, le terme n'accepte pas de complément désignant les parties constitutives qui serait tautologique ( ? *une pinède de pins*). Parfois le complément désignant les parties constitutives devient possible à condition d'être lui-même qualifié (*un jurat d'homes justos* (C), *una familia de parientes bienvenidos* (E)). Quelques-uns acceptent facilement la présence de ce complément désignant les parties constitutives (*un groupe d'amis, la majorité d'amis*). Il existe, enfin, des lexèmes qui ne peuvent se réaliser qu'accompagnés du complément désignant les parties constitutives (*un chapelet de bombes* (F), *un rio de curiosos* (E)). La présence ou absence de ce complément s'est révélée importante vis-à-vis du choix de l'accord.

D'autre part, en fonction de ces différents types d'emplois en tant que collectifs ou quantifieurs, j'ai proposé une typologie des termes. *Le corps médical* ou *le public* sont des collectifs qui ne deviennent jamais quantifieurs ; les expressions de pourcentage, *25%*, ou désignant des parties restent des expressions de quantité qui n'ont jamais un emploi de collectif ; mais il ne manque pas de lexèmes qui fonctionnent tantôt comme des collectifs (*une foule, une armée*), tantôt comme des quantifieurs (*une foule de métamorphoses, une armée d'engins*).

Par ailleurs, chaque terme manifeste des tendances d'accord. *Le reste des enfants*, en français, choisit nettement le singulier, alors que *la plupart des enfants* le pluriel. Des lexèmes comme *la majorité*, dans les quatre langues, ou *(l)a gent(e)*, en portugais, en espagnol et en catalan, naviguent entre les deux.

J'ai examiné les cibles d'accord dans différentes situations syntaxiques, en commençant par la construction directe sujet + verbe. Sur l'ensemble des lexèmes retenus pour faire la comparaison, lorsque le syntagme nominal source remplit la fonction sujet, au moins un tiers peuvent provoquer des accords de pluriel bien qu'ils aient une morphologie de singulier. Mais la proportion varie selon les langues : le portugais et l'espagnol ont seulement un tiers des exemples avec accord au pluriel ; le français 40%, et le catalan 50%. Le simple fait d'ajouter un « dispositif » du type *il y a ... qui* implique pour les quatre langues un changement des tendances d'accord.

Dans les relatives, source et cible se trouvent dans un même grand constituant (*Une génération qui ont perdu*). Pour les quatre langues, le pluriel est l'accord choisi dans 60% des cas.

La cible peut être un verbe « voisin », c'est-à-dire le verbe d'une nouvelle construction (*Elle voit un groupe. Ils s'arrêtent*) ou un verbe de la même construction (*Il regarde la famille quand ils mangent*). Le choix du pluriel s'impose alors, laissant peu de place à l'accord au singulier.

Il en va de même lorsqu'il s'agit d'un accord par inférence entre deux éléments nominaux : le singulier est très rare (*C'est une famille. Je les vis monter*).

Une partie importante de mon corpus atteste qu'un syntagme nominal source peut être concerné par plusieurs cibles, par exemple par un verbe recteur et par un verbe « voisin » (*Ce couple est très sympathique. Ils m'ont accueilli*).

Le paramètre du contexte syntaxique s'est révélé plus important que je ne le pensais au départ : j'ai pu observer un glissement du choix du singulier dans les constructions directes sujet + verbe vers le pluriel dans le cas des relatives, qui s'accroît quand il s'agit des accords avec les verbes « voisins » et les accords par inférence entre deux éléments nominaux.

Je propose une structure lexico-grammaticale des termes analysés, c'est-à-dire, l'ensemble des possibilités que chacun des termes possède. La structure sera établie à partir du lien entre lexique et grammaire. Par exemple, voici la structure typique de « foule » :

1. une foule va venir
2. une foule ? vont venir<sup>2</sup>
3. une foule de personnes va venir
4. une foule de personnes vont venir
5. une foule d'exemples va arriver
6. une foule d'exemples vont arriver
7. une foule ---- va venir<sup>3</sup>
8. une foule ---- vont venir

Faut-il en conclure que la syllepse est un cas de variation ? Il est vrai que dans certains cas oui, mais on observe que le choix du singulier ou du pluriel est significatif dans d'autres cas. D'autre part, il a fallu aussi tenir

<sup>2</sup> Le signe ? marque le doute sur l'existence d'une des possibilités.

<sup>3</sup> Le trait ---- rappelle qu'il n'y a pas de tête lexicale.

compte de processus de grammaticalisation. Certains lexèmes, comme *la foule*, sont encore en évolution, d'autres, comme *la plupart*, sont arrivés à la fin de ce processus.

L'étude de la syllepse, enfin, oblige à raffiner la notion de nombre. Le système du nombre est plus complexe que l'opposition singulier-pluriel à laquelle nous sommes habitués, comme le montrent des études récentes. Dans le cas de l'accord de syllepse, il faut tenir compte du cryptotype « collectif ».

## Le pronom anaphorique et l'architecture de l'oral en finnois et en français

Outi TIAINEN-DUVALLON

Thèse dirigée par Claire Blanche-Benveniste

Ce travail étudie l'emploi des pronoms anaphoriques du point de vue de la structuration syntaxique des textes oraux. L'objectif est de décrire, d'une part, l'émergence des constructions syntaxiques dans les productions orales et de montrer, d'autre part, comment l'emploi et l'interprétation des pronoms de troisième personne sont intimement liés à l'organisation syntaxique des textes. La base de données est un recueil d'enregistrements de conversations en finnois et en français.

L'étude se concentre sur les pronoms finnois *se* 'il/elle ~ ce' et *hän* 'il/elle' en position préverbale et sur les pronoms sujets français de type *il*. Ces pronoms sont présentés comme des expressions référentielles originales, non tributaires des autres. Au sein des systèmes pronominaux des deux langues, ils sont dotés de certaines propriétés inhérentes. Les pronoms français expriment le trait sémantique [ $\pm$ classifié] et les pronoms finnois le trait [ $\pm$ logophorique]. Ces traits font qu'il y a des divergences dans le choix des pronoms dans les contextes locaux, mais ils ne semblent pas peser sur l'emploi des pronoms comme des contraintes non locales. Dans les deux langues, les pronoms en question sont aptes à des emplois anaphoriques.

Les analyses syntaxiques s'inscrivent dans le cadre théorique de l'Approche Pronominale (Blanche-Benveniste *et al.*, 1987). La description est fondée sur la notion de verbe. Le verbe est considéré comme un élément recteur qui ouvre des places de construction autour de lui. L'analyse des données commence par l'examen des procédés par lesquels les constructions verbales sont mises en place dans les textes oraux des deux langues. L'attention est portée, d'une part, à l'émergence de l'axe syntagmatique des énoncés verbaux et, d'autre part, aux réalisations paradigmatiques des places de rection. Ainsi, l'étude intègre à la description syntaxique des phénomènes de « production » comme les réparations d'énoncés et différents types d'insertions qui dédoublent la

chaîne parlée ou qui suspendent momentanément le déroulement d'un énoncé cadre.

Les insertions parenthétiques sont conçues comme une troisième dimension textuelle, c'est-à-dire comme un site à partir duquel les locuteurs peuvent gérer les opérations en cours sur l'axe syntagmatique ou paradigmatique de l'énoncé cadre. L'examen des passages de texte plus longs met l'accent sur le phénomène de récursivité des moules syntaxiques. Les rééditions d'un verbe recteur et de sa construction forment dans les textes oraux des unités de structuration, appelées *configurations syntaxiques*. Pour mettre en valeur ces organisations structurelles, le travail applique aux extraits de corpus le mode de présentation « en grilles syntaxiques ». L'analyse des données permet de voir que les mécanismes de production de textes oraux se ressemblent dans les deux langues, qui possèdent pourtant des micro-syntaxes assez différentes.

Les pronoms sont examinés sous un double rôle : d'une part, ce sont des indicateurs syntaxiques, plus sûrs que les formes lexicales des noms, qui permettent d'identifier les places de rection autour d'un verbe, et, d'autre part, ce sont des expressions référentielles, dotées d'un minimum de charge descriptive. Le travail défend l'idée selon laquelle les pronoms de troisième personne sont des expressions référentielles non marquées qui, face à l'instabilité des désignations lexicales, permettent de saisir un référent d'une manière plus stable. Les exemples analysés montrent que les pronoms peuvent être utilisés non seulement là où les désignations lexicales sont déjà mises au point, mais aussi là où elles sont encore en cours, en train d'être négociées ou même momentanément indisponibles.

L'étude décrit les diverses façons dont l'ancrage référentiel des pronoms anaphoriques se détermine dans le contexte linguistique caractérisé par la complexité syntaxique. L'interprétation référentielle des pronoms est cruciallement guidée par les informations véhiculées par leur construction hôte et par la position de celle-ci dans l'organisation structurelle du texte. Pour les pronoms utilisés dans les séquences insérées, l'énoncé cadre constitue un domaine d'interprétation particulièrement saillant. Pour les pronoms dont la construction hôte forme une configuration syntaxique avec d'autres séquences dans le texte, c'est la configuration syntaxique qui leur fournit un précieux domaine d'interprétation.

Le travail souligne le fait que les domaines d'interprétation formés par les configurations syntaxiques ne sont pas toujours des espaces textuels linéaires. La distance entre le pronom et sa source d'interprétation dans le texte ou la simple présence dans le contexte linguistique de plusieurs



sources d'interprétation sémantiquement compatibles avec le pronom  
auraient moins d'importance sur l'emploi des pronoms anaphoriques que  
l'on ne l'a souvent pensé.