

EXPÉRIMENTATION ET MODÉLISATION EN LINGUISTIQUE

Benjamin STORME
Université de Leiden
b.p.p.storme@hum.leidenuniv.nl

RÉSUMÉ

Cet article présente et motive l'utilisation de deux méthodes issues des sciences naturelles pour la linguistique : l'expérimentation et la modélisation. L'expérimentation vise à tester une hypothèse de façon contrôlée et est particulièrement adaptée pour statuer sur les liens de causalité entre variables linguistiques. La modélisation linguistique vise quant à elle à proposer une analyse entièrement explicite d'un phénomène dans un modèle statistique et permet de faire des prédictions précises tout en étant robuste à la variabilité inhérente aux données linguistiques. L'article montre comment ces approches permettent d'éclairer deux débats actuels en linguistique : la question des motivations de l'écriture inclusive et la question du rôle des principes fonctionnalistes dans le façonnement des systèmes linguistiques.

expérimentation, modélisation, syncrétisme, genre grammatical, écriture inclusive, liaison

ABSTRACT

This article presents and motivates the use of two methods from the natural sciences in linguistics: experiments and modeling. Experiments provide a way to test a hypothesis in a highly controlled setting and are particularly suited to identify causal relations among linguistic variables. Modeling aims to propose an explicit analysis of a phenomenon in a statistical model and makes it possible to make precise predictions while being robust to the variability inherent in real-world linguistic data. The paper shows how these methods can help answer two central questions in current linguistic debates: the question of the motivations for gender-inclusive language and the question of the role of functionalist principles in shaping linguistic systems.

experiments, modeling, syncretism, grammatical gender, gender-inclusive language, French liaison

1. LA LINGUISTIQUE COMME SCIENCE DU LANGAGE

La linguistique est souvent définie comme la science du langage. Cette définition est parfois prise dans un sens large qui identifie la linguistique à l'ensemble des savoirs sur le langage et les langues. Mais de plus en plus, cette définition est prise dans un sens plus étroit désignant une approche spécifique, à savoir la démarche hypothético-déductive, et des méthodes bien particulières d'élaboration du savoir héritées des sciences de la nature, incluant l'expérimentation et la modélisation.

L'expérimentation en linguistique vise à tester une hypothèse de façon contrôlée, au moyen de stimuli soigneusement construits, et se distingue d'une approche plus observationnelle s'appuyant sur des corpus de données naturelles. La modélisation linguistique quant à elle vise à proposer une analyse entièrement explicite d'un phénomène linguistique dans un modèle statistique et se distingue des analyses plus

traditionnelles en linguistique, qui sont typiquement formulées de façon informelle et qualitative.

Le but de cet article est de motiver et illustrer l'intérêt, pour l'étude des langues, de l'expérimentation et de la modélisation. Ces méthodes seront présentées à travers deux thèmes de recherche sur lesquels j'ai travaillé récemment, et qui sont centraux dans les débats de société et les débats scientifiques actuels autour du langage : la question des motivations de l'écriture inclusive et la question de la place du fonctionnalisme en linguistique. Le but sera de montrer comment l'expérimentation et la modélisation permettent d'apporter des réponses rigoureuses à ces questions.

La partie 2 de cet article est consacrée à une présentation générale de l'expérimentation et de la modélisation, où j'explique pourquoi ces méthodes sont particulièrement adaptées respectivement pour évaluer les rapports de causalité postulés dans les théories linguistiques et pour rendre compte de la variation observée dans les données linguistiques.

Les parties 3 et 4 portent sur le premier grand thème abordé dans cet article : l'écriture inclusive. Cette question de recherche sera utilisée principalement pour illustrer l'intérêt de l'expérimentation pour statuer sur la causalité.

La partie 3 introduit l'une des questions centrales autour de l'écriture inclusive : est-ce que le genre grammatical d'un mot influe sur les inférences que font les locuteurs sur le genre social de son référent ? On verra que cette question est complexe car de nombreux facteurs au-delà du genre grammatical peuvent avoir un effet sur ces inférences, comme les stéréotypes de genre ou les effets de compétition entre des formes linguistiques alternatives.

La partie 4 présentera quelques études expérimentales qui tentent de déterminer le rôle du genre grammatical dans les inférences de genre social tout en contrôlant ces facteurs d'explication alternatifs.

Les parties 5 et 6 portent sur le deuxième grand thème empirique abordé dans cet article : la question des explications fonctionnalistes en linguistique. Cette question de recherche sera utilisée principalement pour illustrer l'intérêt de la modélisation. Mais on verra que l'expérimentation a également un rôle à jouer pour établir la causalité des principes fonctionnalistes sur la forme linguistique.

La partie 5 contextualisera cette question et présentera un modèle qui a été proposé à partir des années 1990 pour formaliser les théories fonctionnalistes, à savoir le modèle des grammaires à contraintes. Cette partie présentera également les implémentations probabilistes de ces modèles qui ont été proposées à la fin des années 90 et dans les années 2000 pour rendre ces modèles robustes à la variation inhérente aux données linguistiques.

La partie 6 présentera deux études de cas tirées de mes propres recherches qui illustrent la pertinence des modèles de grammaires à contraintes pour étudier à la fois la variation à travers les langues, avec une étude typologique sur le syncrétisme morphologique, et la variation à l'intérieur des langues, avec une étude sur la réalisation variable de la liaison en français. Cette dernière étude a été choisie spécifiquement parce qu'elle combine les deux méthodologies mises en avant dans cet article, l'expérimentation et la modélisation.

La partie 7 conclut en récapitulant les enseignements méthodologiques de cet article sur l'expérimentation et la modélisation ainsi que les apports théoriques sur l'écriture inclusive et le fonctionnalisme.

2. POURQUOI L'EXPÉRIMENTATION ET LA MODÉLISATION EN LINGUISTIQUE ?

2.1. Données et théories en linguistique

L'entreprise scientifique consiste à prendre d'un côté des données, à savoir des faits observables, et de l'autre des théories, qui permettent d'expliquer et prédire ces données¹. Par exemple, la théorie de la sélection naturelle en biologie prend pour données les propriétés biologiques observées des espèces. Et elle en rend compte à partir d'une théorie : cette théorie postule un mécanisme de tri de la diversité génétique des individus par l'environnement, les individus les plus adaptés à leur milieu ayant de plus grandes chances de se reproduire avec succès.

La linguistique fonctionne de la même façon, avec des données d'un côté et des théories de l'autre. Les données peuvent être des textes, des enregistrements de parole, des mesures de temps de réaction face à un stimulus linguistique ou encore des mesures d'activation des zones du cerveau.

Les théories, quant à elles, peuvent inclure des cadres théoriques très généraux, comme la théorie des exemplaires, qui postule que le savoir linguistique d'un individu consiste en une multitude d'expériences linguistiques mémorisées (les « exemplaires »)², ou la théorie générative, qui postule un savoir linguistique beaucoup plus abstrait, et en partie inné, fait d'un inventaire de sons, d'un lexique et de règles de combinaison de ces unités sonores et lexicales³. Plus modestement, on peut aussi appeler théories des généralisations descriptives beaucoup plus spécifiques, comme la généralisation selon laquelle le latin a six cas (nominatif, vocatif, accusatif, génitif, datif, ablatif) ou que le français a des voyelles nasales (comme *on* ou *an*). Les données sont concrètes et observables alors que les théories sont abstraites et ne sont pas directement observables.

Le linguiste utilise les données linguistiques pour tester ou proposer une théorie sur les principes qui façonnent les langues et leur usage. On distingue deux grands modes d'interaction entre données et théories : l'approche inductive et l'approche déductive. Dans l'approche dite inductive, on part des données pour générer des théories ou des généralisations. Par exemple, l'observation que la deuxième déclinaison latine présente une seule forme en *-o* (par exemple *domino* « maître ») dans des fonctions où les noms de la troisième déclinaison en ont deux (par exemple *regi* et *rege* « roi ») conduira le linguiste à poser deux cas abstraits, le datif et l'ablatif, exprimés de façon ambiguë dans la deuxième déclinaison et non ambiguë dans la troisième. À l'inverse, la présence de deux formes en *-us* et *-e* dans la deuxième déclinaison dans les fonctions de sujet de la phrase et dans les apostrophes là où la troisième déclinaison n'en utilise qu'une (par exemple *rex* « roi ») conduira le linguiste à distinguer deux cas, nominatif et vocatif, qui sont exprimés de façon non ambiguë dans la deuxième déclinaison mais de façon ambiguë dans la troisième. Ainsi, alors que les deux déclinaisons ne présentent que cinq formes distinctes au singulier, ces cinq formes correspondent à six cas abstraits. Il s'agit bien ici d'une généralisation abstraite et non directement observable que l'on dégage à partir de la comparaison minutieuse des formes et de leurs contextes d'emploi.

Dans l'approche déductive, on part des théories et on les teste à partir de nouvelles données. Pour reprendre l'exemple du latin, une fois que l'on a établi de façon inductive

1. Frank *et al* (2023 : chapitre 2).

2. Pierrehumbert (2001).

3. Chomsky & Halle (1968).

que les deuxième et troisième déclinaisons ont six cas, on peut se demander de façon déductive si la première déclinaison est structurée de la même façon, et si donc notre théorie des cas du latin est généralisable à d'autres cas. En collectant les données issues de textes, on se rendrait compte que l'on observe quatre formes distinctes pour la première déclinaison (*vita*, *vitam*, *vitae*, *vitā*), mais que deux d'entre elles sont observées avec deux fonctions grammaticales distinctes (*vita*, qui apparaît dans les emplois du nominatif et du vocatif, et *vitae*, qui apparaît dans les emplois du génitif et du datif). Ces données permettraient donc de valider la théorie des six cas, combinée à une théorie permettant à plusieurs cas d'être exprimés par une forme unique dans certains contextes (nominatif et vocatif d'un côté, génitif et datif de l'autre).

Quelle est la place de l'expérimentation et de la modélisation dans cette démarche qui relie données et théories ? Comme on le verra ci-dessous, l'expérimentation est une façon spécifique de collecter des données, qui permet de mieux cerner les liens de causalité entre les concepts linguistiques posés par le théoricien (partie 2.2). La modélisation est quant à elle une façon spécifique de formuler des théories, qui permet de faire des prédictions précises par rapport aux données que l'on peut s'attendre à observer (partie 2.3).

2.2. Expérimentation et causalité

2.2.1. Une typologie schématique des données linguistiques

En linguistique, on utilise plusieurs types de données que l'on peut classer de façon schématique en fonction de trois critères : la nature intrinsèque des données, le degré de contrôle opéré par le linguiste dans leur collecte et le degré de conscience du participant qui fournit les données.

Concernant la nature des données, on peut avoir affaire à des données comportementales (comme des productions textuelles ou orales, des temps de lecture, des mouvements oculaires) ou des données cérébrales (comme des images obtenues par des méthodes d'imagerie par résonance magnétique). Concernant le critère du degré de contrôle, on a, à un bout du spectre, des données complètement naturelles issues par exemple d'émissions radiophoniques, sans aucune intervention du chercheur. À l'autre bout du spectre se trouvent les données expérimentales, qui se caractérisent par un important contrôle de la part du linguiste. Par exemple, le linguiste peut contrôler très précisément les mots et structures de phrases qui vont être produites en faisant lire des phrases à un locuteur. Enfin, le troisième critère réfère au caractère plus ou moins conscient du locuteur fournissant les données : les données peuvent être observées, sans que le locuteur en ait aucune conscience, ou les données peuvent être déclarées, auquel cas le locuteur fournit un jugement conscient ou métalinguistique sur une expression linguistique.

Traditionnellement, les données dites « de corpus » jouent un rôle majeur en linguistique. Pour reprendre la typologie présentée ci-dessous, les données de corpus sont des données comportementales (souvent sous forme de textes), naturelles (produites dans un contexte naturel, sans intervention du linguiste) et observées (on observe le comportement linguistique directement, sans passer par le jugement métalinguistique des locuteurs). Les textes latins mentionnés plus haut dans la discussion sur les six cas grammaticaux du latin constituent des données de corpus.

Il y a des raisons historiques pour la prévalence des données de corpus en linguistique, comme l'importance de la linguistique historique dans le développement de la linguistique comme science au XIX^e siècle, l'orientation traditionnelle des descriptions

grammaticales vers les «grands auteurs» et, plus récemment, l'essor des méthodes d'intelligence artificielle qui s'appuient sur de grandes quantités de données textuelles. Mais il y a aussi des raisons scientifiques plus profondes, dans la mesure où les corpus de données naturelles reflètent le mieux les conditions naturelles d'utilisation du langage, sans intervention du linguiste.

2.2.2. Le problème de la causalité

Cependant les données naturelles présentent une limite : elles ne permettent pas aussi facilement de statuer sur les liens de causalité que les données obtenues dans un contexte expérimental. Or la causalité est centrale dans les théories scientifiques, y compris dans les théories linguistiques.

Ce point peut être illustré au moyen d'un exemple parlant, bien que non linguistique, à savoir l'idée selon laquelle l'argent fait le bonheur⁴. Cette idée peut se concevoir comme une théorie du bonheur, qui postule un mécanisme causal entre une propriété patrimoniale d'un individu (sa richesse) et une propriété psychologique de ce même individu (son état de bonheur). Comment tester cette théorie en suivant l'approche déductive décrite plus haut ? Imaginons un scénario où l'on aurait accès à toutes les données personnelles en ligne des individus résidant sur un territoire donné. Par exemple, on aurait accès à leurs fiches d'impôts, ce qui permettrait d'établir leur richesse. On aurait également accès à des informations nous permettant d'inférer leur degré de bonheur, comme des avis de satisfaction en ligne ou encore des courriels. Supposons que l'on observe dans cet ensemble de données une corrélation positive entre richesse estimée et bonheur estimé. Pourrait-on pour autant conclure que la richesse cause le bonheur et ainsi valider notre théorie du bonheur ?

Non, et ce pour de nombreuses raisons. D'abord, les indicateurs de richesse et de bonheur utilisés ne reflètent peut-être pas la richesse réelle et le bonheur réel des individus. On pourrait imaginer par exemple que certains individus dans notre échantillon sont parvenus à échapper à l'impôt en cachant une partie de leur richesse dans les paradis fiscaux. Ensuite, les avis sur internet ou les courriels ne donnent sans doute qu'une vision très parcellaire de l'état psychologique des individus. Par ailleurs, on pourrait se demander si l'échantillon étudié est représentatif de l'ensemble de l'humanité.

Mais il y a une autre raison qui nous intéresse plus particulièrement dans le contexte d'une discussion sur l'expérimentation : c'est la différence bien connue entre causation et corrélation. Même si les indicateurs mesuraient fidèlement la richesse et le bonheur et que l'échantillon utilisé était représentatif de la population globale, on ne pourrait pas conclure à partir de ces données à un lien de causalité de la richesse vers le bonheur (richesse \rightarrow bonheur).

Il existe en effet des scénarios de causalité alternatifs qui sont compatibles avec une corrélation positive entre richesse et bonheur. Par exemple, la causalité pourrait aller du bonheur vers la richesse (bonheur \rightarrow richesse). On pourrait en effet imaginer que les gens plus heureux ont aussi tendance à mieux réussir dans la vie, et par conséquent à être plus riches. Dans l'hypothèse selon laquelle c'est le bonheur qui cause la richesse (et non l'inverse), on s'attendrait également à une corrélation positive entre bonheur et richesse, mais avec une causalité inversée par rapport au premier scénario. Par ailleurs, on pourrait également imaginer un autre scénario explicatif impliquant un facteur tiers qui causerait indépendamment plus de bonheur et plus de richesse, sans qu'il y ait de

4. Cet exemple est emprunté à Frank *et al* (2023 : chapitre 1).

lien de causalité direct entre richesse et bonheur. Par exemple, il pourrait s'agir de l'ouverture d'esprit, qui causerait indépendamment plus de bonheur (ouverture d'esprit → bonheur) et plus de richesse (ouverture d'esprit → richesse). On parle dans ce cas de facteur de confusion (*confounding factor* en anglais). On en verra des exemples dans l'étude sur l'écriture inclusive (voir la partie 4).

Comment savoir donc si la richesse cause le bonheur ? Idéalement, il faudrait pouvoir répliquer le monde réel, et avoir ainsi deux mondes, un monde 1 et un monde 2. Dans le monde 1, on introduirait une manipulation minimale de la richesse des individus, en leur donnant par exemple une somme fixe, mais en ne changeant rien d'autre. Dans le monde 2, on ne ferait rien. On observerait ensuite comment le bonheur évolue dans les mondes 1 et 2. Si la richesse cause le bonheur, on devrait alors observer un bonheur moyen plus élevé dans le monde 1 par rapport au monde 2. Il serait alors possible de conclure car l'effet des autres variables aurait été *contrôlé* grâce à l'étude d'individus identiques dans les deux mondes à leur richesse près.

Bien sûr, en pratique, il est impossible de reproduire le monde à l'identique et de n'en manipuler qu'une seule propriété. L'approche expérimentale cherche à se rapprocher au maximum de ce scénario idéal. Dans une expérience, le monde 1 où l'on introduit une manipulation s'appelle le groupe «test». L'autre groupe s'appelle le groupe «contrôle». Comme on ne peut pas avoir exactement les mêmes individus dans les deux groupes, on procède en général à un échantillonnage aléatoire : on va prendre *au hasard* dans la population des individus qui seront assignés au groupe contrôle et d'autres qui seront assignés au groupe test. Si l'échantillon est suffisamment large, les deux groupes devraient être assez proches dans leur composition, avec des proportions semblables de gens heureux, riches, et ouverts d'esprit. Ainsi, si l'on observe après avoir donné une somme d'argent au groupe test une augmentation significative du bonheur par rapport au groupe contrôle, alors on pourra conclure que c'est bien l'augmentation de la richesse qui est en cause, les autres variables (le niveau de bonheur et l'ouverture d'esprit) n'ayant pas été manipulées. L'étude observationnelle ne contrôle pas aussi bien ces facteurs d'explication alternatifs et ne permet donc pas de conclure à une causalité.

Les données de corpus utilisées traditionnellement en linguistique posent encore un autre problème, qui peut rendre certaines hypothèses difficiles à tester. Les données naturelles ont en effet une structure particulière, décrite par la loi de Zipf, du nom du scientifique qui a le premier décrit cette structure dans la première moitié du xx^e siècle (George K. Zipf). D'après la loi de Zipf, un petit nombre d'expressions linguistiques (des mots, des structures de phrases particulières) ont une fréquence d'occurrence très élevée par rapport aux autres expressions linguistiques⁵. La loi de Zipf implique qu'il faut des quantités de données très importantes pour étudier un phénomène rare.

Par exemple, supposons que l'on veuille vérifier à partir d'un corpus si le latin a bien six cas pour tous ses noms. À cause de la structure zipfienne du lexique, certains mots vont apparaître de nombreuses fois sous toutes leurs formes casuelles tandis que les mots rares n'apparaîtront peut-être que sous quelques formes casuelles⁶. Pour ces mots, il sera difficile d'affirmer avec certitude que toutes les désinences casuelles sont disponibles. Dans un contexte expérimental où les données sont contrôlées par le linguiste, ce problème ne se pose pas car l'expérimentateur peut amener un locuteur à produire toutes les formes casuelles d'un mot, même si celui-ci est rare, en lui faisant

5. Zipf (1949 : chapitre 2).

6. Lignos & Yang (2016 : 769).

par exemple lire des phrases contenant les formes voulues ou en créant des contextes appropriés pour que ces formes soient produites.

2.2.3. Étude de cas : expérience d'apprentissage d'une langue artificielle

Illustrons plus concrètement l'intérêt de l'expérimentation en linguistique pour démêler la causalité de la corrélation. Nous allons regarder une étude qui porte sur le fonctionnalisme, puisque c'est un thème sur lequel nous reviendrons plus tard (voir parties 5 et 6). Il s'agit d'une étude expérimentale dont le but général est d'établir si la forme d'un système linguistique est causée par sa fonction⁷.

L'objet d'étude spécifique est le marquage différentiel du cas. Dans certaines langues, comme le latin, la fonction grammaticale d'un groupe nominal (sujet, objet, etc.) est marquée morphologiquement par un cas. Mais dans de nombreuses langues, ce marquage n'est pas systématique et peut être conditionné par des propriétés sémantiques des noms. En particulier, les langues tendent davantage à marquer l'objet avec une forme spécifique quand cet objet dénote un animé que lorsqu'il dénote un inanimé. Cette tendance peut être illustrée avec l'espagnol, où l'objet du verbe est précédé de la préposition *a* avec des noms dénotant des animés, comme dans l'exemple (1a), mais pas avec des noms dénotant des inanimés, comme dans l'exemple (1b), où la préposition est agrammaticale. Il y a donc une asymétrie dans le marquage de l'objet en espagnol : la marque explicite est utilisée pour marquer les objets animés (la préposition *a*) mais pas pour marquer les objets inanimés. Ceux-ci sont en effet exprimés sans préposition ni suffixe casuel, comme les sujets.

- 1 a. *Pedro* *besó* *a* *Lucía*
 Pierre embrasser.Passé Prép Lucia

«Pierre a embrassé Lucie.»

- b. *Pedro* *besó* (**a*) *el* *retrato*
 Pierre embrasser.Passé Prép le portrait

«Pierre a embrassé le portrait.»

La théorie fonctionnaliste explique l'asymétrie entre objet animé et objet inanimé par une différence dans le potentiel de désambiguïsation du marquage de l'objet. Les sujets sont typiquement animés, puisque les verbes transitifs renvoient en général à des actions réalisées par des animés (par ex. «embrasser»). Quand une phrase avec un verbe transitif contient un nom animé et un nom inanimé, comme *Pedro* «Pierre» et *el retrato* «le portrait» dans (1b), on peut généralement inférer les rôles grammaticaux ou sémantiques correctement sans qu'il y ait besoin de les marquer explicitement : le nom animé sera généralement le sujet du verbe (ou, en termes sémantiques, l'agent de l'action) et le nom inanimé l'objet du verbe (ou, en termes sémantiques, le patient de l'action). Cependant, dans le cas où le verbe transitif a deux noms animés comme arguments, comme *Pedro* «Pierre» et *Lucía* «Lucie» dans (1a), il y a une ambiguïté car on ne peut pas savoir en général qui a fait l'action sur qui. Par exemple, *Pedro* pourrait aussi bien embrasser *Lucía* que *Lucía Pedro*. Ainsi, d'un point de vue communicatif, il serait plus important de marquer l'objet pour les animés que pour les inanimés.

7. Fedzechkina, Jaeger & Newport (2012).

Il faut noter qu'en espagnol l'ordre des mots permet également de désambiguïser, mais ce n'est pas le cas de toutes les langues avec un marquage différentiel de l'objet. Par exemple, le latin a un marquage différentiel de l'objet dans le sens où les noms masculins et féminins ont des formes distinctes pour le nominatif et l'accusatif mais pas les neutres. En latin, l'ordre est moins fixe, et donc l'ambiguïté plus problématique. Le fait que l'ordre des mots permette aussi de désambiguïser en espagnol n'est pas forcément un problème pour la théorie fonctionnaliste, dans la mesure où il y a un avantage communicatif à marquer l'information de façon redondante⁸.

Comment tester si l'emploi du marquage différentiel de l'objet est causé par son potentiel de désambiguïstation ? Dans une étude publiée dans les *Proceedings of the National Academy of Sciences*, Fedzechkina et ses co-auteurs proposent de le faire de façon expérimentale, c'est-à-dire en cherchant à contrôler au maximum les facteurs explicatifs alternatifs. Ils utilisent le paradigme expérimental de l'apprentissage d'une langue artificielle. Ce paradigme est particulièrement intéressant parce qu'il permet de contrôler précisément les propriétés de la langue, puisque la langue en question est inventée pour les besoins de l'expérience. Le but de l'étude est de tester si un locuteur qui apprend une langue avec un marquage du cas qui n'est pas optimal d'un point de vue fonctionnel (c'est-à-dire qui ne respecte pas le principe du marquage différentiel du cas décrit plus haut) a tendance à remodeler cette langue pour la rendre plus efficiente. Autrement dit, est-ce que l'apprenant tend à conditionner le marquage casuel sur le caractère animé/inanimé des noms ?

Rentrons un peu plus dans le détail du *design* expérimental de cette étude. Sur une durée de quatre jours, les participants à l'expérience ont été exposés à une langue artificielle ayant les propriétés suivantes. Cette langue a un ordre des mots variable, sujet-objet-verbe (SOV) dans 60 % des phrases et objet-sujet-verbe (OSV) dans les 40 % restants. La variabilité de l'ordre des mots a pour conséquence que l'ordre des mots ne peut pas être utilisé pour désambiguïser, comme en latin : tout le poids de la désambiguïstation va ainsi reposer sur le marquage morphologique des fonctions grammaticales (les cas). Toutes les phrases décrivent une action réalisée par un agent animé, mais le patient de l'action (donc l'objet d'un point de vue grammatical) est soit animé (il s'agit par exemple une personne) soit inanimé (il s'agit par exemple d'une chaise). Cette manipulation permet de limiter l'ambiguïté aux cas où le verbe a deux arguments animés. Quand un des arguments est inanimé, on sait forcément qu'il s'agit du patient. Enfin, troisième manipulation importante : l'objet est marqué de façon variable (dans 60 % des phrases), mais ce marquage n'est pas conditionné par le caractère animé ou inanimé de l'objet : un objet inanimé a autant de chances d'être marqué qu'un objet animé. La variabilité du marquage permet aux participants d'envisager deux options de marquage (présence ou absence de marquage), aussi bien pour les animés que pour les inanimés. Ce qui est crucial, c'est que l'absence de conditionnement du marquage de l'objet sur le caractère animé/inanimé des arguments est sous-optimale du point de vue de la communication : en effet, marquer l'objet inanimé n'a pas d'intérêt communicatif fort, dans la mesure où le rôle d'objet peut être inféré directement du fait que l'argument est inanimé (les inanimés ne sont jamais sujets). En revanche, marquer l'objet animé a un intérêt car cela permet de savoir lequel des deux animés est le sujet et lequel est l'objet. En l'absence d'ordre des mots fixe, il y a une ambiguïté quand l'objet est animé mais pas quand il est inanimé. L'hypothèse fonctionnaliste fait une prédiction :

8. Mahowald *et al* (2023).

les participants devraient avoir tendance à plus marquer l'objet quand il est animé que lorsqu'il est inanimé s'ils cherchent à réduire l'ambiguïté pour l'interlocuteur.

L'expérience contient trois phrases : une phase d'exposition, une phase de compréhension et une phase de production⁹. Dans la phase d'exposition, des participants de langue maternelle anglaise entendent et répètent des phrases de la langue cible qui décrivent une action apparaissant dans une vidéo et impliquant un agent et un patient. Dans la phase de compréhension, les participants entendent des phrases de la langue cible et, pour chaque phrase entendue, doivent indiquer parmi deux images qui leur sont présentées laquelle correspond à l'agent de l'action dénotée par cette phrase. Dans la phase de production, les participants voient une action dans une vidéo et doivent décrire cette action dans la langue cible. Cette dernière étape est la plus importante : elle permet de voir comment les participants ont appris la langue et s'ils l'ont remodelée en accord avec l'hypothèse fonctionnaliste, c'est-à-dire en marquant plus fréquemment les objets animés par rapport aux objets inanimés. Les chercheurs mesurent la proportion de phrases dans les productions des participants qui présentent un marquage de l'objet, en distinguant les cas où l'objet est animé et où il est inanimé.

Les résultats montrent qu'au fur et à mesure de l'apprentissage les participants apprennent à marquer l'objet et se rapprochent en moyenne de la proportion de 60 % de marquage de l'objet de la langue cible. Mais, et c'est le résultat le plus intéressant, les objets animés et inanimés sont traités différemment : les objets animés sont plus souvent marqués (et donc exprimés d'une façon non ambiguë par rapport aux sujets animés) que les objets inanimés, en accord avec les prédictions de l'hypothèse fonctionnaliste. Les participants remodelent la langue qu'ils apprennent pour la rendre plus optimale, économisant leur effort quand il n'y a pas d'ambiguïté (avec les objets inanimés) et ajoutant une marque morphologique explicite en cas d'ambiguïté (avec les objets animés).

L'apport de l'expérimentation devrait apparaître assez clairement ici. Elle permet de contrôler très précisément les propriétés de la langue qui nous intéressent, comme l'ordre des mots ou les contextes d'ambiguïté. Par ailleurs, les chercheurs font une manipulation minimale qui implique des phrases avec ou sans ambiguïté sur les fonctions syntaxique (sujet/objet). Cela permet de statuer sur le lien de causalité entre fonction (désir d'éviter l'homophonie) et forme (marquage de l'objet).

Cependant il est toujours difficile de contrôler tous les facteurs qui pourraient influencer sur un résultat et une expérience seule ne suffit jamais pour établir un lien causal. Dans l'étude que l'on vient de présenter, il y a un paramètre qui n'a pas été suffisamment contrôlé et qui pourrait être problématique pour l'interprétation des résultats. Les participants étaient anglophones. Or l'anglais a un marquage différentiel du cas sur les pronoms personnels : l'objet a un marquage différentiel par rapport au sujet pour les humains (la fonction sujet, *he/she*, a un marquage différent de la fonction objet, *him/her*) mais pas pour les inanimés (*it* est utilisé pour le sujet et l'objet). L'émergence d'une asymétrie animé/inanimé dans le marquage de l'objet dans l'apprentissage de la langue artificielle pourrait être un effet de contamination de la langue maternelle des participants¹⁰. Il faut noter que ce problème ne remet pas en cause l'approche expérimentale,

9. Je me focalise sur la première étude expérimentale.

10. Les auteurs anticipent ce problème en notant que le marquage différentiel du cas en anglais n'est pas productif et se trouve limité aux pronoms (Fedzechkina, Jaeger & Newport 2012 : 17899). Cependant il paraît difficile de complètement exclure que cette asymétrie ait pu jouer un rôle dans les résultats.

mais plutôt amène à contrôler davantage les facteurs explicatifs. Le problème pourrait être réglé en incluant des participants d'une langue n'ayant pas une telle asymétrie.

2.3. Modélisation et variation

Les données expérimentales, comme on vient de le voir, sont un type de données particulièrement approprié pour tester les mécanismes causaux postulés par des théories. La modélisation, quant à elle, consiste à formuler les théories de façon très précise afin de faire des prédictions plus fines et de rendre compte de la variation inhérente aux données. Cette partie introduit la modélisation (partie 2.3.1) avant d'illustrer cette méthode avec une étude sur l'apprentissage de la syntaxe (partie 2.3.2).

2.3.1. Comment formuler une théorie : de la description verbale aux modèles statistiques

Les théories ou les généralisations linguistiques sont typiquement formulées de façon verbale. Illustrons avec un exemple : le schwa ou *e* muet en français, que l'on transcrit typiquement [ə] en alphabet phonétique international. En français, le mot *cerise* peut être prononcé avec ou sans [ə] dans la première syllabe : on peut ainsi dire *cerise* ou *c'rise*. Ce comportement n'est pas limité à ce mot, et on pourra dire plus généralement que « le schwa s'élide entre deux consonnes en français ».

Cette généralisation, répétée en (2a), peut également être formulée de façon formelle, en utilisant par exemple la règle phonologique en (2b). Cette règle a un sens très proche de (2a), avec trois éléments qui sont bien distingués : la cible du processus (à gauche de la flèche), le résultat (à droite de la flèche) et le contexte d'application du processus (après la barre oblique). Dans ce cas précis, la cible est ə, le résultat Ø (qui signifie l'absence de tout son) et le contexte est C_C (position interconsonantique).

- 2 a. « Le schwa s'élide entre deux consonnes. »
 b. $\text{ə} \rightarrow \emptyset / \text{C_C}$

Bien que les sens de ces deux expressions semblent identiques, la règle en (2b) a l'avantage d'avoir une interprétation non ambiguë, contrairement à (2a). Les phrases des langues naturelles, comme le français, donnent souvent lieu à des interprétations multiples. Cette ambiguïté pose un problème quand on veut évaluer une théorie formulée uniquement verbalement. Par exemple, dans la description en (2a), le présent peut être compris soit dans une lecture universelle (le schwa *s'élide nécessairement* entre deux consonnes) soit dans une lecture existentielle (le schwa *peut s'élider* entre deux consonnes). Quelle est la lecture qui est ici sous-entendue ? Ce problème ne se pose pas pour la règle formelle en (2b), qui par construction va être interprétée de façon universelle : cette règle décrit un processus d'élision catégorique, qui s'applique dès que la cible se trouve dans l'environnement adéquat (C_C).

Mais l'ambiguïté de (2a) ne se limite pas à ces deux lectures. Il y a une autre ambiguïté qui concerne le caractère exhaustif ou non des contextes d'application du processus d'élision : s'applique-t-il entre deux consonnes *et potentiellement dans d'autres contextes* (lecture littérale) ou s'applique-t-il *uniquement* entre deux consonnes (lecture plus forte) ? Ce genre d'ambiguïtés est lié au fait que les phrases des langues naturelles peuvent être enrichies d'inférences pragmatiques, qui sont liées à la façon dont on raisonne par rapport aux intentions communicatives de notre interlocuteur. Ces principes pragmatiques ne s'appliquent pas à l'interprétation des règles formelles, comme (2b). Ces règles n'admettent qu'une interprétation littérale : la règle

n'exclut pas que le processus puisse s'appliquer dans d'autres contextes (auquel cas une autre règle sera nécessaire).

Une théorie formalisée a l'avantage d'être moins ambiguë par rapport à une théorie formulée verbalement. Cependant les formalisations du type de la règle phonologique en (2b), qui sont dominantes dans le courant générativiste en linguistique, posent un autre problème : elles ne permettent pas la variation. En effet, comme expliqué plus haut, cette règle est interprétée comme s'appliquant catégoriquement. Appliquée au français, cette règle prédit ainsi que *cerise* devrait se prononcer *c'rise* systématiquement. Or on sait qu'il y a de la variabilité : ces modèles formels se heurtent donc au caractère continu ou probabiliste des données. Pour reprendre la question du schwa, on peut citer ces mots de Benoît de Cornulier :

À chaque instant, il existe entre l'élision obligatoire et l'impossible, une infinité mouvante de degrés qu'il est absurde de quantifier en quelques nombres entiers. Tel est le continu qui échappe, par essence, à la réduction à une combinatoire abstraite de phonèmes discrets et alignés. (de Cornulier 1975 : 105)

La variabilité dans la réalisation du schwa dont parle de Cornulier peut être illustrée avec des données de corpus. Utilisons comme exemple le corpus de français oral collecté à Nyon en Suisse¹¹. Dans ce corpus, le schwa tend à s'élider quand il est précédé et suivi d'une seule consonne (VC_CV), mais il y a des exceptions : huit mots ne font pas l'élision dans ce contexte sur un total de 307 occurrences de schwa potentiel attestées dans le corpus. À l'opposé, le schwa tend à rester stable quand il est précédé de deux consonnes et suivi d'une consonne (VCC_CV), mais il y a aussi des exceptions : trois mots élident le schwa dans ce contexte sur un total de 45 occurrences. L'analyse avec une règle phonologique comme (2b) ne rend pas compte de cette variabilité observée dans des données réelles.

Le problème ne serait pas si aigu si la variabilité était juste le résultat d'une moyenne à travers les participants. On pourrait par exemple imaginer que les huit mots qui ne font pas l'élision entre deux consonnes dans le corpus de Nyon ont été prononcés par la même personne, qui, à la différence des autres, n'aurait pas de règle d'élision du schwa dans sa grammaire. Mais en réalité cette variabilité n'est pas un artefact créé par l'agrégation des données individuelles : l'environnement entre consonnes (C_C) est traité de façon variable par un même locuteur. Pour le dire autrement, il y a de la variation *intraindividuelle* dans l'application de la règle¹².

Cette difficulté des grammaires formelles à rendre compte de la variation a été décrite d'une façon imagée et frappante par le linguiste Edward Sapir dans sa fameuse phrase *All grammars leak* (que l'on pourrait traduire par « Toutes les grammaires prennent l'eau »), dans l'extrait suivant :

Were a language ever completely « grammatical » it would be a perfect engine of conceptual expression. Unfortunately, or luckily, no language is tyrannically consistent. All grammars leak. (Sapir 1921 : 39)

« Si une langue se trouvait être complètement "grammaticale", elle serait un moteur parfait d'expression conceptuelle. Malheureusement, ou par chance, aucune langue n'est tyranniquement cohérente. Toutes les grammaires prennent l'eau. » (traduction personnelle)

11. Racine, Durand & Andreassen (2016).

12. Voir par ex. Bayes, Kaplan & Kaplan (2016) et Storme (2021) sur la variabilité intraindividuelle dans le traitement du schwa en français.

Les grammaires formelles sont-elles donc condamnées à l’incapacité à rendre compte de la variabilité inhérente aux langues ? Non, car il existe des variantes *probabilistes* de ces grammaires qui sont robustes au bruit.

Par exemple, dans le cas de la règle phonologique d’élision, on peut contraster une variante catégorique de cette règle, répétée en (3a), avec une variante probabiliste, en (3b). La seule différence tient à l’ajout d’un quatrième élément à la règle, en plus de la cible, du résultat et du contexte, à savoir la fréquence d’application de la règle. Ici, cette fréquence est égale 0,7, à titre purement indicatif. Cela signifie que cette règle ne s’applique pas dans tous les cas où la cible se trouve dans le contexte défini par la règle, mais dans 70 % des cas où ces conditions sont remplies.

- 3

a.

$\emptyset \rightarrow \emptyset/C_C$

variante catégorique

b.

$\emptyset \rightarrow \emptyset/C_C$

(fréquence d’application = 0,7) variante probabiliste

Les grammaires probabilistes, en permettant de façon inhérente de la variation, vont être robustes au bruit. Ce tournant probabiliste permet également d’utiliser les outils quantitatifs fournis par les méthodes d’inférence statistique pour comparer différentes théories et identifier celle qui rend le mieux compte des données. Par exemple, pour illustrer cela de façon très simple, une grammaire avec une règle d’élision s’appliquant avec une fréquence de 0,7 rendra mieux compte d’un corpus où le schwa est éliminé dans 65 % des cas entre deux consonnes qu’une grammaire avec une règle d’élision s’appliquant avec une fréquence de 0,3 (0,65 étant plus proche de 0,7 que de 0,3).

2.3.2. Étude de cas : modélisation de la syntaxe

On va maintenant illustrer l’intérêt de la modélisation avec un exemple issu de la syntaxe. La question de recherche qui va nous intéresser est une question fondamentale en syntaxe : y a-t-il des syntagmes dans la phrase, c’est-à-dire des regroupements de mots intermédiaires entre les unités syntaxiques minimales (les mots) et les unités maximales (les phrases), ou bien les phrases sont-elles directement formées à partir de mots sans unités complexes intermédiaires ? Cette étude de cas est intéressante car on peut présenter l’argument d’une façon intuitive mais également de façon plus rigoureuse, en utilisant la modélisation statistique au moyen de règles syntaxiques probabilistes, semblables aux règles phonologiques probabilistes de la partie 2.3.1.

On va donc comparer deux théories de la structure de la phrase. D’abord une théorie linéaire, qui postule que la syntaxe d’une langue est un ensemble de patrons de phrases, constitués chacun d’une séquence linéaire de catégories syntaxiques. Par exemple, la phrase déclarative en (4) serait générée par un patron de phrase de la forme $P \rightarrow \text{Déf } N \text{ } V \text{ } \text{Adj}$. Cette règle stipule qu’une phrase (P) est constituée d’une séquence contenant un déterminant (D), un nom (N), un verbe (V) et un adjectif (Adj). Dans cette théorie, la structure des phrases est plate, c’est-à-dire qu’il n’y a pas de niveau de représentation intermédiaire entre le niveau maximal (la phrase) et le niveau minimal (les mots).

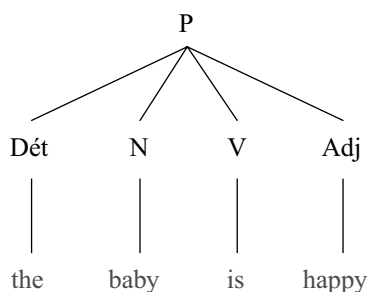
4 *The baby is happy.*

Cette théorie linéaire de la phrase peut être contrastée avec une théorie hiérarchique, qui pose que les phrases ne sont pas directement constituées de mots, mais de syntagmes. Les syntagmes sont eux-mêmes des groupements de mots ou d’autres syntagmes. Dans cette théorie, la phrase (4) n’est pas produite par un seul patron, comme dans la théorie linéaire, mais par trois patrons ou règles distinctes : une règle

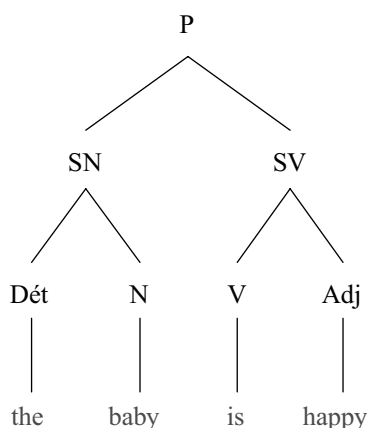
qui forme un syntagme nominal à partir d'un déterminant et d'un nom ($SN \rightarrow Det N$), une règle qui forme un syntagme verbal à partir d'un verbe et d'un adjectif ($SV \rightarrow V Adj$) et enfin une règle qui crée une phrase à partir d'un syntagme nominal et d'un syntagme verbal ($P \rightarrow SN SV$).

Les deux théories prédisent deux structures syntaxiques distinctes pour la phrase en (4) : la structure linéaire en (5a) pour la grammaire linéaire et la structure arborescente, hiérarchique en (5b) pour la grammaire hiérarchique avec des syntagmes.

5 a.



b.



Quelle est la bonne théorie ? *A priori*, la théorie hiérarchique paraît plus complexe. On a trois règles au lieu d'une. Mais nous allons voir que les apparences peuvent être trompeuses.

Considérons la formulation des questions en anglais. À la phrase déclarative en (4) correspond la phrase interrogative en (6). Quelle est la règle qui nous permet de former l'interrogative à partir de la déclarative ? Avec le verbe *être*, il y a une inversion du sujet en anglais. La formulation technique de cette inversion est très différente dans les deux théories. Dans la théorie linéaire, on peut dire que la question est formée en déplaçant le premier verbe en comptant à partir de la gauche au début de la phrase. La théorie hiérarchique offre non seulement un ordre linéaire mais aussi hiérarchique des éléments. C'est cet ordre hiérarchique que l'on va ici mobiliser : on peut dire en effet que la question en (6) est formée en déplaçant le verbe principal de la phrase au début. Le verbe principal de la phrase est défini de la façon suivante : il s'agit du verbe (V) qui se trouve directement sous le syntagme verbal (SV) constituant immédiat de la phrase (P).

6. *Is the baby happy?*

Laquelle est la bonne approche ? Encore une fois, la théorie hiérarchique paraît plus complexe. Mais si l'on considère des phrases plus complexes, on va voir qu'elle est en définitive plus simple. Considérons par exemple une phrase complexe, comme en (7a), où le nom est suivi d'une subordonnée relative (*who is tall*).

- 7 a. *The man who is tall is happy.*
 b. **Is the man who tall is happy?*
 c. *Is the man who is tall happy?*

La théorie linéaire prédit que la question correspondant à (7a) devrait être la phrase interrogative (7b), où le premier verbe a été placé au début de la phrase. Or cette question est agrammaticale, comme indiqué par l'étoile (*). La théorie hiérarchique prédit que c'est le verbe principal qui devrait se retrouver en première position dans l'interrogative : il s'agit du deuxième verbe *is* de la phrase (7a) selon l'ordre linéaire. La phrase interrogative résultante, en (7c), est grammaticale. On voit que la théorie hiérarchique a une notion d'ordre hiérarchique qui est pertinente dans la formation des questions. La théorie linéaire n'a pas ce concept, ce qui rend très difficile la formulation du processus de formation de la question.

On pourrait cependant amender l'hypothèse linéaire pour rendre compte de ce genre de phrases en complexifiant la règle de formation de la question. Par exemple, on pourrait spécifier que le premier verbe observé après un relatif ne compte pas dans le calcul. Ainsi, avec cet amendement, le *is* principal serait bien le premier verbe. Cependant la règle devrait encore être complexifiée davantage pour prendre en compte la possibilité de coordonner deux verbes sous un relatif, comme dans (8a). Là encore, c'est le verbe principal qui se retrouve en première position dans l'interrogative (8b). Le verbe suivant le premier verbe dans la relative ne peut pas monter en première position dans l'interrogative, comme montré en (8c). Dans ce cas-là, il faudrait dire que les deux premiers verbes ne comptent pas. On voit cependant que la règle devient beaucoup plus compliquée. Et l'apparent avantage de simplicité de la grammaire linéaire n'est plus aussi clair.

- 8 a. *The man who is tall and was here yesterday is happy.*
 b. *Is the man who is tall and was here yesterday happy?*
 c. **Was the man who is tall and here yesterday is happy?*

L'avantage de la simplicité que présentait la théorie linéaire disparaît à mesure que l'on considère des phrases plus complexes, et la théorie hiérarchique, qui semblait initialement inutilement alambiquée, se révèle finalement plus simple : elle peut rendre compte des différents types de questions avec une règle unique de déplacement du verbe principal (défini en termes hiérarchiques) vers le début de la phrase.

Nous venons de présenter l'argument en faveur de la grammaire hiérarchique de façon intuitive, avec quelques exemples. Il se trouve que cet argument a aussi été formalisé dans le cadre de la modélisation statistique dans une étude publiée en 2011 dans le journal *Cognition*¹³. Un argument entièrement explicite a l'intérêt d'aller au-delà de la simple intuition. Nous nous penchons maintenant sur cette étude.

Avant d'entrer dans le détail de l'étude, voyons comment les auteurs ont procédé globalement. À partir d'un corpus de phrases annotées avec des catégories syntaxiques (nom, adjectif, verbe, etc.), ils ont inféré la meilleure grammaire linéaire qui rend compte des données et la meilleure grammaire hiérarchique qui rend compte des données. Une

13. Perfors & Tenenbaum (2011).

grammaire linéaire correspond à un ensemble de patrons de phrases semblables à ceux vus précédemment ($P \rightarrow \text{Dét N V Adj}$). Une grammaire hiérarchique correspond à un ensemble de règles de formation de syntagmes. Puis les auteurs ont comparé ces grammaires avec l'inférence statistique bayésienne, qui est en gros une métrique qui permet de comparer des théories en fonction de deux critères : la simplicité de la théorie (ou, plus techniquement, la probabilité *a priori* d'une théorie) et l'adéquation de la théorie aux données (ou, plus techniquement, la probabilité des données d'après la théorie). La théorie qui satisfait le mieux ces deux critères est jugée plus probable. Deuxièmement les auteurs ont étudié la manière dont les deux grammaires se généralisaient au-delà du corpus. La grammaire qui permet de mieux rendre compte des types de phrases en dehors du corpus est jugée supérieure. C'est une approche très semblable à ce qui a été présenté précédemment, mais cette fois avec une argumentation entièrement explicite.

Les données utilisées dans l'étude sont issues de CHILDES, un corpus documentant la parole enfantine et souvent utilisé dans les études d'acquisition du langage. Les auteurs se sont concentrés spécifiquement sur la parole des adultes adressée aux enfants. La raison de ce choix est que les auteurs cherchent dans leur étude à modéliser l'apprentissage de la syntaxe par un enfant, avec l'idée qu'un enfant apprend la grammaire d'une langue un peu à la façon dont un scientifique aborde son objet d'étude, c'est-à-dire en comparant des théories alternatives pour rendre compte des données observées. Mais on peut aussi utiliser les résultats de cette étude pour comparer les théories de façon scientifique, sans envisager la question de l'acquisition.

Dans ce corpus, les auteurs extraient toutes les phrases (au nombre de 25 876) et ne retiennent que les types syntaxiques de phrase, c'est-à-dire les séquences de catégories syntaxiques. Par exemple, une phrase comme *The boy is nice* n'est pas retenue telle quelle dans l'analyse finale, mais sous la forme Dét N V Adj . Finalement, le corpus inclut 2338 types distincts. Les auteurs ne prennent en compte que les catégories syntaxiques des mots car ils s'intéressent uniquement à la modélisation de la syntaxe dans cette étude. Les 2338 types sont ensuite classés en six sous-corpus selon la fréquence des types. Par exemple, un premier sous-corpus ne comprend que les types de phrases qui sont attestés plus de 500 fois dans le corpus général. Un deuxième sous-corpus ne comprend que les types de phrases attestés plus de 300 fois, etc. Cette démarche va permettre de tester les grammaires linéaire et hiérarchique sur différents sous-corpus, et d'évaluer comment la grammaire peut être généralisée d'un sous-corpus plus petit à un sous-corpus strictement plus grand. L'approche est semblable à celle que l'on a adoptée plus haut, en regardant d'abord comment les théories linéaire et hiérarchique rendent compte d'une phrase simple, comme (4), puis de phrases plus complexes, comme (7a) et (8a).

Après l'établissement des corpus, vient la phase d'induction grammaticale. Commençons par décrire la méthode utilisée pour la théorie hiérarchique. Ici, les auteurs utilisent un analyseur syntaxique automatique qui, à partir des types de phrases observés, va inférer la grammaire hiérarchique qui rend les données observées les plus probables. Cet analyseur syntaxique va poser des syntagmes intermédiaires pour rendre compte des séquences de symboles linéaires. Par exemple, il va poser une règle qui dit qu'une phrase peut être constituée d'un syntagme nominal suivi d'un syntagme verbal ($P \rightarrow \text{SN SV}$). Et il va donner une probabilité à cette règle car toutes les phrases du corpus ne peuvent pas être analysées comme un SN suivi d'un SV. En effet, certaines phrases du corpus seront mieux analysées comme un SV seul ($P \rightarrow \text{SV}$), par exemple les phrases à l'impératif (comme *Come here !*). Puis il posera également une règle de formation d'un SN, par exemple «un SN contient un déterminant suivi d'un nom»

(SN \rightarrow Dét N). Là aussi cette règle viendra avec une probabilité particulière, car tous les SN n'ont pas cette forme. On voit ici l'importance d'avoir une modélisation qui est spécifiquement *statistique* : cela permet de rendre compte du fait qu'il existe différentes façons de former des phrases.

De la même façon les auteurs infèrent la grammaire linéaire qui rend le mieux compte des données. Dans ce cas-là, la démarche est beaucoup plus simple : il suffit d'enregistrer chaque séquence de catégories syntaxiques avec sa fréquence d'apparition dans le corpus et de créer la règle correspondante. Par exemple, si le type de phrase P \rightarrow Dét N V Adj représente 20 % des types de phrases du corpus, alors on aura une règle de la forme suivante : P \rightarrow Dét N V Adj (fréquence d'application = 0,2).

Maintenant que l'on a obtenu la meilleure grammaire hiérarchique probabiliste et la meilleure grammaire linéaire probabiliste pour rendre compte des données, on va pouvoir comparer ces deux grammaires entre elles pour déterminer laquelle est la meilleure, d'après les deux critères de l'inférence bayésienne vus plus haut. Le premier critère valorise les grammaires qui sont plus simples. On va compter le nombre de symboles dans chacune des deux grammaires et établir à partir de ce calcul une probabilité *a priori* de chacune des grammaires, en pénalisant les grammaires en proportion de leur nombre de symboles.

Le deuxième critère va favoriser la grammaire qui rend les données les plus probables, ou pour le dire autrement qui maximise la probabilité des données observées dans le corpus. Cette probabilité peut être calculée à partir des probabilités d'application des différentes règles grammaticales. Par exemple, supposons que la grammaire hiérarchique ait une règle de formation des phrases P \rightarrow SN SV avec une probabilité d'application de 0,6 et une règle de formation des SV \rightarrow V SN avec une probabilité d'application de 0,3. D'après cette grammaire, la probabilité d'observer le type de phrase SN V SN dans un corpus est obtenue en multipliant la probabilité d'appliquer les deux règles, cette probabilité sera donc égale à $0,6 \times 0,3 = 0,18$. Le score global d'une grammaire est obtenu en combinant ces deux critères (probabilité *a priori* d'une grammaire et probabilité des données d'après cette grammaire) au moyen d'un produit¹⁴, qui nous donne la probabilité *a posteriori* de la théorie¹⁵.

Le résultat obtenu par les auteurs sur la comparaison entre grammaires hiérarchique et linéaire est très intéressant. On observe que la grammaire linéaire a un meilleur score global sur le corpus de types de phrases le plus restreint, qui correspond aux 8 types de phrases les plus fréquents, car elle est à la fois plus simple et rend les données plus probables. En réalité, la grammaire linéaire rend toujours les données plus probables parce qu'elle enregistre exactement la fréquence des différents types. Mais à mesure que le corpus grandit et que des types de phrases plus complexes et plus rares sont pris en compte, la comparaison bascule en faveur du modèle hiérarchique : la complexité de la grammaire hiérarchique croît moins vite que celle de la grammaire linéaire à mesure que le nombre de types de phrases distincts augmente. Cela est lié au fait que la grammaire hiérarchique est capable de réutiliser des patrons déjà inférés pour analyser de nouveaux types alors que la grammaire linéaire doit toujours créer de nouvelles règles pour des types de phrases non préalablement observés.

14. Dans l'article, les auteurs utilisent en fait une somme, mais c'est parce qu'ils appliquent une transformation logarithmique sur les probabilités. Or une multiplication devient une somme après transformation logarithmique (cf. $\log(a*b) = \log(a) + \log(b)$).

15. *A posteriori* car il s'agit de la probabilité de la théorie *après* avoir vu les données.

Pour illustrer cela plus concrètement, si la grammaire hiérarchique a une règle $SN \rightarrow \text{Det SAdj N}$ et une règle $SAdj \rightarrow \text{Adv Adj}$ qu'elle a induites à partir d'un sous-corpus, alors elle va prédire automatiquement la possibilité d'une séquence Det Adv Adj N (comme *the very nice boy sleeps*) sans créer de nouveaux patrons. En effet ce type de phrase est prédit directement à partir des deux règles déjà existantes : il suffit de les combiner entre elles. En revanche, pour rendre compte des mêmes données, la grammaire linéaire sera obligée d'ajouter un nouveau patron de phrase ($P \rightarrow \text{Det Adv Adj N V}$).

Le deuxième test utilisé par les auteurs est la capacité d'une grammaire à être généralisable au-delà du corpus. Ce test montre que la grammaire hiérarchique a une bien meilleure capacité à rendre compte des données non encore observées que la grammaire linéaire, pour les raisons vues au-dessus. À titre d'exemple, la grammaire hiérarchique qui a été entraînée sur le corpus du niveau 5 parvient à analyser 87 % des types de phrases du niveau plus inclusif 6. Par contraste, la grammaire linéaire obtenue à partir du corpus de niveau 5 n'est capable d'analyser que 11 % des types de phrases du corpus plus inclusif de niveau 6. C'est parce que 89 % des types sont nouveaux, et la grammaire linéaire est incapable de les analyser sans créer de nouvelles règles.

Pour conclure, on a vu que la modélisation statistique donne la possibilité de formuler les théories linguistiques de façon non seulement explicite mais aussi probabiliste, permettant de rendre compte de la variabilité observée dans les langues, par exemple la variabilité dans les types syntaxiques utilisés. Une fois que les théories linguistiques sont modélisées de façon statistique, on peut ensuite utiliser les outils offerts par les statistiques (comme l'inférence bayésienne) pour comparer les prédictions de ces théories d'un point de vue quantitatif. L'argumentation devient alors entièrement explicite et permet de dépasser le niveau purement intuitif qui caractérise souvent les débats sur les théories linguistiques dans les approches non formelles.

3. DÉBATS AUTOUR DU LANGAGE INCLUSIF

Maintenant que nous avons vu quel était l'intérêt des méthodes d'expérimentation et de modélisation en linguistique d'un point de vue général, nous allons passer à des applications plus concrètes, en commençant par la méthode expérimentale. Comme terrain d'application, un sujet particulièrement intéressant nous est offert par le langage inclusif, ou sa variante dans le domaine de l'écrit, l'écriture inclusive.

Qu'est-ce que le langage inclusif ? Il s'agit d'une façon générale d'un ensemble de stratégies visant à réduire l'utilisation du masculin générique comme forme neutre d'un point de vue du genre. D'un point de vue théorique, le langage inclusif pose la question suivante : le genre grammatical d'un mot influe-t-il sur les inférences portant sur le genre de son référent ? Par exemple, la forme masculine du nom *chirurgien* dans la phrase « L'hôpital recherche un chirurgien » est-elle interprétée de façon complètement neutre ? Ou bien a-t-on tendance à inférer, en entendant cette phrase, que l'hôpital recherche plutôt un homme pour ce poste en chirurgie ?

Cette question est à la racine de tous les débats autour du langage inclusif mais, comme on va le voir, il n'est pas aisé d'y répondre. En effet il y a d'autres facteurs à côté du genre grammatical d'un mot qui peuvent influencer sur les inférences de genre que l'on fait, comme les stéréotypes de genre associés à certains noms. Pour reprendre l'exemple de la chirurgie, il se pourrait que ce soient majoritairement des hommes qui exercent cette profession. Ainsi le biais masculin dans l'interprétation pourrait être lié à ce stéréotype plutôt qu'au genre grammatical masculin, ou alors les deux effets

pourraient s'additionner. On a vu dans la partie précédente que la méthode expérimentale est particulièrement indiquée pour démêler les liens de causalité entre des variables. Ainsi l'expérimentation devrait pouvoir jouer un rôle important pour arbitrer entre ces théories alternatives.

Nous allons d'abord replacer la question du langage inclusif dans le contexte plus global des systèmes de genre dans les langues (partie 3.1) et examiner les arguments qui ont été mis en avant par les tenants et les opposants au langage inclusif (partie 3.2). On verra que les écrits sur le sujet de l'écriture inclusive font des affirmations souvent trop vagues pour être directement testables, ce qui renforce la nécessité d'une formalisation.

3.1. Le langage inclusif dans la typologie des systèmes de genre

Le langage inclusif, et sa manifestation écrite en français, l'écriture inclusive, peuvent se définir comme un ensemble de stratégies visant à éliminer l'utilisation de la forme morphologiquement masculine comme forme sémantiquement neutre en français. Parmi ces stratégies, on trouve notamment (i) la féminisation des noms de fonctions, grades, métiers et titres, comme en (9a, b), (ii) l'utilisation conjointe du féminin et du masculin pour évoquer des groupes mixtes ou des personnes dont le genre n'est pas connu ou pertinent, comme en (9c, d), (iii) l'utilisation de mots épiciens, qui ont la même forme au féminin et au masculin, en (9e), ou encore (iv) l'accord de proximité, en (9f)¹⁶.

- | | |
|--|--|
| 9 a. <i>Madame la ministre</i> | (au lieu de <i>Madame le ministre</i>) |
| b. <i>Cette femme est une autrice remarquable</i> | (au lieu de <i>un auteur remarquable</i>) |
| c. <i>les candidats et candidates</i> | (au lieu de <i>les candidats</i>) |
| d. <i>les candidat-e-s</i> | (au lieu de <i>les candidats</i>) |
| e. <i>les membres du corps professoral</i> | (au lieu de <i>les professeurs</i>) |
| f. <i>Les candidats et candidates sont parties</i> | (au lieu de <i>partis</i>) |

Le langage inclusif répond donc à un problème assez spécifique, qui est lié aux caractéristiques du système de genre utilisé en français. Il est utile de prendre un peu de recul par rapport au français pour envisager sa position dans la typologie des systèmes de genre.

On peut d'abord définir le genre grammatical comme une propriété d'un mot qui se manifeste dans des phénomènes d'accord grammatical¹⁷. Par exemple, le mot «personne» est féminin et cela se manifeste dans le fait que ce nom requiert des modificateurs à la forme féminine, par exemple «*la* personne», «*une* personne», «personne *âgée*». Le genre d'un nom est souvent corrélé à des propriétés sémantiques de son référent (par exemple, les noms féminins qui dénotent des animés en français renvoient souvent à des référents de genre féminin) ou à des propriétés phonologiques de sa forme (par exemple les mots latins en *-a* sont souvent féminins). Mais il y a aussi une part d'arbitraire. Par exemple, les mots *sable* et *table* en français renvoient tous deux à des inanimés (une matière naturelle et un artefact, respectivement) et ont une forme phonologique assez proche. Malgré cette similarité, ils ont des genres différents, masculin et féminin respectivement. On parle de genre sémantique quand le genre est entièrement déterminé par des propriétés sémantiques. Par exemple, pour les pronoms

16. Haut conseil à l'égalité entre les femmes et les hommes (2022).

17. Corbett (2013a).

en anglais, on parle de genre sémantique parce que l'on peut largement déduire de la forme du pronom le genre de son référent (*he* renvoie à un homme, *she* à une femme, *it* à un objet), à quelques exceptions près.

Le genre grammatical n'est pas un phénomène majoritaire à travers les langues. Par exemple, dans l'échantillon de 257 langues utilisé dans le chapitre sur le genre dans le *World Atlas of Language Structures*, seulement 112 ont un genre grammatical. Les autres langues, comme le turc, n'ont aucun système de genre grammatical. Parmi les 112 langues qui ont un genre grammatical, 84 ont un genre grammatical basé sur le genre social, comme en français. Les autres langues ont un genre grammatical basé sur d'autres propriétés, comme l'opposition animé/inanimé. C'est le cas par exemple du peul de Maasina, une langue parlée au Mali, qui a un système de genre riche mais où la distinction masculin/féminin n'est pas pertinente¹⁸. Environ un tiers des langues de l'échantillon ont donc un genre grammatical fondé sur le genre social, et présentent donc potentiellement le problème du français par rapport au langage inclusif.

Mais il ne suffit pas d'avoir un genre grammatical basé sur le genre social pour que le problème du langage inclusif se pose. En fait, le langage inclusif répond à une situation très spécifique. Dans une langue avec des distinctions de genre, ces distinctions peuvent être limitées à certains contextes. En anglais, les distinctions de genre sont largement limitées aux pronoms (*he*, *she*, *it*). Parmi les pronoms, les distinctions de genre sont limitées au singulier (*he*, *she*, *it* vs *they*). On parle de *synchrétisme* quand un trait grammatical n'est pas exprimé morphologiquement dans tous les contextes grammaticaux d'un mot. Quelques contextes où le genre est synchrétique sont les pluriels en anglais (*they*), les noms de métiers ou les titres en français, qui sont traditionnellement utilisés au masculin (*Madame le ministre*) et les contextes génériques (par exemple «le candidat doit dire s'il a un casier judiciaire vierge»). En cas de synchrétisme, le français utilise la forme masculine. Mais on peut remarquer que ce n'est pas forcément le cas à travers les langues. À ma connaissance, il n'existe pas d'études typologiques approfondies sur les stratégies de marquage du genre en cas de synchrétisme. Le masculin semble très utilisé dans les langues européennes dont nous avons l'habitude. Mais d'autres stratégies sont utilisées dans d'autres langues. Par exemple, en buin, une langue parlée sur l'île Bougainville en Papouasie–Nouvelle-Guinée, c'est la forme féminine qui est utilisée au pluriel pour renvoyer à un groupe mixte (là où le français utilise le masculin)¹⁹. En vanimo, une autre langue de Papouasie–Nouvelle-Guinée, une forme distincte du masculin et du féminin est utilisée pour les groupes mixtes²⁰.

Pour récapituler, on peut résumer les grands traits de la typologie des systèmes de genre discutée précédemment à l'aide de quelques paramètres : présence/absence d'un système de genre grammatical (français vs turc), genre grammatical fondé/non fondé sur le genre social (français vs peul de Maasina), genre grammatical entièrement/partiellement fondé sur le genre social (anglais vs français), et forme utilisée en cas de synchrétisme de genre grammatical (le masculin en français, le féminin en buin et une forme ni masculine ni féminine en vanimo). Le langage inclusif répond donc à une situation très spécifique, observée en français notamment : un système de genre basé sur le genre social, où la forme masculine est utilisée en cas de synchrétisme.

18. Corbett (2013b).

19. Laycock (2003 : xv).

20. Plank & Schellinger (1997 : 76).

3.2. Arguments pour et contre le langage inclusif dans l'espace public

Nous allons maintenant passer en revue quelques arguments qui ont été fournis en faveur de et contre l'écriture inclusive dans l'espace public, avant de les examiner dans la partie suivante avec des données expérimentales.

L'argument typique en faveur de l'écriture inclusive est l'idée que le masculin est surreprésenté dans la langue, et que cela conduit à une invisibilisation des femmes²¹. Cet argument pose cependant un problème, que ne manquent pas de pointer les détracteurs de l'écriture inclusive²² : il présuppose que la forme masculine favorise une lecture masculine de son référent. Or il s'agit d'une question empirique à laquelle il n'y a pas de réponse claire *a priori*. En effet, il existe en français des emplois clairement neutres du masculin, par exemple en (10a), avec l'utilisation du *il* impersonnel, en (10b), où le masculin réfère à un contenu propositionnel (le fait qu'Anne est malade) et en (10c), où la forme masculine peut être utilisée pour renvoyer au groupe constitué d'Anne et Pierre, sans que l'on fasse l'inférence qu'Anne est un homme.

10 a. *Il pleut.*

b. *Anne est malade, je le sais.*

c. *Anne et Pierre, ils sont partis tout à l'heure.*

d. *Le garçon, il est parti.*

Les données linguistiques du français suggèrent une déconnexion entre genre grammatical et genre social pour les inanimés et une connexion partielle pour les animés ou humains, avec une forme morphologiquement masculine qui a un sens générique (pouvant renvoyer à un référent homme ou femme, ou mixte dans le cas d'un groupe) et un sens spécifiquement masculin (renvoyant uniquement à un référent homme). Le sens générique est illustré en (10c) tandis que le sens spécifiquement masculin est illustré en (10d).

L'hypothèse des défenseurs de l'écriture inclusive est que, même dans ces emplois supposément génériques, comme en (10c), la forme masculine crée un biais interprétatif en faveur du masculin. Ainsi, en entendant une forme masculine dans un contexte générique, un locuteur francophone serait plus susceptible de penser que l'on parle d'un homme que d'une femme. Cependant, certains linguistes, comme Claude Hagège, suggèrent que ce biais masculin éventuel ne serait pas lié à la langue elle-même, mais plutôt aux stéréotypes sociaux : « ce n'est pas la langue qui est sexiste mais les comportements sociaux », affirme-t-il ainsi dans une tribune du *Monde*²³. On peut illustrer le problème concrètement à partir du texte suivant :

On est samedi après-midi. Un père conduit son fils à un anniversaire. Alors qu'il consulte son téléphone, il manque de percuter un animal et emboutit la voiture contre un arbre. Le père meurt sur le coup, le fils est gravement blessé et amené aux urgences. Là, le médecin censé l'opérer réagit : « Je ne peux pas opérer cet enfant, c'est mon fils. »

Ce texte pose typiquement un problème d'interprétation pour les lecteurs, puisqu'il semble y avoir une contradiction entre l'information selon laquelle un garçon a perdu son père dans un accident de voiture et l'information selon laquelle il doit être ensuite opéré par son père. Le problème disparaît pourtant aussitôt si on accepte que l'expression masculine « le médecin » est ici à interpréter de façon générique. Le fait que le

21. Par exemple Haddad (2017).

22. Par exemple Hagège (2017).

23. Hagège (2017).

texte pose un problème d'interprétation signifie que les lecteurs s'orientent automatiquement sur une lecture spécifique et masculine du mot. Cependant la source de cette inférence masculine n'est pas évidente d'emblée : est-elle liée au fait que la forme du mot est masculine, comme le dirait un défenseur de l'écriture inclusive, ou bien à un stéréotype masculin associé à la chirurgie, comme le dirait un détracteur de l'écriture inclusive comme Hagège ? Autrement dit, la source de l'inférence masculine est-elle linguistique ou sociale ?

Outre ce débat sur le rôle du genre grammatical dans les inférences de genre social en français, il y a un autre débat qui anime défenseurs et détracteurs de l'écriture inclusive : la question de sa complexité. Chez les détracteurs de l'écriture inclusive, on met l'accent sur l'idée que le langage inclusif est complexe, et que cela peut avoir des conséquences néfastes sur l'apprentissage de la langue²⁴.

Mais là aussi, la réponse à cette question n'est pas aussi simple qu'il pourrait y paraître. En effet, l'écriture inclusive correspond à une diversité de stratégies, comme on l'a vu en (9), et ces stratégies ne sont pas toutes équivalentes en termes de complexité formelle. Le point médian est sans doute plus complexe que l'utilisation de formes épicles par exemple.

De plus, une plus grande complexité formelle n'implique pas toujours une plus grande complexité cognitive. Des études montrent que des locuteurs qui ont un lexique plus large, et qui connaissent donc plus de mots, ont un traitement cognitif des mots du lexique plus rapide que les locuteurs avec un plus petit lexique mental²⁵. Par exemple, ils sont capables de reconnaître les mots plus rapidement. De façon assez contre-intuitive, une plus grande complexité du lexique se traduit donc par un traitement cognitif plus rapide des mots. Si l'on extrapole ces résultats, la complexité lexicale résultant de l'adoption de stratégies d'écriture inclusive comme la féminisation des noms de métiers pourrait en fait se traduire par une accélération du traitement lexical.

Un autre résultat de cette étude sur la vitesse de traitement lexical est intéressant dans le contexte de l'écriture inclusive. L'article montre en effet que la fréquence d'utilisation d'un mot a un effet sur la vitesse de traitement : plus un mot est utilisé fréquemment dans la langue, plus il est traité rapidement par les locuteurs. Si on applique ce résultat à l'écriture inclusive, on s'attend donc à ce que les formes inclusives novatrices, comme *autrice* à la place du masculin générique *auteur*, soient d'abord relativement difficiles à traiter en raison de leur faible fréquence. Mais à mesure qu'elles se répandent dans la population et deviennent plus fréquentes, leur traitement cognitif devrait se trouver facilité.

Enfin, certaines stratégies d'écriture inclusive pourraient s'avérer moins coûteuses d'un point de vue cognitif que le masculin générique. C'est le cas de l'accord de proximité, illustré en (11a), qui est une stratégie d'écriture inclusive, par rapport à l'accord de résolution, illustré en (11b), qui utilise le masculin générique.

- | | |
|---|------------------------|
| 11 a. <i>les départements et régions françaises</i> | (accord de proximité) |
| b. <i>les départements et régions français</i> | (accord de résolution) |

En quoi l'accord de proximité pourrait-il présenter un avantage cognitif ? Parce qu'il implique une dépendance syntaxique plus courte que l'accord de résolution.

24. Par exemple, l'Académie française (2017) écrit, dans un avis critique sur l'écriture inclusive : « La multiplication des marques orthographiques et syntaxiques que [l'écriture inclusive] induit aboutit à une langue désunie, disparate dans son expression, créant une confusion qui confine à l'illisibilité. »

25. Brysbaert, Mandera & Keuleers (2018).

Par exemple, dans (11a), l'accord de l'adjectif se fait avec le nom linéairement le plus proche, donc au féminin. Dans (11b), il faut remonter plus loin pour décider que la forme doit être le masculin. Or la localité des dépendances syntaxiques joue un rôle dans la difficulté de traitement cognitif²⁶ : une dépendance non locale, comme l'accord de résolution en (11b), a un coût plus important en termes de mémoire qu'une dépendance locale comme l'accord de proximité en (11a).

Pour conclure, on a vu que l'écriture inclusive répond à une configuration très particulière du système de genre grammatical en français, à savoir l'utilisation de formes masculines en contexte de syncrétisme de genre. Les débats autour de l'écriture inclusive s'appuient sur des hypothèses linguistiques qui ne sont pas évidentes *a priori* et qui doivent donc être évaluées empiriquement. L'expérimentation fournit des méthodes particulièrement adaptées pour tester ces hypothèses, comme on va le voir dans la partie suivante.

4. LANGAGE INCLUSIF : APPORT DE L'EXPÉRIMENTATION

Les débats autour du langage inclusif se cristallisent donc autour de deux hypothèses empiriques. On trouve d'abord du côté de la défense du langage inclusif l'hypothèse que le genre grammatical d'un mot influe sur les inférences de genre (H1). Du côté de la critique du langage inclusif, on met en avant l'hypothèse que l'écriture inclusive crée des problèmes à la lecture et pour l'apprentissage (H2). Comme on l'a vu dans la partie précédente, ces hypothèses nécessitent d'être testées avec attention parce que, dans un cas, il y a d'autres facteurs comme les stéréotypes qui pourraient expliquer les inférences de genre en plus ou même à la place du genre grammatical, et dans l'autre, parce que le lien entre complexité formelle et complexité cognitive n'est pas évident.

Pour tester les hypothèses (H1) et (H2), l'approche expérimentale est particulièrement appropriée parce que, comme on l'a vu dans la partie 2, les données expérimentales permettent de répondre à des questions portant sur la causalité. Dans la partie 4.1, on se penche d'abord sur une étude observationnelle, non expérimentale, qui teste une hypothèse beaucoup plus ambitieuse, à savoir l'hypothèse d'un lien entre genre grammatical et inégalité femmes-hommes à l'échelle de pays entiers. On montre qu'elle pose exactement le genre de problèmes qu'une approche expérimentale cherche à éviter. Dans la partie 4.2, on passe en revue plusieurs études expérimentales qui ont testé l'hypothèse (H1) en contrôlant l'effet des stéréotypes ou de la compétition morphologique. Dans la partie 4.2, on présente une étude qui a testé l'hypothèse (H2). Le but de cette partie n'est pas de donner une liste exhaustive des études portant sur le sujet, mais plutôt de se concentrer sur quelques études emblématiques, en insistant sur les apports de la méthode expérimentale²⁷.

4.1. Effet du genre grammatical sur l'inégalité femmes-hommes : une étude observationnelle

L'idée centrale des défenseurs de l'écriture inclusive est qu'il y a un lien entre la façon de marquer le genre dans une langue et l'égalité des sexes dans la société où cette langue est parlée. Cette partie se concentre sur une étude observationnelle qui

26. Gibson (1998).

27. Pour une approche plus exhaustive, on peut se référer au livre de Gygax, Zufferey & Ute (2021) sur le sujet.

a cherché à tester ce lien entre marquage du genre dans une langue et indice d'égalité femmes-hommes dans le pays où la langue est parlée²⁸.

Dans cette étude, les langues sont classées en trois groupes : les langues qui n'ont pas de genre grammatical (comme le turc), les langues qui ont un genre *grammatical* basé sur l'opposition masculin-féminin (comme le français) et les langues qui ont un genre *sémantique* basé sur l'opposition masculin-féminin (comme l'anglais). Les pays sont caractérisés d'après la langue majoritaire dans le pays et l'analyse inclut ainsi 134 pays-langues. Le but de l'étude est de voir si l'appartenance d'une langue à un de ces trois groupes a une influence sur l'égalité femmes-hommes dans les pays où ces langues sont parlées. Notons déjà que la raison pour laquelle les auteurs ont choisi de se focaliser sur cette variable n'est pas tout à fait claire dans la mesure où, comme on l'a vu dans la partie 3.1, ce n'est pas tellement le système de genre grammatical en soi qui peut poser problème, mais plutôt l'utilisation d'une forme masculine en contexte syncrétique. Mais nous reviendrons sur ce point plus loin.

Pour mesurer l'égalité femmes-hommes dans un pays, les auteurs utilisent un indice établi par le Forum économique mondial : le *Global Gender Gap index*. Cet indice évalue le fossé entre femmes et hommes dans les sphères de l'économie, l'éducation, la politique et la santé. Chaque pays a un score compris entre 0 et 1 (0 et 1 représentant respectivement le plus petit et le plus grand score d'égalité femmes-hommes). Dans l'analyse, les auteurs introduisent d'autres variables de contrôle en plus du système de genre grammatical : les traditions religieuses, les types de gouvernements, la position géographique et l'indice de développement humain. Le but est d'établir si le type de genre grammatical contribue à l'inégalité femmes-hommes indépendamment de ces facteurs.

Les résultats montrent qu'il existe une corrélation entre l'indice d'égalité de genre et le marquage du genre linguistique : les pays avec des langues non genrées (comme le turc) ou sémantiquement genrées (comme l'anglais) ont en moyenne un indice d'égalité femmes-hommes supérieur aux pays avec des langues grammaticalement genrées (comme le français). Cet indice est égal en moyenne à 0,72 pour les langues à genre sémantique, à 0,70 pour les langues sans genre linguistique et à 0,67 pour les langues avec un genre grammatical.

Ces résultats semblent apporter une confirmation à l'hypothèse selon laquelle il peut y avoir un lien entre un système linguistique de genre et la structure sociale d'un pays. Cependant ils sont assez difficiles à interpréter. En effet, ce qui est critiqué dans les langues avec un genre grammatical, comme on l'a rappelé plus haut, c'est spécifiquement l'utilisation de la forme masculine comme forme par défaut, que le système de genre en question soit grammatical (comme en français) ou sémantique (comme en anglais). Par ailleurs, on peut tout à fait imaginer un système de genre grammatical où la forme par défaut ne serait pas le masculin, mais le féminin ou encore une forme autre (voir la partie 3.1 pour des exemples de langues utilisant d'autres formes en contexte syncrétique). Il faudrait donc montrer que les langues avec genre grammatical utilisent généralement des masculins par défaut. À ma connaissance, il n'y a pas à l'heure actuelle d'étude typologique qui établisse cette corrélation.

Cette étude pose également un autre problème, qui nous ramène au cœur des motivations de la démarche expérimentale par rapport à une démarche observationnelle (voir partie 2.2.2). Parmi les variables de contrôle utilisées par les auteurs, on trouve la variable Religion, qui comprend les valeurs suivantes : christianisme, islam, judaïsme,

28. Prewitt-Freilino, Caswell & Laakso (2012).

bouddhisme, hindouisme, irréligion et autre. Cette variable ne distingue donc pas catholicisme et protestantisme. Ce choix est potentiellement problématique, dans la mesure où il pourrait y avoir une corrélation entre catholicisme/protestantisme et genre lexical/sémantique. En Europe, les langues romanes (à genre grammatical) sont parlées dans des pays plutôt catholiques alors que les langues germaniques (qui utilisent plus fréquemment un genre sémantique) sont plutôt parlées dans des pays protestants. La variation de l'indice Global Gender Gap entre pays avec genre grammatical et pays avec genre sémantique pourrait potentiellement s'expliquer par la différence entre catholicisme et protestantisme. On se retrouve donc dans la situation évoquée dans la partie 2.2.2 où une variable alternative pourrait expliquer les résultats, en l'occurrence une variable de type culturel par rapport à une variable proprement linguistique.

Pour conclure, l'étude apporte un résultat sur une corrélation entre système linguistique de genre et inégalité sociale des sexes. Cependant cette corrélation est difficile à interpréter. D'abord, l'étude ne se concentre pas sur la variable linguistique la plus pertinente, qui serait l'utilisation de masculins génériques plutôt que l'absence/présence de genre. Par ailleurs l'effet du système grammatical est potentiellement confondu avec celui de la religion. D'un point de vue plus général, cette étude illustre le type de problèmes que posent des études corrélationnelles à grande échelle qui cherchent à évaluer les effets de la langue sur la pensée ou les structures sociales et politiques. Il est souvent difficile de distinguer les effets proprement linguistiques des effets culturels : les langues ayant une parenté génétique correspondent souvent aussi à des sociétés culturellement proches²⁹. Étant donné ces difficultés, nous nous tournons à présent vers des études expérimentales, qui permettent de mieux contrôler l'effet de ces différents facteurs.

4.2. Effet du genre grammatical sur les inférences de genre : études expérimentales

De nombreuses études expérimentales ont été menées pour évaluer l'effet du genre grammatical sur les inférences de genre social. Cette partie se concentre sur des études qui ont évalué cet effet en contrôlant le rôle des stéréotypes et le rôle de la compétition linguistique avec des expressions alternatives.

4.2.1. Études qui contrôlent l'effet des stéréotypes

Commençons par les études ayant contrôlé le rôle des stéréotypes. Comme on l'a vu dans la partie 3.2, il est important de contrôler soigneusement l'effet des stéréotypes de genre quand on évalue le rôle du genre grammatical. En effet, certains linguistes comme Hagège ont proposé que seuls les stéréotypes ont un effet sur l'interprétation. On peut citer une étude pionnière sur le sujet datant de 2008³⁰ sur la façon dont les locuteurs francophones, germanophones et anglophones interprètent les référents pour des noms de métier au pluriel, comme *les travailleurs sociaux* et l'équivalent anglais *the social workers*. Le français et l'allemand diffèrent de l'anglais dans leur système linguistique de genre : ce sont des langues à genre grammatical, qui utilisent les masculins de façon générique pour renvoyer notamment à des groupes mixtes (cf. la règle «c'est le masculin qui l'emporte»). Par exemple, l'expression *les travailleurs*

29. Voir Roberts, Winters & Chen (2015) pour un autre exemple sur le lien entre marquage du futur et comportements économiques.

30. Gygax, Gabriel, Sarasin, Oakhill & Garnham (2008).

sociaux peut être utilisée en principe de façon générique pour renvoyer à un groupe pouvant comprendre des hommes et des femmes. L'anglais est une langue à genre sémantique, qui n'associe pas de genre grammatical à ses noms et utilise des formes épiciques ou neutres pour renvoyer à des groupes mixtes. Par exemple, l'expression *the social workers* n'a pas de genre spécifique d'un point de vue linguistique.

L'originalité de cette étude tient au fait qu'elle teste les effets de ces systèmes linguistiques de genre sur les inférences de genre social tout en contrôlant l'effet des stéréotypes associés aux noms. Les auteurs utilisent une première étude³¹ qui a établi les stéréotypes de genre associés à différents métiers, indépendamment de la question du genre grammatical. Ces stéréotypes sont établis en demandant à des participants d'évaluer la proportion de femmes et d'hommes dans les différents métiers. Une fois que ces stéréotypes sont établis, les auteurs sélectionnent trois types de noms qu'ils vont utiliser dans l'étude sur l'effet du genre grammatical : des noms correspondant à des métiers stéréotypiquement féminins (*travailleurs sociaux*), des métiers neutres du point de vue du genre (*chanteurs*) et des métiers stéréotypiquement masculins (*techniciens*).

Dans l'étude portant sur l'effet du genre grammatical, on présente aux participants des phrases qui contiennent une forme en principe générique renvoyant à un nom de métier au pluriel (par exemple *les travailleurs sociaux* au masculin générique, ou *the social workers*). Puis on leur demande s'ils acceptent une continuation de cette phrase qui implique que le groupe désigné par le nom au pluriel inclut soit des hommes soit des femmes. Par exemple, pour tester si une continuation masculine est acceptée, on va demander aux participants s'il leur paraît possible de continuer la phrase en (12a) avec la phrase (12b). Pour tester si un nom masculin générique est compatible avec une lecture féminine, on va leur demander s'il leur paraît possible de continuer la phrase en (12a) avec la phrase (12c). La même chose est demandée aux germanophones et francophones, avec les exemples donnés en (13) et (14).

- 12 a. *The social workers were working through the station.*
 b. *Since sunny weather was forecast, several of the men weren't wearing a coat.*
 c. *Since sunny weather was forecast, several of the women weren't wearing a coat.*
- 13 a. *Les assistants sociaux marchaient dans la gare.*
 b. *Du beau temps étant prévu, plusieurs hommes n'avaient pas de veste.*
 c. *Du beau temps étant prévu, plusieurs femmes n'avaient pas de veste.*
- 14 a. *Die Sozialarbeiter liefen durch den Bahnhof.*
 b. *Wegen der schönen Wetterprognose trugen mehrere der Männer keine Jacke.*
 c. *Wegen der schönen Wetterprognose trugen mehrere der Frauen keine Jacke.*

Les résultats de cette étude suggèrent que ces différences au niveau de la structure des langues correspondent à des différences d'interprétation dans les trois langues. En effet, les résultats montrent que les locuteurs francophones et germanophones ont tendance à interpréter les noms de métier présentés dans l'étude sous la forme masculine générique comme référant à des groupes d'hommes, même quand le métier en question est associé à un stéréotype féminin (par exemple *les assistants sociaux*) : les participants francophones et germanophones trouvent les continuations masculines (les phrases en 13b et 14b) plus naturelles que les continuations féminines (les phrases

31. Gabriel, Gygas, Sarrasin, Garnham & Oakhill (2008).

en 13c et 14c) quel que soit le stéréotype, et y compris quand le stéréotype est féminin, comme en (13) et (14). En l'absence de genre grammatical dans leur langue, les anglophones donnent en revanche des réponses qui sont en accord avec les stéréotypes : quand le stéréotype est masculin, l'interprétation masculine est privilégiée et quand le stéréotype est féminin, l'interprétation féminine est privilégiée. Par exemple, les anglophones jugent la continuation en (12c) plus naturelle que la continuation en (12b), dans la mesure où le stéréotype correspondant au métier de l'assistance sociale est féminin. Les résultats révèlent donc un contraste très fort entre les réponses d'un côté des locuteurs de langues utilisant le masculin générique (les francophones et germanophones) par rapport aux locuteurs de langues où il n'y a pas de masculin générique (les anglophones).

Une étude de 2013³² confirme ces résultats et montre en plus que des participants bilingues font des inférences différentes par rapport au genre social du référent d'un nom selon qu'ils lisent en français ou en anglais. Ces résultats complémentaires sont importants : ils montrent que la tendance à interpréter les masculins génériques de façon spécifique est un biais proprement linguistique, plutôt qu'un biais culturel lié à des locuteurs spécifiques.

Les deux études mentionnées ci-dessus, ainsi que d'autres études similaires³³, ont établi que les masculins génériques biaisaient l'interprétation vers des référents masculins, indépendamment des stéréotypes. Mais qu'en est-il des formes d'écriture inclusive ? Permettent-elles de corriger ce biais, comme le prétendent les défenseurs de l'écriture inclusive ?

C'est la question à laquelle a cherché à répondre une étude menée en 2023³⁴ en regardant si l'utilisation de formes inclusives, avec coordination (*musiciens et musiciennes*) ou avec le point médian (*musicien·ne·s*), permettait de diminuer le biais masculin observé avec les masculins génériques (*musiciens*). La méthode est très proche de celle des études citées ci-dessus. On présente aux participants un texte avec un mot sous l'une des trois formes (coordination, point médian, masculin générique), en faisant varier le stéréotype du nom (stéréotype masculin, par ex. *électricien*, stéréotype neutre, par ex. *musicien*, stéréotype féminin, par ex. *caissier*).

Les résultats de cette étude de 2023 sont très intéressants. Ils montrent que la forme masculine générique introduit un biais masculin pour les métiers stéréotypiquement féminins (*caissier*) ou neutres (*musicien*), mais pas pour les métiers stéréotypiquement masculins (*électricien*). Une interprétation possible pour l'absence d'effet avec les métiers stéréotypiquement masculins serait un effet plafond : les inférences masculines étant déjà très fortes en raison du stéréotype masculin, il n'y a pas une grande marge de manœuvre pour un effet du genre grammatical.

Les résultats pour les formes inclusives sont encore plus intéressants. Ces formes sont en effet interprétées de façon neutre uniquement pour les stéréotypes neutres. Pour les stéréotypes masculin et féminin, les formes inclusives ont un effet égalitaire par-delà les stéréotypes : elles conduisent en effet à un biais masculin dans les métiers stéréotypiquement féminins (les participants surévaluent le nombre d'hommes dans les métiers typiquement féminins, comme *caissier*) et à un biais féminin dans les métiers stéréotypiquement masculins (les participants surévaluent le nombre de femmes dans

32. Sato, Gyga & Gabriel (2013).

33. Voir Gyga, Zufferey & Ute (2021) pour un aperçu plus exhaustif.

34. Xiao, Strickland & Peperkamp (2023).

les métiers typiquement masculins, comme *électricien*). Ce qui est intéressant, c'est que les formes inclusives ne reflètent pas les stéréotypes de façon fidèle mais les atténuent.

4.2.1. Une étude qui contrôle l'effet de la compétition morphologique

Les études citées ci-dessus montrent toutes que le masculin générique est associé à un biais masculin dans l'interprétation. Cependant ce résultat est établi en utilisant des formes masculines qui ont un compétiteur morphologique féminin. Par exemple, la forme masculine *les musiciens* a un compétiteur féminin : *les musiciennes*. On pourrait imaginer que le biais masculin n'est pas lié directement à la forme masculine mais indirectement à l'existence d'un compétiteur féminin. Par un raisonnement pragmatique, les participants pourraient avoir donné plus de réponses masculines parce qu'ils se seraient attendus à voir la forme féminine si le locuteur voulait désigner spécifiquement des femmes. On se retrouve donc face à un problème désormais familier : la présence d'un facteur de confusion (voir partie 2.2.2).

La compétition joue sans doute un rôle dans les inférences masculines des masculins génériques. Il a été en effet montré dans une étude de 2008³⁵ que les masculins génériques créent un biais masculin plus fort en présence de formes féminines dans le même texte qu'en l'absence de ces formes féminines. La question qui se pose alors est la suivante : est-ce que le biais masculin pourrait s'expliquer entièrement par la présence de formes féminines alternatives dans la langue ? Ou bien existe-t-il un biais masculin des formes masculines même en l'absence de compétiteur morphologique ?

C'est la question que se pose une étude à paraître dans le journal *Linguistics Vanguard* et que j'ai menée avec une collègue de l'Université de Lausanne, Laura Delaloye Saillen³⁶. Dans cette étude, nous nous sommes concentrés sur des noms à interprétation générique mais qui n'ont pas de compétiteur morphologique en raison de leur invariabilité en genre. Il s'agit de mots comme *personne* (féminin) et *individu* (masculin) ou encore *vedette* (féminin) et *talent* (masculin). Ces mots ont un genre grammatical fixe, mais une interprétation sémantique générique, pouvant renvoyer aussi bien à des femmes qu'à des hommes. Ces mots présentent encore un autre intérêt : ils permettent également de tester le rôle du genre grammatical féminin dans les inférences de genre social, ce qui n'était pas possible dans les études antérieures.

Le protocole expérimental utilisé dans cette étude est très proche de celui des études antérieures, s'appuyant sur des données de jugement. Les participants étaient confrontés à des phrases comme (15a) ou (15b) et devaient juger si, selon eux, la phrase parlait plutôt d'une femme ou d'un homme, en utilisant une échelle de 1 à 7. Des paires de mots avec des sens très proches, comme *vedette* et *talent* ou *personne* et *individu*, ont été utilisées afin de contrôler au maximum les effets de stéréotypes qui pourraient s'ajouter aux effets du genre grammatical. De plus, les mots d'une même paire étaient insérés dans un même contexte phrastique, comme en (15), afin de minimiser les différences entre les deux phrases à juger.

- 15 a. *Une vedette de la chanson a été invitée pour présider le jury.*
 b. *Un talent de la chanson a été invité pour présider le jury.*

Les résultats montrent qu'il y a bien un biais masculin avec les génériques masculins, comme *individu* ou *talent*, mais pas avec les noms féminins correspondants, comme *personne* et *vedette*. Ces résultats sont intéressants parce qu'ils permettent d'établir que

35. Gygax & Gabriel (2008).

36. Storme & Delaloye Saillen (2024).

le genre grammatical introduit un biais même en l'absence de compétiteur morphologique direct. Par ailleurs, ils suggèrent qu'il n'existe un biais grammatical qu'avec le masculin, le féminin étant, lui, interprété de façon neutre. Cependant, comme nous le suggérons dans l'article, ce deuxième résultat est à prendre avec précaution parce que nous n'avons pas contrôlé l'effet des stéréotypes à travers les contextes phrastiques et les paires de mots. Les phrases et les paires de mots utilisées comme stimuli pourraient présenter un biais masculin global. Ce pourrait être le cas par exemple si on pense plutôt à des hommes pour présider un jury, dans les phrases en (15), et si on pense qu'un homme a plus de chance d'avoir du succès dans la chanson, et par conséquent d'être qualifié de vedette/talent de la chanson.

4.3. Effet de l'écriture inclusive sur la vitesse de lecture : une étude expérimentale

Les études sur l'effet de l'écriture inclusive sur la rapidité du traitement cognitif, et plus spécifiquement sur la lecture, sont beaucoup moins nombreuses en comparaison des études sur le biais masculin des masculins génériques. Je ne mentionnerai qu'une de ces études ici, qui a été menée en 2007³⁷. L'hypothèse des auteurs est que les formes d'écriture inclusive, en particulier les formes impliquant une typographie particulière comme le tiret (*les musicien-ne-s*) causent un ralentissement à la première lecture de ces formes mais qu'il y a un effet d'habituation rapide quand le lecteur rencontre d'autres occurrences, en vertu de l'effet de la fréquence lexicale discuté dans la partie 3.2.

L'étude de 2007 teste l'effet de plusieurs stratégies sur la vitesse de lecture : le masculin (*les pharmaciens*), le féminin (*les pharmaciennes*), une forme épïcène avec un tiret (*les pharmacien-ne-s*) et une forme épïcène avec coordination (*les pharmaciens et pharmaciennes*). Cinq noms de métier ont été choisis, deux correspondant à un stéréotype féminin (*infirmiers, esthéticiens*), deux correspondant à un stéréotype masculin (*maçons, avocats*), et un nom de métier correspondant à un stéréotype neutre (*pharmaciens*). Chaque participant lisait un texte contenant trois occurrences de l'un des 20 noms sur un ordinateur. Les textes étaient présentés en plusieurs parties, le participant devant cliquer sur un bouton pour passer de l'une à l'autre. Ce dispositif permettait d'obtenir un temps de lecture pour chaque partie du texte et de calculer ainsi le temps de lecture moyen par mot par participant. Une fois les données collectées, les auteurs ont calculé l'effet des différentes stratégies d'écriture (masculin générique, féminin, forme épïcène avec un tiret et coordination) sur le temps de lecture pour chacune des trois occurrences, en comparant le temps de lecture effectif des passages contenant ces formes avec le temps moyen attendu.

Les résultats montrent que la première occurrence du nom est associée à un ralentissement du temps de lecture pour la forme épïcène avec un tiret (*les pharmacien-ne-s*) par rapport à la forme masculine (*les pharmaciens*). Cela va dans le sens des détracteurs de l'écriture inclusive, qui lui reprochent une lourdeur à la lecture. Cependant cet effet disparaît sur les occurrences suivantes du mot, suggérant une adaptation rapide des participants, en accord avec les prédictions des auteurs.

Les auteurs ne trouvent pas d'impact de la stratégie de coordination (*les pharmaciens et pharmaciennes*) sur la vitesse de lecture. Mais cette absence d'effet est sans doute due au fait que les auteurs utilisent le temps de lecture moyen d'un mot dans leur calcul. Or la forme coordonnée contient trois mots. Il faudrait comparer le temps

37. Gygax & Gesto (2007).

de lecture de ces trois mots (*pharmaciens et pharmaciennes*) par rapport au temps de lecture d'un mot unique (*pharmaciens*). Les résultats négatifs sur la forme d'écriture inclusive avec coordination sont donc à interpréter avec précaution.

4.4. Conclusion sur le langage inclusif et l'expérimentation

L'expérimentation permet de mieux cerner les liens de causalité entre variables par rapport aux études observationnelles, en raison du plus grand degré de contrôle qu'elle permet. Les études mentionnées dans cette partie suggèrent qu'il y a bien un effet du genre grammatical masculin sur la façon dont on interprète le genre social du référent correspondant, et cela même quand on contrôle le rôle des stéréotypes et le rôle de la compétition morphologique. L'étude que j'ai menée avec Laura Delaloye Saillen suggère par ailleurs qu'on ne trouve pas de biais symétrique avec les noms féminins génériques. Mais comme cela a été dit, ce résultat est pour l'instant à prendre avec précaution, en raison de l'absence de contrôle des stéréotypes dans cette étude.

Les données expérimentales présentées ici ont aussi des limites qu'il faut bien garder en tête mais qui sont largement indépendantes de l'aspect expérimental. Ces études utilisent essentiellement des données déclarées (des jugements). Or les données déclarées ne reflètent pas nécessairement les inférences réelles que font les individus. Par ailleurs, ces études s'appuient souvent sur des échantillons de convenance, comme des étudiants. Cela peut limiter la possibilité de généraliser les résultats à l'ensemble de la population.

De façon générale, une étude ne suffit pas pour établir un résultat. Il faut plusieurs études convergentes. Une autre limite est que les études mentionnées se concentrent sur des effets d'interprétation à court terme. Mais on peut se demander quels sont les effets à plus long terme de ces biais grammaticaux. Sont-ils suffisants pour construire une vision du genre social fondamentalement biaisée ? Je renvoie à la littérature riche sur l'écriture inclusive pour approfondir ces questions³⁸.

5. MODÉLISATION DES THÉORIES FONCTIONNALISTES AVEC LES GRAMMAIRES À CONTRAINTES

À côté de l'expérimentation, nous avons introduit une autre méthode empruntée aux sciences naturelles : la modélisation. Dans la partie 2, nous avons vu comment cette méthode, et en particulier la modélisation statistique, permettait de formuler les théories linguistiques de façon entièrement explicite tout en rendant compte de la variabilité inhérente aux données linguistiques. Le but de cette partie 5 est d'introduire une façon spécifique de modéliser les données linguistiques : les grammaires probabilistes à contraintes. Les grammaires à contraintes correspondent à une formalisation des analyses de type fonctionnaliste en linguistique, et sont ainsi particulièrement adéquates pour le linguiste qui cherche non seulement à décrire les phénomènes mais aussi à les expliquer.

La partie 5.1 introduit et motive brièvement l'hypothèse fonctionnaliste en linguistique. La partie 5.2 montre comment cette théorie peut être formalisée avec la notion de contrainte, en se focalisant sur des exemples issus de la phonologie. La partie 5.3 présente la mécanique des grammaires à contraintes. La partie 5.4 explique comment ces grammaires à contraintes peuvent être rendues probabilistes pour rendre compte de la variation.

38. Gygax, Zufferey & Gabriel (2021).

5.1. Le fonctionnalisme en linguistique

L'hypothèse centrale du courant fonctionnaliste en linguistique est que la forme linguistique peut s'expliquer en partie par la fonction communicative du langage. André Martinet, l'un des plus éminents représentants de ce courant en France, explique par exemple :

L'évolution linguistique peut être conçue comme régie par l'antinomie permanente entre les besoins communicatifs de l'homme et sa tendance à réduire au minimum son activité mentale et physique. (Martinet 1980 [1960¹] : 176)

Cette hypothèse a été utilisée notamment pour expliquer l'observation de Zipf selon laquelle la longueur des mots est inversement corrélée à leur fréquence, les mots plus fréquents ayant tendance à être plus courts³⁹ : les mots fréquents étant plus saillants dans la mémoire et donc plus rapidement accessibles pour un interlocuteur, il est possible de réduire l'effort articulatoire nécessaire à leur production, et ainsi de les raccourcir. On a vu aussi que l'hypothèse fonctionnaliste était centrale pour expliquer les résultats de l'étude sur le marquage différentiel du cas mentionnée dans la partie 2.2.3 de cet article : les participants marquent davantage la différence entre sujet et objet quand il y a une ambiguïté potentielle pour l'interlocuteur (avec deux animés comme arguments du verbe d'action) que lorsqu'il n'y en a pas (avec un animé et un inanimé comme arguments du verbe d'action)⁴⁰.

Cependant les analyses d'inspiration fonctionnaliste ont pu être jugées prématurées, notamment par Chomsky dans l'extrait suivant, où il oppose à cette approche une approche formelle plus précise et plus attentive aux détails dans la description des phénomènes linguistiques :

It seems to me that the most hopeful approach today is to describe the phenomena of language and of mental activity as accurately as possible, to try to develop an abstract theoretical apparatus that will as far as possible account for these phenomena and reveal the principles of their organization and functioning, without attempting, for the present, to relate the postulated mental structures and processes to any physiological mechanisms or to interpret mental function in terms of « physical causes ». (Chomsky 1968 : 12)

« Il me semble que l'approche la plus prometteuse aujourd'hui consiste à décrire les phénomènes du langage et de l'activité mentale aussi correctement que possible, à développer un système théorique abstrait qui rendra compte autant que possible de ces phénomènes et révélera les principes de leur organisation et de leur fonctionnement, sans tenter, pour le moment, de relier les structures et les processus mentaux postulés à quelque mécanisme physiologique ou d'interpréter la fonction mentale en termes de "causes physiques". » (traduction personnelle)

Parmi les griefs de Chomsky envers l'approche fonctionnaliste, on trouve l'idée que la présence massive d'ambiguïté dans les langues est incompatible avec l'hypothèse d'une optimisation de la communication à travers les langues. Un autre grief concerne l'absence de formalisation de ces théories, qui contraste avec ce qu'on trouve dans les analyses générativistes, où les phénomènes linguistiques sont modélisés de façon précise au travers de règles, comme le type de règle syntaxique utilisé dans l'étude de l'apprentissage de la syntaxe que l'on a vue dans la partie 2.3.2⁴¹.

39. Zipf (1965 [1935¹] : 25).

40. Fedzechkina, Jaeger & Newport (2012).

41. Perfors & Tenenbaum (2011).

Cependant, dès les années 70, on observe un effort de formalisation des théories fonctionnalistes. L'exemple le plus emblématique est sans doute l'article de Liljencrants et Lindblom paru en 1972 dans le journal *Language*⁴². Cet article propose un modèle mathématique entièrement explicite qui dérive, à partir de l'idée que les langues cherchent à maximiser le contraste entre les sons qu'elles utilisent, les grandes tendances observées dans la structure des inventaires vocaliques à travers les langues, et notamment la prédominance des voyelles cardinales [a], [i] et [u], qui se trouvent être les voyelles les plus distinctes qui puissent être produites par l'appareil vocal humain. Plus récemment, un article paru en 2019 dans le journal *Trends in Cognitive Sciences* recense les travaux de modélisation du langage d'inspiration fonctionnaliste, allant du lexique jusqu'à la syntaxe⁴³.

5.2. Motivations pour les grammaires à contraintes

Dans cet article, je me concentre sur un modèle formel spécifique pour les approches fonctionnalistes : les grammaires à contraintes, connues aussi sous le nom de « théorie de l'optimalité ». Ce modèle a été proposé dans les années 90 par Alan Prince et Paul Smolensky⁴⁴ comme une alternative aux grammaires à règles qui avaient dominé la période antérieure⁴⁵.

Dans les grammaires à contraintes, comme dans le système basé sur des règles décrit dans la partie 2.3.2 pour la syntaxe, la grammaire associe des formes sous-jacentes, correspondant à la forme abstraite des mots (notées entre /.../), à des formes de surface, correspondant à la forme concrète des mots en contexte (notées entre [...])⁴⁶. L'association entre formes sous-jacentes et formes de surface est désormais régulée non plus par des règles, mais par des contraintes. Il y a deux types de contraintes : les contraintes de fidélité, dont le but est de préserver les distinctions lexicales, et les contraintes de marque, dont le but est de bannir les structures difficiles d'un point de vue articulatoire, perceptuel, cognitif, etc. Les contraintes sont en conflit : on ne peut pas préserver toutes les distinctions lexicales possibles sans créer des structures marquées en surface. Les langues varient dans leur façon de résoudre ces conflits, c'est-à-dire dans l'importance qu'elles accordent à ces différentes contraintes. Ce modèle se veut à la fois un modèle de la grammaire des individus et un modèle qui peut rendre compte de la typologie des langues.

Les grammaires à contraintes répondaient à une difficulté des modèles à règles à rendre compte de l'unité fonctionnelle des processus linguistiques (les « conspirations »), en particulier des processus phonologiques. Cette approche était aussi motivée par un besoin de rendre la théorie phonologique non seulement descriptive mais aussi explicative, besoin qui se manifeste déjà très clairement dans l'article de Liljencrants et Lindblom cité plus haut. Comme on le verra, le formalisme des contraintes convient bien aux recherches d'inspiration fonctionnaliste, puisqu'il permet de formuler comme des primitives de la grammaire les besoins communicatifs des locuteurs.

42. Liljencrants & Lindblom (1972).

43. Gibson, Futrell, Piantadosi, Dautriche, Mahowald, Bergen & Levy (2019).

44. Prince & Smolensky (1993).

45. Pour des exemples de grammaires à règles, voir la partie 2.3.2 sur la modélisation syntaxique avec des règles probabilistes. En phonologie, les grammaires à règles ont été proposées par Chomsky & Halle (1968).

46. Dans les grammaires syntaxiques de la partie 2.3.2, les règles syntaxiques associent aussi des structures abstraites, les syntagmes, à des structures concrètes, les séquences de mots.

La notion de « conspiration » des processus phonologiques, qui est une des motivations centrales des grammaires à contraintes en phonologie, peut être illustrée avec des données du français. En français, on observe en effet un ensemble de processus phonologiques, qui ont des manifestations très disparates mais semblent être motivés par la même contrainte. Il s'agit de l'élision (la chute du /a/ dans *l'amie* au lieu de la forme attendue **la amie*), de l'épenthèse (l'ajout d'une consonne [t] dans le verbe *pianoter*, formé sur le nom *piano* avec le suffixe verbal -er) ou encore de la liaison (l'utilisation de la forme avec liaison *gran[t]* dans le groupe nominal *grand ami*). Ces trois processus sont très différents dans leur manifestation mais semblent pourtant converger vers le même but : éviter la présence d'un hiatus, c'est-à-dire d'une séquence de deux voyelles. Dans une analyse avec des règles, il faut trois règles indépendantes (élision, épenthèse et liaison). Dans une analyse avec des contraintes, ces processus permettent d'éviter un hiatus, et donc de satisfaire une unique contrainte que l'on va noter *VV et qui pénalise une séquence de deux voyelles dans les formes de surface. Cette contrainte est violée par les formes **la amie* (sans élision du [a]), **pianoer* (sans épenthèse du [t]) et *grand ami* (sans [t] de liaison), mais pas par les formes observées en français : *l'amie* (avec élision), *pianoter* (avec épenthèse du [t]) et *gran[t] ami* (avec liaison).

5.3. La mécanique des grammaires à contraintes

Dans cette partie, nous allons illustrer plus précisément le fonctionnement des grammaires à contraintes⁴⁷, avec un exemple construit et simplifié. Imaginons deux langues A et B qui présentent la distinction suivante. La langue A a une alternance impliquant deux formes [bat] et [ba] pour une racine sous-jacente /bat/. Cette langue élide les consonnes finales. La consonne finale de la racine /bat/ est élidée dans le mot non suffixé, [ba], mais est protégée dans la forme suffixée [bata], formée avec le suffixe /-a/, comme montré en (16a). Dans la langue B, le /t/ à la fin de la racine /bat/ apparaît dans toutes les formes du mot, qu'il soit non suffixé dans la forme [bat] ou suffixé dans la forme [bata], comme montré en (16b).

16 a. Langue A : /bat/ → [ba], /bat-a/ → [bata]

b. Langue B : /bat/ → [bat], /bat-a/ → [bata]

Comment cette différence entre les langues A et B est-elle modélisée dans le modèle à contraintes ? Intuitivement la langue A n'aime pas les consonnes finales. On pose donc une contrainte, dite de marque, qui pénalise les consonnes finales. Cette contrainte sera notée *C#. La langue B quant à elle préfère protéger ses consonnes. On pose donc une contrainte, dite de fidélité, qui pénalise l'élision des consonnes. Cette contrainte sera notée Max(C). Dans la langue A, la contrainte de marque *C# est plus forte que la contrainte de fidélité Max(C), ce qui conduit à l'élision de la consonne finale : /bat/ → [ba]. Dans la langue B, la contrainte de marque *C# est moins forte que la contrainte de fidélité Max(C), ce qui conduit au maintien de la consonne finale : /bat/ → [bat].

Dans un système avec des règles, on caractériserait la différence entre les langues A et B de la façon suivante. La langue A aurait une règle qui élide les consonnes (C → Ø / _ #) alors que la langue B n'aurait pas cette règle. Dans le système avec contraintes, le processus d'élision n'est pas exprimé directement dans une règle mais émerge de l'importance relative des deux contraintes : dans la langue A avec élision du [t] final, la contrainte de marque qui pénalise les consonnes finales (*C#) a une plus grande importance que la contrainte de fidélité qui pénalise l'élision (Max(C)). Dans

47. Pour une introduction plus détaillée, voir Kager (1999).

ce système, un processus phonologique est observé quand une contrainte de marque domine une contrainte de fidélité. Dans la langue B, qui n'a pas d'élimination, la contrainte de fidélité contre l'élimination (Max(C)) a une plus grande importance que la contrainte de marque qui pénalise les consonnes finales (*C#). La différence entre les langues A et B se résume donc à une différence entre l'importance relative de deux contraintes présentes dans les deux langues, comme montré en (17).

17 a. Langue A : *C# > Max(C)

b. Langue B : Max(C) > *C#

Une différence importante avec le modèle basé sur les règles tient à ce que les contraintes ne génèrent pas directement la forme de surface à partir de la forme sous-jacente. Elles se contentent d'évaluer un ensemble de candidats pour une forme sous-jacente donnée. Le système linguistique doit donc inclure une composante supplémentaire qui génère tous les candidats possibles pour une forme sous-jacente donnée. En pratique, on considère toujours un nombre fini de candidats qui varient minimalement par rapport à la forme sous-jacente (par ex. [ba], [bat], [bad], [ban], [pat], etc. pour une forme sous-jacente /pat/). En effet, les candidats qui diffèrent trop de la forme sous-jacente (par exemple [gub] pour une forme sous-jacente /pat/) seront souvent lourdement pénalisés par les contraintes de fidélité.

L'architecture générale d'une grammaire à contraintes comprend quatre composantes⁴⁸ :

1. le lexique, qui contient les représentations lexicales ou formes sous-jacentes des morphèmes, qui servent d'*input* au générateur ;
2. le générateur (Gen), qui génère les candidats pour une forme sous-jacente et les soumet à l'évaluateur ;
3. l'évaluateur (Eval), qui sélectionne parmi les candidats le candidat optimal pour la forme sous-jacente, à partir des rapports de priorité qu'a établis cette langue entre des contraintes phonologiques universelles ;
4. l'ensemble de toutes les contraintes phonologiques (Con).

Ainsi le générateur prend un *input* et génère les candidats, comme montré en (18a). Puis l'évaluateur prend cet ensemble de candidats et choisit parmi eux le candidat optimal, comme montré en (18b).

18 a. Gen(*input*) = {candidat 1, candidat 2, candidat 3, ...}

b. Eval({candidat 1, candidat 2, candidat 3, ...}) = *output*

Concentrons-nous maintenant sur l'évaluation des candidats. Il y a en fait deux façons d'évaluer le coût d'un candidat par rapport à un ensemble de contraintes, qui correspondent à deux variantes de grammaires à contraintes : la théorie de l'optimalité (*Optimality Theory*), où les contraintes sont *classées*, et la grammaire harmonique (*Harmonic Grammar*), où les contraintes sont *pondérées*. Ces deux variantes sont très proches conceptuellement mais diffèrent de façon subtile dans leurs prédictions et leur fonctionnement⁴⁹. On ne rentrera pas ici dans les détails de cette comparaison et on se concentrera uniquement sur la grammaire harmonique, qui a une variante probabiliste très facile à mettre en application (voir partie 5.4).

Dans la grammaire harmonique, le coût d'une forme linguistique est calculé au moyen d'une somme pondérée du nombre de violations de toutes les contraintes par

48. Kager (1999 : 18-19)

49. Voir par exemple Magri (2013).

cette forme. Le candidat avec le plus petit coût gagne. Ce coût est noté traditionnellement h (pour harmonie). On note ainsi $h(x)$ le coût d’une forme x . Pour une forme linguistique x , avec n contraintes C_i (i compris entre 1 et n), et un poids w_i (w pour *weight*) pour chaque contrainte C_i , on définit le coût de x de la façon suivante :

$$h(x) = C_1(x) \times w_1 + C_2(x) \times w_2 + \dots + C_n(x) \times w_n$$

où $C_i(x)$ correspond au nombre de fois que la forme x viole la contrainte C_i .

Pour illustrer, supposons que l’on veuille dériver la prononciation [ba] de /bat/ dans la langue A. L’analyse est présentée dans le Tableau 1. L’évaluation est représentée dans un tableau, incluant l’*input* /bat/ en haut de la colonne de gauche, les candidats [bat] et [ba] (sous l’*input*) et les contraintes avec leur poids respectif dans les colonnes deux et suivantes. L’harmonie est calculée dans la colonne de droite comme la somme pondérée des violations de contraintes pour chacun des candidats considérés, [ba] et [bat]. Pour dériver une préférence pour [ba], il faut que la contrainte pénalisant la consonne finale (*C#) ait un poids plus élevé que la contrainte pénalisant l’élision d’une consonne (Max(C)). L’harmonie est plus petite pour le candidat [ba], ce qui correspond à un coût plus petit pour cette forme.

Tableau 1. Analyse dérivant la prononciation [ba] de /bat/.

/bat/	*C# $w=2$	Max(C) $w=1$	Harmonie (h)
[bat]	1	0	2
[ba]	0	1	1

Pour obtenir la langue B, il suffirait de donner un poids plus important à Max(C) par rapport à *C#. C’est alors la forme fidèle à l’*input*, [bat], qui l’emporterait, comme montré dans le Tableau 2.

Tableau 2. Analyse dérivant la prononciation [bat] de /bat/.

/bat/	*C# $w=1$	Max(C) $w=2$	Harmonie (h)
[bat]	1	0	1
[ba]	0	1	2

Qu’a-t-on gagné en faisant cette analyse avec des contraintes par rapport à une simple description de la prononciation de /bat/ comme impliquant une élision des consonnes finales dans un cas et aucun processus dans l’autre cas ? Le grand avantage est que l’on a gagné en abstraction et donc en généralité. On explique des formes linguistiques en fonction de principes plus fondamentaux et avec une interprétation linguistique : d’une part le désir de maintenir des distinctions lexicales (ce qui pousse à préserver les consonnes finales) et d’autre part le désir d’évacuer des formes qui sont difficiles à prononcer ou percevoir (ce qui pousse à effacer les consonnes finales). On a donc une analyse qui est proprement linguistique et explicative. Par ailleurs on a une caractérisation très intuitive de la variation linguistique comme différentes façons de résoudre un conflit entre différents principes linguistiques : dans la langue A, *C# est plus importante que Max(C), dans la langue B, Max(C) est plus importante que *C#. On verra aussi dans la partie 6.1 que ce modèle permet de faire des prédictions non triviales sur la typologie du syncrétisme morphologique.

5.4. Les grammaires à contraintes probabilistes

Les grammaires à contraintes sont également intéressantes parce qu'il existe des variantes probabilistes qui permettent de rendre compte de la variation observée au sein d'une langue. Comme on l'a vu dans la partie 2.3.1, la variation est omniprésente dans les langues, y compris à l'échelle d'un même individu qui peut varier dans sa prononciation d'un même mot. Le modèle présenté dans la partie 5.3 ne peut rendre compte de la variation dans les usages des individus ou d'un même individu car il est déterministe : il ne dérive qu'une prononciation possible pour un *input* donné. En effet, on a vu que la grammaire harmonique ne dérivait qu'un seul candidat pour une forme sous-jacente donnée, celle qui a le plus petit coût. On retombe ainsi sur le problème de Sapir abordé dans la partie 2.3.1 : les grammaires ne parviennent pas à rendre compte de la variabilité (*All grammars leak*). Mais, comme les grammaires à règles, les grammaires à contraintes ont des variantes probabilistes permettant de surmonter ce problème.

Il existe deux grandes façons de rendre les grammaires harmoniques probabilistes⁵⁰. Soit on spécifie que les poids des contraintes sont des variables aléatoires, permettant ainsi à une contrainte de l'emporter ou non sur une autre contrainte pour deux évaluations distinctes du même mot⁵¹. Soit on utilise une approche qui permet directement de dériver des probabilités à partir des scores d'harmonie, comme dans les grammaires MaxEnt (pour *maximum entropy*)⁵². Les deux approches diffèrent de façon subtile dans leurs prédictions⁵³. On va utiliser ici la seconde approche car elle est plus facile à mettre en œuvre d'un point de vue technique.

Dans les grammaires MaxEnt, on utilise les coûts (ou scores d'harmonie) pour définir, non pas un candidat vainqueur unique, mais une distribution de probabilité pour les différents candidats. Pour résumer, on utilise une procédure qui permet qu'un candidat avec un plus petit coût ait une probabilité plus grande qu'un candidat avec un coût plus important. Pour arriver à ce résultat, trois opérations sont appliquées à l'harmonie. On prend d'abord l'opposé de l'harmonie, $-h(x)$, ce qui permet de transformer un coût plus petit en une valeur plus grande (puisque un *plus petit* coût devrait correspondre à une *plus grande* probabilité d'être utilisé). Ensuite, on prend l'exponentielle du résultat, $e^{-h(x)}$, qu'on appelle la valeur maxent du candidat. L'exponentiation permet non seulement de passer à des valeurs positives (les probabilités sont toujours positives, comprises entre 0 et 1) mais aussi de pénaliser de façon exponentielle (cela crée une plus grande asymétrie entre des candidats bons et des candidats moins bons). Enfin, on normalise le résultat en le divisant par la somme des valeurs maxent pour tous les candidats. Cette étape permet d'obtenir que les probabilités de l'ensemble des candidats somment à 1.

Ainsi, pour deux uniques candidats x et y correspondant à une même forme sous-jacente, on obtient leur probabilité à partir de leur harmonie respective $h(x)$ et $h(y)$ de la façon suivante :

$$p(x) = \frac{e^{-h(x)}}{e^{-h(x)} + e^{-h(y)}}$$

$$p(y) = \frac{e^{-h(y)}}{e^{-h(x)} + e^{-h(y)}}$$

50. Hayes (2017).

51. Voir par exemple Boersma & Hayes (2001).

52. Hayes & Wilson (2008). Je renvoie à cet article et aux références qui y sont citées pour un arrière-plan plus théorique sur la notion d'entropie.

53. Flemming (2021).

Appliquons ces formules à l'harmonie calculée pour les formes [ba] et [bat] correspondant à l'*input* /bat/ dans le Tableau 2. On obtient ainsi les probabilités prédites par la grammaire pour les deux prononciations :

$$p([bat]) = \frac{e^{-h([bat])}}{e^{-h([bat])} + e^{-h([ba])}} = \frac{e^{-1}}{e^{-1} + e^{-2}} = 0,73$$

$$p([ba]) = \frac{e^{-h([ba])}}{e^{-h([bat])} + e^{-h([ba])}} = \frac{e^{-2}}{e^{-1} + e^{-2}} = 0,27$$

Cette approche prédit donc que [bat] devrait être plus fréquent que [ba] dans une langue donnée si la contrainte qui pénalise [bat] (*C#) a un poids de 2 alors que la contrainte qui pénalise [ba] (Max(C)) a un poids de 1.

Concrètement, dans une analyse de données montrant de la variation, on fera le raisonnement inverse : à partir des fréquences observées, on remontera aux poids des contraintes correspondantes, en utilisant des méthodes d'inférence statistique comme la régression logistique. La modélisation statistique permet ainsi, à partir de données bruitées et variables, de remonter aux principes sous-jacents qui façonnent les systèmes linguistiques.

6. ÉTUDE DE CAS : VARIATION ENTRE LES LANGUES ET À L'INTÉRIEUR DES LANGUES

Dans cette dernière partie, nous nous tournons vers des applications des grammaires à contraintes pour rendre compte de la variation à travers les langues et au sein d'une même langue. La première étude porte sur la typologie du syncrétisme et montre comment les grammaires à contraintes peuvent dériver quelques généralisations observées dans ce domaine à travers les langues. La deuxième étude porte sur la variation dans la réalisation de la consonne de liaison en français et montre comment cette variation peut être analysée comme résultant d'un conflit entre deux contraintes dans un modèle de grammaire probabiliste.

6.1. La variation entre les langues : la typologie du syncrétisme

L'ambiguïté lexicale est omniprésente dans les langues. Par exemple, le morphème -s de l'anglais marque à la fois le pluriel (*cars* «voitures»), le possessif (*John's car* «la voiture de John») et la troisième personne du singulier au présent (*he swims* «il nage»). Mais l'ambiguïté n'est pas complètement arbitraire et a une distribution qui est contrainte à travers les langues⁵⁴. Dans cette partie, nous allons nous pencher sur un type d'ambiguïté particulier, le syncrétisme, et montrer comment on peut expliquer la typologie du syncrétisme avec un système très simple de contraintes que j'ai proposé dans un article publié dans le *Journal of Linguistics*⁵⁵. La partie 6.1.1 présente deux généralisations typologiques sur le syncrétisme, connues sous le nom d'universaux 44 et 45 de Greenberg⁵⁶. La partie 6.1.2 présente le modèle basé sur contraintes que j'ai proposé pour rendre compte du syncrétisme. La partie 6.1.3 montre comment ce modèle dérive les universaux 44 et 45.

54. Par exemple, voir Piantadosi, Tily & Gibson (2012).

55. Voir Storme (2022a).

56. Voir Greenberg (1963).

6.1.1. Le syncrétisme et les universaux 44 et 45 de Greenberg

Dans les langues, certaines distinctions grammaticales ne sont pas marquées dans certains contextes. Par exemple, en anglais, les pronoms singuliers de la troisième personne ont une distinction de genre : le masculin *he* «il» s'oppose au féminin *she* «elle». Mais cette opposition n'existe pas au pluriel, où seule la forme *they* est utilisée. Cette forme peut être utilisée pour renvoyer à un groupe constitué uniquement d'hommes, un groupe constitué uniquement de femmes et un groupe mixte. Dans ce cas, on dit que le genre est syncrétique dans le pluriel des pronoms (voir partie 3.1).

Les langues varient dans leur façon d'exprimer les distinctions grammaticales, avec des degrés plus ou moins importants de syncrétisme. Par exemple, dans certaines langues comme l'espagnol, les pronoms de troisième personne ne sont syncrétiques ni au singulier ni au pluriel : la distinction de genre entre masculin et féminin est présente dans les deux nombres, comme le montre le Tableau 3a. Dans d'autres langues, comme l'anglais ou le touareg de Tahhagart, la distinction de genre n'est présente que dans un seul contexte, le singulier et le pluriel respectivement, comme le montrent les Tableaux 3b et 3c. Enfin, certaines langues comme le turc n'expriment la distinction de genre dans aucun de leurs pronoms, qu'il s'agisse du singulier ou du pluriel, comme le montre le Tableau 3d. Dans ce cas, on peut parler d'un syncrétisme maximal.

Tableau 3. Exemples de paradigmes de pronoms de troisième personne avec différents degrés de syncrétisme de genre.

	SG	PL
MASC	<i>él</i>	<i>ellos</i>
FEM	<i>ella</i>	<i>ellas</i>

(a) Espagnol

	SG	PL
MASC	<i>he</i>	<i>they</i>
FEM	<i>she</i>	<i>they</i>

(b) Anglais

	SG	PL
MASC	<i>ənta</i>	<i>əntaniɖ</i>
FEM	<i>ənta</i>	<i>əntanətiɖ</i>

(c) Touareg de Tahhagart

	SG	PL
MASC	<i>o</i>	<i>onlar</i>
FEM	<i>o</i>	<i>onlar</i>

(d) Turc

Les quatre paradigmes du Tableau 3 représentent les quatre stratégies de syncrétisme théoriquement possibles à partir d'un système présentant une distinction binaire (ici le genre) dans deux contextes (ici le singulier et le pluriel). Cependant, toutes ces stratégies d'expression du genre ne sont pas représentées de la même façon à travers les langues. En particulier, les langues qui n'ont une distinction de genre qu'au pluriel (comme le touareg) sont très peu nombreuses par rapport aux langues qui n'ont une distinction de genre qu'au singulier (comme l'anglais). Cette généralisation a été

pour la première fois décrite par Greenberg en 1963 sous le nom d'universel 45. Cet universel stipule que si une langue a des distinctions de genre au pluriel, alors elle a aussi des distinctions de genre au singulier. Cette généralisation a été confirmée plus tard empiriquement : dans un échantillon de 360 langues décrit dans le *World Atlas of Language Structures*, on trouve 42 langues avec une distinction de genre limitée au singulier mais seulement une langue avec une distinction de genre limitée au pluriel (il s'agit du dagaare, une langue parlée au Ghana)⁵⁷.

Globalement, le syncrétisme est souvent favorisé dans certains types de contextes à travers les langues. On peut citer comme autre exemple l'expression des distinctions de genre en fonction de la personne, qui est décrite par l'universel 44 de Greenberg : si une langue a des distinctions de genre dans les personnes du discours (les personnes 1 et 2 : *je, tu*), alors en général elle a aussi des distinctions de genre dans la troisième personne (*il/elle*). Par exemple, dans un échantillon de 378 langues décrit dans le *World Atlas of Language Structure*, on trouve 18 langues avec une distinction de genre à la troisième personne du singulier seulement, mais seulement deux langues avec une distinction de genre aux personnes du discours seulement (les personnes 1 et 2)⁵⁸. Les personnes du discours favorisent donc le syncrétisme de genre par rapport à la troisième personne, de la même façon que le pluriel le favorise par rapport au singulier.

Ces tendances observées à travers les langues appellent une explication : pourquoi certains contextes favorisent-ils le syncrétisme ? Pourquoi les généralisations impliquant le syncrétisme ont-elles une forme implicationnelle, comme les universaux 44 et 45 ? Et de façon plus fondamentale, quelle est la source du syncrétisme ?

6.1.2. Modélisation à l'aide des grammaires à contraintes

Dans mon article paru en 2022 dans le *Journal of Linguistics*, je propose que le syncrétisme est façonné par le conflit entre deux contraintes antagonistes : une contrainte qui pénalise les paradigmes morphologiques qui contiennent beaucoup de formes (qui correspond aux contraintes de marque introduites précédemment) et une contrainte qui pénalise les erreurs d'interprétation dans la communication (qui correspond aux contraintes de fidélité introduites précédemment). La première contrainte favorise des petits paradigmes morphologiques, et donc le syncrétisme. En effet, plus un paradigme présente de syncrétismes, plus le nombre de formes qu'il contient est réduit, et ainsi plus il est facile à mémoriser et à utiliser. La seconde contrainte favorise quant à elle les paradigmes morphologiques qui reflètent de façon transparente les distinctions grammaticales sous-jacentes de la langue dans tous les contextes. Cette contrainte défavorise donc le syncrétisme.

D'un point de vue formel, l'interaction de ces contraintes est modélisée en utilisant une grammaire harmonique comme on l'a vu dans la partie 5.3. La contrainte qui favorise le syncrétisme assigne une violation à un paradigme pour tout contexte où une distinction grammaticale binaire entre deux valeurs t_1 et t_2 ⁵⁹ (par exemple entre masculin et féminin) est marquée morphologiquement, comme détaillé en (19).

57. Siewierska (2013).

58. Siewierska (2013).

59. Le modèle proposé dans Storme (2022a) se concentre sur les traits grammaticaux binaires, comme le genre (qui oppose masculin ou féminin) ou l'aspect grammatical (qui oppose aspect perfectif et imperfectif).

19 Contrainte favorisant le syncrétisme (contrainte sur la taille des paradigmes)

- a. Assigne une pénalité de 0 pour tout contexte où la distinction entre deux traits grammaticaux t_1 et t_2 n'est pas exprimée morphologiquement.
- b. Assigne une pénalité de 1 pour tout contexte où la distinction entre deux traits grammaticaux t_1 et t_2 est exprimée morphologiquement.

Si le syncrétisme permet de réduire la taille d'un paradigme, il pose un problème d'ambiguïté pour l'interlocuteur. La contrainte qui défavorise le syncrétisme dans un contexte donné assigne une violation qui est égale à la probabilité que l'interlocuteur commette une erreur d'interprétation dans ce contexte, comme détaillé en (20).

20 Contrainte défavorisant le syncrétisme (contrainte sur l'ambiguïté des paradigmes)

Pour chaque contexte, assigne une pénalité égale à la probabilité d'identifier incorrectement t_1 et t_2 dans ce contexte.

Dans le cas où une distinction morphologique est faite, cette probabilité est égale à zéro. Par exemple, si une langue distingue masculin et féminin dans un contexte donné, alors il sera possible pour l'interlocuteur d'identifier sans erreur le genre. Dans le cas où aucune distinction n'est faite, on peut considérer que la probabilité d'erreur sera égale à la probabilité de la valeur grammaticale la moins probable dans ce contexte. En effet, la stratégie optimale de désambiguïsation en cas de syncrétisme consiste à choisir la forme la plus probable. Dans un système où les traits grammaticaux sont binaires, l'erreur correspond donc à la probabilité de choisir le trait le moins probable. Cette probabilité d'erreur peut donc se noter comme dans l'équation suivante :

$$p(\text{erreur}, \text{contexte}) = \min\{p(t_1, \text{contexte}), p(t_2, \text{contexte})\}$$

où $p(\text{erreur}, \text{contexte})$ note la probabilité de faire une erreur d'interprétation dans le contexte en question et $p(t_1, \text{contexte})$ et $p(t_2, \text{contexte})$ notent la probabilité de chacun des traits dans chaque contexte. La fonction $\min\{a, b\}$ renvoie le plus petit nombre entre a et b .

Dans la grammaire harmonique, les langues diffèrent par le fait d'attribuer des poids différents à différentes contraintes. Ainsi on s'attend à ce que les langues puissent différer dans le poids qu'elles vont accorder à la contrainte qui favorise le syncrétisme en (19) et par rapport à celle qui défavorise le syncrétisme en (20). Notons w_S le poids de la contrainte favorisant le syncrétisme (w pour *weight* « poids » et S pour *size* « taille ») et w_A le poids de la contrainte défavorisant le syncrétisme (w pour *weight* « poids » et A pour *ambiguity* « ambiguïté »).

Dans un contexte donné c_i , une langue va favoriser ou défavoriser le syncrétisme en fonction du rapport entre les deux coûts. Les contextes dans lesquels l'analyse prédit que syncrétisme et absence de syncrétisme ont le même coût sont définis par l'égalité suivante :

$$w_S = P(\text{erreur}, c_i)w_A$$

Cette équation équivaut à l'équation ci-dessous, par la définition de la probabilité d'erreur donnée plus haut :

$$w_S = \min\{P(t_1, c_i), P(t_2, c_i)\}w_A$$

Et, par la définition de la probabilité conditionnelle issue de la théorie des probabilités, cette équation équivaut elle-même à l'équation suivante :

$$w_S = \min\{P(t_1|c_i), P(t_2|c_i)\}P(c_i)w_A$$

La probabilité conditionnelle $P(t_i|c_i)$ donne la probabilité d'observer le trait grammatical t_i (par exemple, le masculin) en sachant que l'on est dans le contexte c_i , en faisant abstraction de la probabilité de ce contexte. La probabilité du contexte est exprimée par $P(c_i)$.

On peut maintenant réordonner cette équation de façon à obtenir l'équation d'une courbe exprimant les probabilités $P(c_i)$ et $\min\{P(t_1|c_i), P(t_2|c_i)\}$ pour lesquelles il y a autant intérêt à avoir du syncrétisme qu'une absence de syncrétisme :

$$\min\{P(t_1|c_i), P(t_2|c_i)\} = \frac{1}{P(c_i)} \times \frac{w_S}{w_A}$$

(en postulant que les dénominateurs sont non nuls)

Cette courbe peut être transformée en une droite, plus facile à interpréter, en appliquant la transformation logarithmique de chaque côté de l'équation :

$$\log[\min\{P(t_1|c_i), P(t_2|c_i)\}] = -\log P(c_i) + \log\left(\frac{w_S}{w_A}\right)$$

Cette droite est représentée en pointillés dans la Figure 1, avec un choix spécifique pour les poids w_S et w_A ⁶⁰. Elle caractérise les valeurs de probabilité pour le contexte et les traits grammaticaux pour lesquelles le syncrétisme et l'absence de syncrétisme ont le même coût. Au-dessus de la droite, on trouve les contextes pour lesquels l'absence de syncrétisme est préférable. En dessous de la droite, on trouve les contextes pour lesquels le syncrétisme est préférable.

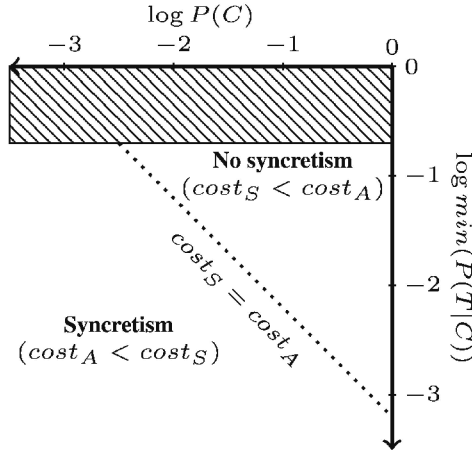


Figure 1. Prédications du modèle à propos des contextes favorisant et défavorisant le syncrétisme (graphique reproduit de Storme 2022a)

Le graphique montre que, d'après les prédictions du modèle, le syncrétisme devrait être préférable dans les contextes qui sont moins fréquents (voir l'axe des x) et dans les

60. La zone grisée en haut du graphique correspond à des valeurs de probabilité supérieures à 0.5. Ces valeurs ne sont pas possibles car la plus petite valeur entre deux valeurs de probabilité sommant à 1 ne peut être supérieure à 0.5.

contextes où il y a une plus grande asymétrie dans la probabilité des traits grammaticaux (voir l'axe des y). Cela se comprend intuitivement : la présence de syncrétisme devrait être moins dommageable dans un contexte qui apparaît moins souvent et dans lequel il y a une plus forte asymétrie entre les probabilités des deux traits t_1 et t_2 . En effet, dans le cas d'une forte asymétrie dans la probabilité des deux traits grammaticaux, on peut par défaut choisir la valeur la plus probable et ne se tromper que dans un petit nombre de cas.

Supposons maintenant que les contextes grammaticaux et les traits grammaticaux aient globalement la même probabilité à travers les langues, c'est-à-dire par exemple que les humains ont tendance globalement à parler avec la même fréquence de référents animés, inanimés, féminins, masculins, singuliers, pluriels, etc. Dans cette hypothèse, le modèle proposé ici prédit que certains contextes devraient universellement favoriser le syncrétisme. En effet, dans cette hypothèse, la seule source de variation dans les langues tient à la différence de pondération des deux contraintes (la différence entre w_S et w_A), ce qui correspond à l'ordonnée à l'origine de la droite représentée dans la Figure 1. Dans les langues où le poids de la contrainte favorisant le syncrétisme (qui correspond au poids w_S) est très élevé, l'ordonnée à l'origine sera plus grande et le nombre de contextes favorisant le syncrétisme plus grand. Dans les langues où le poids de la contrainte défavorisant le syncrétisme (qui correspond au poids w_A) est très élevé, l'ordonnée à l'origine sera plus petite et le nombre de contextes favorisant le syncrétisme plus petit.

Cette approche dérive également des généralisations de type implicationnel, à la Greenberg. Ceci peut être observé sur la Figure 2. Supposons qu'il y ait deux contextes c_1 et c_2 , caractérisés par une certaine probabilité et par une certaine probabilité conditionnelle des traits grammaticaux t_1 et t_2 . Dans la Figure 2, le contexte c_1 est caractérisé par une plus grande probabilité que c_2 (voir la différence sur l'axe horizontal) mais par une moins grande asymétrie entre les probabilités conditionnelles de t_1 et t_2 . Autrement dit, il y a une moins grande asymétrie entre les valeurs t_1 et t_2 dans le contexte c_1 par rapport au contexte c_2 .

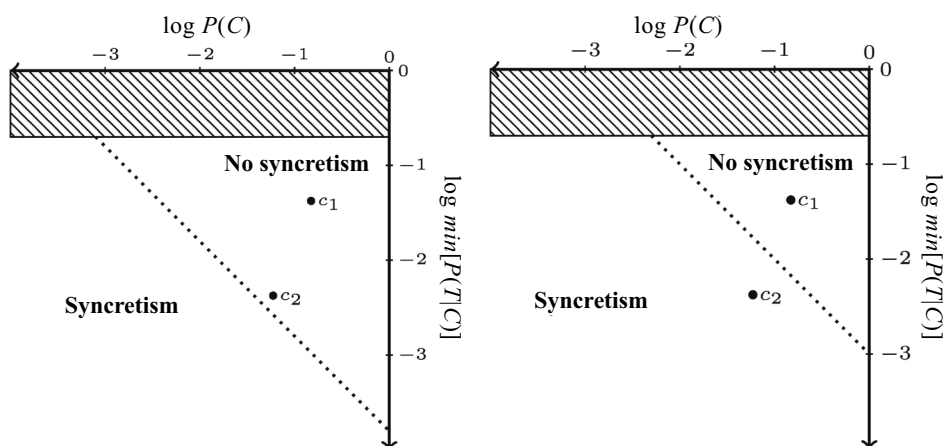


Figure 2. Dérivation de généralisations implicationnelles pour la typologie du syncrétisme dans le modèle à contraintes (graphique reproduit de Storme 2022a)

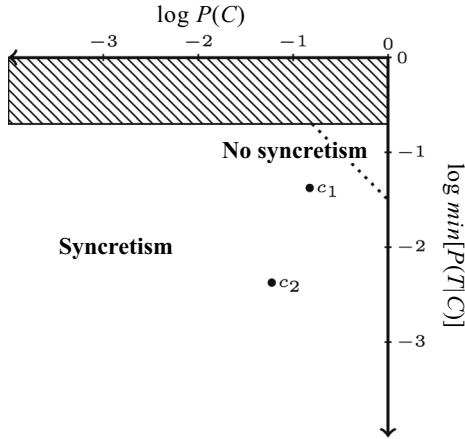


Figure 2. (suite)

Dans cette situation, on voit que la présence d'un synchrétisme dans le contexte c_1 implique de façon asymétrique la présence d'un synchrétisme dans le contexte c_2 . Graphiquement, il n'y a aucune droite avec une pente égale à -1 qui puisse passer à la fois au-dessus de c_1 (c_1 favorise le synchrétisme) et en dessous de c_2 (c_2 favorise l'absence de synchrétisme). Le modèle prédit donc que l'on ne devrait pas observer de langue où le trait grammatical distinguant t_1 et t_2 est exprimé de façon synchrétique dans c_1 uniquement. En revanche, les trois autres configurations sont prédites possibles par le modèle : absence totale de synchrétisme (en haut à gauche), synchrétisme dans c_2 seulement (en haut à droite) et présence d'un synchrétisme dans les deux contextes (en bas). Autrement dit, d'après les prédictions du modèle, le contexte c_2 est un contexte qui favorise nécessairement le synchrétisme par rapport au contexte c_1 .

6.1.3. Application du modèle aux universaux 44 et 45 de Greenberg

Le modèle proposé dans la partie précédente motive non seulement l'existence du synchrétisme mais prédit également que certains contextes devraient favoriser le synchrétisme à travers les langues. Le but de la présente partie est de montrer que le modèle prédit correctement les contextes favorisant le synchrétisme dans les universaux 44 et 45 de Greenberg.

Focalisons-nous d'abord sur l'universel 45 de Greenberg sur le rôle du nombre (singulier, pluriel) dans le synchrétisme de genre. Le modèle prédit que le pluriel devrait favoriser le synchrétisme de genre par rapport au singulier si le pluriel est moins fréquent que le singulier et/ou s'il y a une moins grande asymétrie entre masculin et féminin au pluriel qu'au singulier. Comme je le montre dans l'article du *Journal of Linguistics*, ces deux prédictions sont correctes : le pluriel est non seulement beaucoup moins fréquent que le singulier à travers les langues mais en plus il y a une plus grande asymétrie entre masculin et féminin au singulier qu'au pluriel. En effet, alors qu'au singulier, il est à peu près aussi probable qu'un pronom soit utilisé au masculin qu'au féminin, au pluriel l'asymétrie est très forte en raison de la présence de groupes mixtes. Dans les langues où le masculin est utilisé de façon générique, ces groupes mixtes sont désignés avec la forme masculine du pronom, ce qui fait monter la probabilité d'observer cette forme au pluriel, créant une asymétrie

forte entre masculin et féminin au pluriel. Cette asymétrie est confirmée par des données de corpus issues de l'espagnol.

Focalisons-nous ensuite sur l'universel 44 de Greenberg sur le rôle de la personne (première/deuxième personne, troisième personne) dans le syncrétisme de genre. Le modèle prédit que les personnes de discours (première et deuxième personnes) devraient favoriser le syncrétisme de genre par rapport à la troisième personne si les personnes de discours sont moins fréquentes que la personne trois et/ou s'il y a une moins grande asymétrie entre masculin et féminin dans les personnes du discours que dans la troisième personne. C'est surtout la deuxième prédiction qui est intéressante et reçoit un soutien empirique. En effet, pour les personnes du discours, le référent étant visible (il s'agit du locuteur ou de l'interlocuteur), son genre est connu et la probabilité d'erreur en cas de syncrétisme est ainsi minime. En revanche, pour la troisième personne, le genre n'est pas disponible dans le contexte de la même façon. Aussi y a-t-il une plus grande incertitude, que permet de lever le marquage explicite du genre.

L'article montre également que d'autres modèles proposés pour rendre compte de la typologie du syncrétisme font des prédictions problématiques. En particulier, le modèle diachronique proposé par Haspelmath dans son article de 2006 paru dans le *Journal of Linguistics*⁶¹ analyse le syncrétisme comme le résultat d'un apprentissage imparfait : les combinaisons grammaticales qui sont rares sont plus difficiles à apprendre et ont par conséquent tendance à disparaître d'une génération à l'autre.

Cette analyse explique pourquoi le syncrétisme de genre est favorisé au pluriel par rapport au singulier (universel 45) : la combinaison du pluriel avec la valeur de genre la moins fréquente dans ce contexte (comme le féminin pluriel en français) est rare, et ainsi aurait tendance à échapper aux apprenants, qui étendent dans ce contexte la valeur la plus fréquente (comme le masculin pluriel en français).

Cependant l'analyse diachronique de Haspelmath échoue à rendre compte de l'universel 44. En effet, les deux valeurs de genre (par exemple masculin et féminin) devraient être fréquentes dans la première et la deuxième personne du singulier. Ainsi on ne comprend pas pourquoi les formes genrées devraient avoir tendance à disparaître dans ce contexte. Dans l'approche fonctionnaliste proposée ici, le genre n'est pas particulièrement utile dans les personnes du discours dans la mesure où le genre peut être désambiguïsé facilement en utilisant le contexte extralinguistique.

Pour conclure, le modèle à contraintes permet de formuler l'analyse fonctionnaliste du syncrétisme de façon explicite et de faire des prédictions précises sur les contextes qui devraient favoriser le syncrétisme à travers les langues. Ces prédictions sont corroborées dans deux études de cas, contrairement à celles de l'analyse diachronique.

6.2. La variation dans une langue : la liaison en français

Nous nous tournons maintenant vers un phénomène de variation à l'intérieur d'une langue, pour montrer comment les grammaires à contraintes permettent également de dériver ce genre de variation. L'étude que je vais présenter dans cette partie s'appuie sur des données de type expérimental, et illustre ainsi l'intérêt de combiner expérimentation et modélisation. Cette étude est issue d'un article paru en 2024 dans le journal *Natural Language and Linguistic Theory*⁶².

61. Haspelmath (2006).

62. Storme (2024).

6.2.1. Variabilité de la réalisation de la liaison

L'étude de cas qui va nous occuper ici porte sur le phénomène de la liaison. En français, certains mots ont une variante avec consonne finale qui n'apparaît que devant des mots commençant par une voyelle. Cette consonne finale est appelée consonne de liaison. L'utilisation de la consonne de liaison peut être illustrée avec l'adjectif *grand*, qui finit par une voyelle nasale, sauf lorsqu'il est suivi par un mot à initiale vocalique, comme *ami*. Dans ce cas-là, le mot *grand* est alors prononcé avec une consonne de liaison [t], *un gran[t] ami*.

La liaison est un phénomène linguistique très complexe qui a fait l'objet d'un très grand nombre d'études. Parmi les problèmes que soulève la liaison se trouve sa réalisation variable comme consonne finale ou comme consonne initiale, un phénomène qui a été notamment décrit par le linguiste Pierre Encrevé dans les années 80⁶³. Par exemple, en présence d'une pause entre un mot à liaison et le mot suivant, la consonne de liaison peut être aussi bien prononcée à la fin du mot 1 (*un gran[t]... ami*) qu'au début du mot 2 (*un grand... [t]ami*). Cela contraste avec les consonnes finales et initiales, qui se prononcent toujours respectivement à la fin du mot 1 (*tren[t]... amis*) et au début du mot 2 (*un grand... [t]imide*).

Dans un article paru dans le journal *Natural Language and Linguistic Theory*, je propose que cette variabilité de la liaison s'explique par un effet d'uniformité paradigmatique. L'uniformité paradigmatique requiert qu'un mot soit prononcé de la même façon dans tous les contextes où il apparaît, et a ainsi le même rôle qu'une contrainte de fidélité (voir partie 5.3). En particulier, dans le cas de la liaison, l'uniformité paradigmatique requiert que le mot soit prononcé de la même façon que sa forme de citation. La forme de citation d'un mot correspond à sa forme lorsque le mot constitue un énoncé à lui seul, sans mot qui suive ou précède. Dans le cas des mots avec une forme de liaison, la forme de citation du mot ne contient pas de consonne de liaison. Ainsi *grand* est prononcé [gʁɑ̃], sans consonne [t] finale. L'uniformité paradigmatique pénalise donc l'utilisation de la forme de liaison avec une consonne finale. Devant un mot à initiale vocalique, comme *ami*, cette contrainte peut être satisfaite en évitant la forme de liaison (*grand ami* est ainsi prononcé [gʁɑ̃ami]), ou bien en repoussant la consonne sur le mot 2 (*grand... [t]ami*). Mais le mot 2 est lui-même soumis à la contrainte d'uniformité paradigmatique, ce qui va avoir tendance à repousser la consonne de liaison vers le mot 1 (*gran[t]... ami*). La variabilité de l'attachement prosodique de la liaison émerge donc d'un antagonisme entre les contraintes d'uniformité paradigmatique relatives au mot 1 et au mot 2.

Ce conflit impliquant deux types d'uniformité paradigmatique antagonistes peut être modélisé dans une grammaire à contraintes en exploitant la notion de contraintes de fidélité. En effet l'exigence d'uniformité paradigmatique peut se concevoir comme une contrainte de fidélité. Avant de passer à la modélisation, je vais présenter les résultats d'une étude expérimentale dont le but était de tester le rôle causal de l'uniformité paradigmatique dans la variabilité de la liaison.

6.2.2. Étude expérimentale

Si l'uniformité paradigmatique avec la forme de citation est le mécanisme causal qui explique la variabilité de la liaison, alors on ne devrait pas observer de variabilité pour les mots qui ont une forme de liaison identique à leur forme de citation. Cette prédiction

63. Encrevé (1988).

peut être testée avec les mots à liaison qui utilisent une forme de féminin en contexte de liaison. C'est le cas par exemple de l'adjectif *beau* qui, en contexte de liaison, utilise une forme identique à sa forme féminine *belle*, comme dans *un bel ami*⁶⁴. Ces liaisons diffèrent du type plus classique de liaison, comme *grand*, parce qu'elles s'appuient sur une autre forme du paradigme de l'adjectif, en l'occurrence la forme féminine.

Dans mon article paru dans le journal *Natural Language and Linguistic Theory*, j'ai testé cette hypothèse expérimentalement en présentant à des participants francophones deux prononciations possibles pour les deux types de mots à liaison (*grand* vs *beau/bel*) : une prononciation où la consonne de liaison est attachée au mot 1, comme en (21a) et en (22a), et une prononciation où la consonne de liaison est attachée au mot 2, comme en (21b) et en (22b).

- 21 a. *un gran[t] euh ami*
 b. *un grand euh [t]ami*
 22 a. *un be[l] euh ami*
 b. *un be euh [l]ami*

Les participants devaient indiquer la prononciation vers laquelle allait leur préférence. Les résultats vont dans le sens des prédictions. Pour les liaisons du type *grand*, il y a une variabilité dans les réponses : les deux prononciations (21a) et (21b) sont jugées acceptables. En revanche, pour les liaisons du type *beau/bel*, seule la prononciation (22a) est jugée acceptable. Autrement dit, la consonne de liaison n'est acceptée dans les deux positions par les participants qu'à condition que la forme de liaison diffère de la forme de citation du mot.

6.2.3. Modélisation

Pour s'assurer que la théorie de l'uniformité paradigmaticque peut bien rendre compte de la variabilité de la liaison, les résultats de l'expérience ont été modélisés en utilisant une grammaire à contraintes probabiliste. Je présente ici une version simplifiée du modèle utilisé dans Storme (2024), avec seulement deux contraintes.

La première contrainte, appelée *Ancre-à-droite*, exige que, pour un mot donné, il y ait correspondance entre le segment final de sa forme contextuelle et le segment final de sa forme de citation. Cette contrainte pénalise ainsi l'ajout ou l'élision d'une consonne finale dans la forme contextuelle par rapport à la forme de citation. Cette contrainte va pénaliser par exemple la forme de liaison *gran[t]* avec un [t] final (dans la mesure où ce [t] est absent de la forme de citation de *grand*). La seconde contrainte, appelée *Ancre-à-gauche*, a le même effet, mais sur le début du mot : cette contrainte pénalise l'ajout ou l'élision d'un segment au début de la forme contextuelle d'un mot par rapport à la forme de citation. Elle va pénaliser par exemple la forme [t]*ami* avec un [t] ajouté au début du mot *ami*.

Le Tableau 4 montre comment ces deux contraintes évaluent les deux prononciations proposées aux participants pour quatre types de mots : des mots avec une consonne finale stable (par exemple *magnifique* avec un [k] final), des mots avec une consonne de liaison du type *beau/bel* (par exemple *bel* avec un [l] de liaison, présent dans la forme de citation du féminin), des mots avec une consonne de liaison du type *grand* (par exemple *gran[t]* avec un [t] de liaison, absent de la forme de citation) et des mots avec une consonne initiale (par exemple *sourire* avec un [s] initial).

64. Il s'agit ici d'identité phonologique et non graphique (*bel* et *belle* [bɛl] sont identiques phonologiquement mais pas graphiquement).

Tableau 4. Analyse de la variabilité de la liaison dans le modèle à contraintes

<i>magnifique hôtel</i>	Ancre-à-droite w=1.36	Ancre-à-gauche w=1.29	Harmonie	Probabilité prédite
<i>magnifi</i> [k]# <i>hôtel</i>	0	0	0	0,07
<i>magnifi</i> #[k] <i>hôtel</i>	1	1	2,66	0,93

a. Consonne finale

<i>bel appartement</i>	Ancre-à-droite w=1.36	Ancre-à-gauche w=1.29	Harmonie	Probabilité prédite
<i>be</i> [l]# <i>appartement</i>	0	0	0	0,07
<i>be</i> #[l] <i>appartement</i>	1	1	2,66	0,93

b. Consonne de liaison du type *beau/bel*

<i>grand hommage</i>	Ancre-à-droite w=1.36	Ancre-à-gauche w=1.29	Harmonie	Probabilité prédite
<i>gran</i> [t]# <i>hommage</i>	1	0	1,36	0,52
<i>gran</i> #[t] <i>hommage</i>	0	1	1,29	0,48

c. Consonne de liaison du type *grand*

<i>joli sourire</i>	Ancre-à-droite w=1.36	Ancre-à-gauche w=1.29	Harmonie	Probabilité prédite
<i>joli</i> [s]# <i>sourire</i>	1	1	2,66	0,93
<i>joli</i> #[s] <i>sourire</i>	0	0	0	0,07

d. Consonne initiale

Dans le Tableau 4a, la forme avec la consonne prononcée à la fin du mot 1 (*magnifi*[k]) ne viole aucune de ces contraintes : le [k] final du mot 1 se trouve bien prononcé à la fin du mot 1, en accord avec la contrainte Ancre-à-droite, et le [o] initial du mot 2 se trouve bien prononcé au début du mot 2, en accord avec la contrainte Ancre-à-gauche. En revanche, la forme avec le [k] prononcé au début du mot 2 (*magnifi*#[k] *hôtel*) viole les deux contraintes : le [k] qui se trouve à la fin de la forme de citation de *magnifique* est absent de la forme contextuelle de ce mot, en violation de la contrainte Ancre-à-droite, et se trouve présent au début de la forme contextuelle du mot 2 (*hôtel*), en violation de la contrainte Ancre-à-gauche qui exige qu'aucun segment ne soit ajouté au début d'un mot. L'analyse prédit donc qu'il devrait y avoir une nette préférence pour la première prononciation. D'après l'analyse, les mots à liaison du type *beau/bel* devraient favoriser de la même façon un attachement final de la consonne : dans ces mots, la forme féminine est utilisée de façon supplétive en contexte de liaison, et la forme de citation correspondante présente une consonne finale (le [l] pour l'adjectif *beau/bel*).

Le Tableau 4c montre la façon dont ces contraintes évaluent les deux prononciations pour les mots à liaison du type *grand*. Dans ce cas, chacune des deux prononciations viole une des deux contraintes : la prononciation avec la consonne de liaison à la fin du mot 1 viole la contrainte Ancre-à-droite (une consonne absente de la forme de citation de *grand* est ajoutée à la fin de la forme contextuelle du mot) tandis que la prononciation avec la consonne de liaison au début du mot 2 viole la contrainte Ancre-à-gauche (une consonne absente de la forme de citation de *hommage* est ajoutée au début la forme contextuelle du mot). L'analyse ne prédit pas une asymétrie aussi forte entre les deux

prononciations que dans le cas des consonnes finales stables (*magnifique*, Tableau 4a) et les consonnes de liaison du type *beau/bel* (Tableau 4b).

Le Tableau 4d montre la façon dont les contraintes évaluent les deux prononciations pour les mots commençant par une consonne (par exemple *sourire*). Ici, comme dans le cas des consonnes finales stables et des consonnes de liaison du type *beau/bel*, une préférence nette est prédite pour l'une des deux prononciations, à savoir la prononciation avec la consonne au début du mot 2. En effet prononcer le [s] de *sourire* à la fin du mot 1 viole les deux contraintes : Ancre-à-droite, parce qu'un [s] est ajouté à la fin du mot 1, et Ancre-à-gauche, parce qu'un [s] est enlevé au début du mot 2.

Une fois que cette première étape de l'analyse est faite, on passe ensuite à la modélisation proprement statistique. Dans cette étape, j'ai utilisé le logiciel statistique *R*⁶⁵ et le package *rjags*⁶⁶ pour trouver les poids des deux contraintes qui permettent de rendre compte des fréquences observées dans l'expérience décrite dans la partie 6.2.2 pour les deux prononciations de chaque type de consonne⁶⁷. Les poids qui ont été inférés sont montrés dans les colonnes correspondantes pour les deux contraintes : le poids d'Ancre-à-droite est égal à 1,36 et celui d'Ancre-à-gauche est égal à 1,29. Les fréquences prédites par l'analyse apparaissent dans la colonne de droite du Tableau 4. Les prédictions correspondent aux observations : l'analyse dérive de la variabilité seulement pour la liaison de type *grand*. Par ailleurs les fréquences prédites sont très proches des fréquences observées dans l'expérience, ce qui signifie que le modèle rend bien compte des données.

7. CONCLUSION

Le but principal de cet article était méthodologique : il s'agissait de démontrer l'intérêt des méthodes d'expérimentation et de modélisation pour la linguistique. Dans cette conclusion, je reviens sur ces aspects méthodologiques, mais aussi sur des enseignements que l'on a pu tirer de l'utilisation de ces méthodes pour des questions plus théoriques, à savoir les motivations de l'écriture inclusive et la place du fonctionnalisme en linguistique.

On a vu que l'expérimentation est particulièrement adaptée pour comprendre les mécanismes causaux à l'œuvre dans les systèmes linguistiques. Elle permet en effet d'isoler l'effet de différentes variables explicatives dans un phénomène. Par exemple, à propos de la question de l'écriture inclusive, nous avons vu comment un protocole expérimental qui contrôle l'effet des stéréotypes sur les inférences de genre permet d'établir que le genre grammatical contribue indépendamment à ces inférences. Dans l'étude qui se focalisait sur des noms avec un genre fixe (*une personne, un individu*), on a vu que l'effet du genre grammatical sur les inférences de genre subsistait, ce qui signifie que le genre grammatical contribue aux inférences de genre indépendamment de toute compétition morphologique. Sur la question de la liaison, nous avons vu comment le protocole expérimental comparant deux types de liaison a permis de montrer que la variabilité de la liaison s'explique par la notion d'uniformité paradigmatique : la consonne de liaison n'est pas variable intrinsèquement mais seulement si elle est absente de la forme de citation correspondante. Ce résultat appuie une des hypothèses clés de la théorie fonctionnaliste du langage, à savoir que les locuteurs cherchent

65. R Core Team (2021).

66. Plummer (2016).

67. Les données sont disponibles en ligne dans Storme (2022b).

à minimiser la dissimilarité entre les variantes d'un même mot, ici la variante de liaison et la forme de citation du mot.

Quant à la modélisation, son intérêt réside dans la possibilité qu'elle offre de rendre les théories non seulement complètement explicites mais aussi robustes au bruit. Nous avons illustré en particulier l'intérêt des grammaires à contraintes pour modéliser des théories d'inspiration fonctionnaliste. Les grammaires à contraintes permettent d'intégrer les facteurs d'explication fonctionnalistes (minimiser l'effort, maximiser la similarité entre les variantes d'un même mot) comme des primitives de la grammaire et ont un intérêt pour l'étude de la variation à la fois à travers et au sein des langues. D'un point de vue typologique, cette approche nous a permis de dériver des contraintes observées par Greenberg sur la façon dont les langues expriment des distinctions morphologiques. En particulier, le modèle basé sur des contraintes prédit que certains contextes grammaticaux devraient favoriser l'expression ambiguë des distinctions morphologiques. Plus spécifiquement, il prédit que les contextes qui permettent de mieux désambiguïser les distinctions morphologiques sous-jacentes sont plus susceptibles de ne pas marquer ces distinctions ouvertement. On a vu que des approches purement diachroniques et non téléologiques, comme celle proposée par Haspelmath, ne permettent pas de rendre compte de tous les cas. Enfin, sur la liaison, nous avons vu que la modélisation permet d'analyser des résultats expérimentaux sur la variabilité d'attachement de la consonne de liaison directement en termes de contraintes fonctionnalistes.

RÉFÉRENCES BIBLIOGRAPHIQUES

- ACADÉMIE FRANÇAISE, 2017. « Déclaration de l'Académie française sur l'écriture dite inclusive », en ligne : <https://www.academie-francaise.fr/actualites/declaration-de-lacademie-francaise-sur-lecriture-dite-inclusive> (consulté le 11 novembre 2023).
- BAYLES, A., KAPLAN, A. & KAPLAN, B., 2016. « Inter- and intra-speaker variation in French schwa », *Glossa: A journal of general linguistics* 1, en ligne : <https://doi.org/10.5334/gjgl.54> (consulté le 28 février 2024).
- BOERSMA, P. & HAYES, B. (2001). « Empirical tests of the gradual learning algorithm », *Linguistic Inquiry* 32, 45-46.
- BRYSSAERT, M., MANDERA, P. & KEULEERS, E., 2018. « The word frequency effect in word processing: An updated review », *Current Directions in Psychological Science* 27, 45-50.
- CHOMSKY, N., 1968. *Language and Mind*, New York, Harper & Row.
- CHOMSKY, N. & HALLE, M., 1968. *The sound pattern of English*, MIT Press, Cambridge.
- CORBETT, G., 2013a. « Number of Genders », in M. S. Dryer et M. Haspelmath (éd.), *The World Atlas of Language Structures Online*, Leipzig, Max Planck Institute for Evolutionary Anthropology, en ligne : <http://wals.info/chapter/30> (consulté le 11 novembre 2023).
- CORBETT, G., 2013b. « Sex-based and non-sex-based gender systems », in M. S. Dryer et M. Haspelmath (éd.), *The World Atlas of Language Structures Online*, Leipzig, Max Planck Institute for Evolutionary Anthropology, en ligne : <http://wals.info/chapter/31> (consulté le 11 novembre 2023).

- CORNULIER, B. DE (1975). «Le droit d'e : e et la syllabacité», *Cahiers de linguistique, d'orientalisme et de slavistique* 5-6, 101-117.
- ENCREVÉ, P., 1988. *La liaison avec et sans enchaînement : phonologie tridimensionnelle et usages du français*, Paris, Seuil.
- FEDZCHKINA, M., JAEGER, F. T. & NEWPORT, E. L., 2012. « Language learners restructure their input to facilitate efficient communication », *Proceedings of the National Academy of Sciences* 109, 17897-17902.
- FLEMMING, E., 2021. « Comparing MaxEnt and Noisy Harmonic Grammar », *Glossa : A journal of general linguistics* 6, en ligne : <https://doi.org/10.16995/glossa.5775> (consulté le 17 novembre 2023).
- FRANK, M. C., BRAGINSKY, M., CACHIA, J., COLES, N., HARDWICKE, T. E., HAWKINS, R. D., MATHUR, M. B. & WILLIAMS, R., 2023. *Experimentology: An Open Science Approach to Experimental Psychology Methods*, en ligne : <https://experimentology.io/> (consulté le 6 novembre 2023).
- GABRIEL, U., GYGAX, P., SARRASIN, O., GARNHAM, A. & OAKHILL, J., 2008. « Au pairs are rarely male: Norms on the gender perception of role names across English, French, and German », *Behavior Research methods* 40, 206-212.
- GIBSON, E., FUTRELL, R., PIANTADOSI, S. P., DAUTRICHE, I., MAHOWALD, K., BERGEN, L. & LEVY, R., 2019. « How efficiency shapes human language », *Trends in Cognitive Science* 23, 389-407.
- GREENBERG, J. H. 1963. « Some universals of grammar with particular reference to the order of meaningful elements », in J. H. Greenberg (éd.), *Universals of Human Language*, Cambridge, MIT Press, 73-113.
- GYGAX, P. & GESTO, N., 2007. « Féminisation et lourdeur de texte », *L'Année psychologique* 107, 239-255.
- GYGAX, P., GABRIEL, U., SARRASIN, O., OAKHILL, J. & GARNHAM, A., 2008. « Generically intended, but specifically interpreted: When beauticians, musicians, and mechanics are all men », *Language and Cognitive Processes* 23, 464-485.
- GYGAX, P. & GABRIEL, U., 2008. « Can a group of musicians be composed of women? Generic interpretation of French masculine role names in the absence and presence of feminine forms », *Swiss Journal of Psychology* 67, 143-151.
- GYGAX, P., ZUFFEREY, S. & GABRIEL, U., 2021. *Le cerveau pense-t-il au masculin ?* Paris, Le Robert.
- GIBSON, E., 1998. «Linguistic complexity: Locality of syntactic dependencies», *Cognition* 68, 1-76.
- HADDAD, R., 2017. «Ce que peut apporter l'écriture inclusive à l'égalité femmes-hommes», *Huffington Post*, en ligne : https://www.huffingtonpost.fr/politique/article/ce-que-peut-apporter-l-ecriture-inclusive-a-l-egalite-femmes-hommes_92564.html (consulté le 29 novembre 2023).
- HAGÈGE, C., 2017. «Ce n'est pas la langue qui est sexiste mais les comportements sociaux», *Le Monde*, en ligne : https://www.lemonde.fr/idees/article/2017/12/26/claude-hagege-ce-n-est-pas-la-langue-qui-est-sexiste_5234408_3232.html (consulté le 29 novembre 2023).
- HASPELMATH, M., 2006. « Against markedness (and what to replace it with) », *Journal of Linguistics* 42, 25-70.

- HAUT CONSEIL À L'ÉGALITÉ ENTRE LES FEMMES ET LES HOMMES, 2022. *Guide pratique pour une communication publique sans stéréotypes de sexe*, en ligne : https://www.haut-conseil-egalite.gouv.fr/IMG/pdf/guide_egacom_sans_stereotypes-2022-versionpublique-min.pdf (consulté le 10 novembre 2023).
- HAYES, B. (2017). « Varieties of Noisy Harmonic Grammar », *Proceedings of the 2016 Annual Meeting on Phonology*, Washington, DC, Linguistic Society of America.
- HAYES, B. & WILSON, C., 2008. « A Maximum Entropy Model of phonotactics and phonotactic learning », *Linguistic Inquiry* 39, 379-440.
- KAGER, R., 1999. *Optimality Theory*, Cambridge, Cambridge University Press.
- LAYCOCK, D. C. 2003. *A dictionary of Buin, a language of Bougainville*, Pacific Linguistics, Canberra.
- LIGNOS, C. & YANG, C., 2016. « Morphology and language acquisition », in A. Hippisley & G. Stump (éd.), *Cambridge Handbook of Morphology*, Cambridge, Cambridge University Press, 765-791.
- LILJENCRAITS, J. & LINDBLOM, B., 1972. « Numerical simulation of vowel quality systems: the role of perceptual contrast », *Language* 48, 839-862.
- MAGRI, G., 2013. « Harmonic Grammar has no computational advantages over Optimality Theory: Towards a new toolkit for computational Optimality Theory », *Linguistic Inquiry* 44, 569-609.
- MAHOWALD, K., DIACHEK, E., GIBSON, E., FEDORENKO, E. & FUTRELL, R., 2023. « Grammatical cues to subjecthood are redundant in a majority of simple clauses across languages », *Cognition* 241, 105543.
- MARTINET, A., 1980 [1960¹]. *Éléments de linguistique générale*, Paris, Colin.
- PERFORS, A., TENENBAUM, J. B. & REGIER, T., 2011. « The learnability of abstract syntactic principles », *Cognition* 118, 306-338.
- PIANTADOSI, S. T., TILY, H. & GIBSON, E., 2012. « The communicative function of ambiguity in language », *Cognition* 122, 280-291.
- PIERREHUMBERT, J., 2001. « Exemplar dynamics: Word frequency, lenition and contrast », in J. Bybee & P. Hopper (éd.), *Frequency and the Emergence of Linguistic Structure*, Amsterdam, John Benjamins, 137-157.
- PLANK, F. & SCHELLINGER, W., 1997. « The uneven distribution of genders over numbers: Greenberg nos. 37 and 45 », *Linguistic Typology* 1, 53-101.
- PLUMMER, M., 2016. *rjags: Bayesian Graphical Models using MCMC*, R package version 4-5, en ligne : <https://CRAN.R-project.org/package=rjags> (consulté le 27 novembre 2023).
- PREWITT-FREILINO, J. L., CASWELL, T. A. & LAAKSO, E., 2012. « The gendering of language: A comparison of gender equality in countries with gendered, natural gender, and genderless languages », *Sex Roles* 66, 268-281.
- PRINCE, A. & SMOLENSKY, P., 1993. *Optimality Theory: Constraint interaction in generative grammar*, New Brunswick, NJ, Rutgers Center for Cognitive Science, Rutgers University.
- R CORE TEAM, 2020. *R : A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, en ligne : <https://www.R-project.org/> (consulté le 27 novembre 2023).

- RACINE, I., DURAND, J. & ANDREASSEN, H., 2016. «PFC, codages et représentations : la question du schwa», *Corpus* 15, en ligne : <https://doi.org/10.4000/corpus.3014> (consulté le 28 février 2024).
- ROBERTS, S. G., WINTERS, J. & CHEN, K., 2015. « Future tense and economic decisions: Controlling for cultural evolution », *PLoS ONE* 10, e0132145.
- SAPIR, E., 1921. *Language: An Introduction to the Study of Speech*, New York, Harcourt, Brace & Company.
- SATO, S., GYGAX, P. & GABRIEL, U., 2013. « Gender inferences: Grammatical features and their impact on the representation of gender in bilinguals », *Bilingualism: Language and Cognition* 16, 792-807.
- SIEWIERSKA, A., 2013. « Gender distinctions in independent personal pronouns », in M. S. Dryer & M. Haspelmath (éd.), *The World Atlas of Language Structures Online*, Leipzig, Max Planck Institute for Evolutionary Anthropology, en ligne : <https://wals.info/chapter/44> (consulté le 28 février 2024).
- STORME, B., 2021. « Not only size matters: limits to the Law of Three Consonants in French phonology », *Glossa: a journal of general linguistics* 6, 125, en ligne : <https://doi.org/10.16995/glossa.5882> (consulté le 28 février 2024).
- STORME, B., 2022a. « Implicational generalizations in morphological syncretism: The role of communicative biases », *Journal of Linguistics* 58, 381-421.
- STORME, B., 2022b. « Paradigm uniformity effects on French liaison » [Ensemble de données], *Zenodo*, en ligne : <https://doi.org/10.5281/zenodo.7971975> (consulté le 26 novembre 2023).
- STORME, B., 2024. « Paradigm uniformity effects on French liaison », *Natural Language and Linguistic Theory*, en ligne : <https://doi.org/10.1007/s11049-023-09596-z> (consulté le 27 février 2024).
- STORME, B. & DELALOYE SAILLEN, L., 2024. « Effects of grammatical gender on gender inferences: Evidence from French hybrid nouns », *Linguistics Vanguard*, en ligne : <https://doi.org/10.1515/lingvan-2022-0064> (consulté le 18 mars 2024).
- XIAO, H., STRICKLAND, B. & PEPERKAMP, S., 2023. « How fair is gender-fair language? Insights from gender ratio estimations in French », *Journal of Language and Social Psychology* 42, 82-106.
- ZIPF, G. K., 1949. *Human Behavior and the Principle of Least Effort*, Cambridge, Addison-Wesley.
- ZIPF, G. K., 1965 [1935¹]. *The Psycho-Biology of Language*, Cambridge, The MIT Press.