

Mais où sont donc les femmes ? À la recherche de formes doubles dans un corpus de textes administratifs suisses

Eva Schaeffer-Lacroix¹

¹*CeLiSo* Paris – France

Abstract

Through the adoption of the language law 2007, the Swiss government reaffirms its commitment to gender-sensitive language use in its official texts of administrative nature. This law corroborates and specifies article 8 about "Equality before the law" (Federal Constitution of the Swiss Confederation), complemented in 1981 by the aspect "gender equality". The quantitative research needs of a project about language change in Swiss administration discourse motivated the creation of the multilingual three-part corpus *Bundesblatt/Feuille fédérale/Foglio federale*. This corpus contains documents published by the Swiss Federal Council between 1849 and 2014, and it can be searched with the help of the corpus management system *CQPweb* (Evert, 2010). The project focusses on common personal nouns. Our paper presents a study about non-shortened paired nouns referring to people, such as *Verkäuferinnen und Verkäufer* [saleswomen and salesmen] in the German *Bundesblatt* corpus. This *pattern* is considered as one of the suited linguistic means to display the co-presence of women and men in German texts. The retrieval of such paired nouns in the corpus is hindered by the graphic identity between certain personal nouns and proper nouns or between personal nouns and common nouns which are not referring to people. This implies that the employed search method should help to eliminate the highest possible proportion of results which, for semantic reasons, are considered as not relevant. Amongst other discoveries, we detected a significant distributional difference between both possible noun word orders. Our results support the idea that the declared intention of the Swiss government to guarantee the visibility of women in its official administrative texts matches well with some aspects of gender-friendly language use attested in the *Bundesblatt*.

Résumé

En adoptant en 2007 la *Loi sur les langues*, le gouvernement suisse s'engage à favoriser la formulation non sexiste dans les textes administratifs qu'il fait publier. Cette loi est une sorte de piqûre de rappel du chantier engagé depuis 1981, année de l'inscription de l'égalité des hommes et des femmes dans l'article 8 de la constitution fédérale de la Confédération suisse. En lien avec un projet portant sur les caractéristiques diachroniques et synchroniques du langage administratif suisse, un corpus multilingue a été créé afin d'appuyer les recherches qualitatives de l'équipe par des recherches quantitatives. Il contient des documents publiés par la Chancellerie suisse entre 1849 et 2014 (*Bundesblatt/Feuille fédérale/Foglio federale*). Ses trois parties peuvent être interrogées séparément à l'aide du système de gestion de corpus *CQPweb* (Evert, 2010). Le projet porte en particulier sur les noms communs de personnes. Notre communication s'intéresse à la présence de formes doubles intégrales, telles que *Arbeitnehmerinnen und Arbeitnehmer* [salariées et salariés] dans le *Bundesblatt*, le volet allemand du corpus. Ce *pattern* est considéré comme l'un des moyens linguistiques les plus adéquats pour rendre visible la co-présence des femmes et des hommes dans des textes rédigés en langue allemande. L'identification de ces formes dans le corpus se heurte au fait que leurs constituants peuvent être des homonymes de noms propres ou de noms communs désignant autre chose que des personnes. Il s'agit donc d'élaborer une méthode de recherche permettant d'écarter le mieux possible les résultats non désirés, principalement du point de vue sémantique. Nous avons, entre autres, découvert une distribution inégale significative entre les deux ordres syntaxiques possibles. Nous interprétons les faits observés comme des indices suggérant une adéquation entre la volonté de l'administration suisse de renforcer la visibilité des femmes et l'usage linguistique attesté dans les textes qu'elle diffuse via le *Bundesblatt*.

Key words: discours administratif ; suisse allemand ; féminisation ; corpus ; formes doubles intégrales

1. Introduction

1.1. Arrière-plan de l'étude

En adoptant la *Loi fédérale sur les langues nationales et la compréhension entre les communautés linguistiques* (Chancellerie fédérale, 2007), le gouvernement suisse s'engage explicitement à favoriser la formulation non sexiste dans les textes administratifs qu'il fait publier. Cette loi est une sorte de piqûre de rappel du chantier engagé depuis 1981, année de l'inscription de l'égalité des hommes et des femmes dans la constitution suisse (Commission fédérale pour les questions féminines CFQF, 2009 : volet 3.1.). En effet, depuis les années 70, en ce qui concerne les femmes, le droit suisse évolue de façon constante. Ce cadre légal crée de bonnes conditions pour l'amélioration de leur situation familiale et professionnelle. Des guides ont été publiés à des périodes variées, pour trois des quatre langues officielles suisses, afin de soutenir la mise en pratique des textes de loi allant dans ce sens. La version la plus récente d'un tel guide en langue allemande a paru en 2009 (Schweizerische Bundeskanzlei, 2009). Nous nous y référerons sous la dénomination de *Leitfaden*.

Notre étude s'inscrit dans le projet *Politique linguistique et usage de la langue dans la Suisse plurilingue: noms communs de personne dans le langage administratif*¹ portant sur les noms communs de personnes dans le langage administratif suisse en diachronie et synchronie. Pour les besoins de ce projet, un corpus trilingue a été créé afin d'appuyer les recherches qualitatives de l'équipe par des recherches quantitatives. Il contient des documents publiés par la Chancellerie fédérale suisse entre 1849 et 2014 (*Bundesblatt/Feuille fédérale/Foglio federale*), et il peut être interrogé à l'aide du système de gestion de corpus *CQPweb* (Evert, 2010). Nos observations sont faites à partir du volet allemand de ce corpus, nommé *Corpus Bundesblatt* (Elmiger, 2015b). Il contient plus de 203 millions de tokens.

1.2. Question de travail

Notre communication s'intéresse à la présence de formes doubles intégrales dans le *Corpus Bundesblatt*, telles que *Arbeitnehmerinnen und Arbeitnehmer* [salariées et salariés]. Ce *pattern* est considéré comme l'un des moyens linguistiques les plus adéquats pour rendre visible la co-présence des femmes et des hommes dans des textes rédigés en langue allemande (Elmiger, 2015a ; Schweizerische Bundeskanzlei, 2009 : 19). Afin de voir si ce *pattern* fait partie des stratégies utilisées dans des textes administratifs depuis que le thème de la féminisation de la langue est encadré par la législation suisse, nous souhaitons examiner dans quelle mesure et sous quelles formes il est attesté dans les différentes périodes couvertes par le *Corpus Bundesblatt*.

2. Système d'exploitation de corpus *CQPweb*

Les données du corpus *Bundesblatt/Feuille fédérale/Foglio federale* ont été préparées en vue de leur exploration à l'aide de *CQPweb* (Hardie, 2012). Les fonctionnalités principales que nous avons utilisées sont les suivantes : la création de lignes de concordance, le classement alphabétique des occurrences du mot recherché ou de son environnement proche, la création de listes de fréquence et de tables de distribution (sous forme de graphique ou de tableau), l'identification de collocations et la fonction appelée *thinning* [élagage] permettant la restriction manuelle du nombre des occurrences que l'on souhaite voir s'afficher.

¹ [Projet FNS \(Fond National Suisse\) 143585](#). Université de Genève (2013 à 2016).

3. Caractéristiques du corpus

3.1. Données

Les trois volets du corpus du projet contiennent l'intégralité des feuilles fédérales publiées depuis 1849 en allemand et français et depuis 1971 en italien (Chancellerie fédérale, 1849-2014). Elles paraissent à un rythme hebdomadaire. On y trouve, entre autres, les rapports du Conseil fédéral aux Chambres, les arrêtés et les lois qui ont été votés par les Chambres, les décisions de la Chancellerie fédérale en matière d'initiatives populaires ou de référendums et les notifications des unités de l'administration fédérale et des tribunaux (Chancellerie fédérale, 2015).

| Nom | <i>Bundesblatt</i> | <i>Feuille fédérale</i> | <i>Foglio federale</i> |
|------------------|--------------------|-------------------------|------------------------|
| Nombre de tokens | 203 585 806 | 239 125 036 | 85 223 085 |
| Nombre de textes | 65 671 | 62 366 | 21 400 |
| Période couverte | 1849-2014 | 1849-2014 | 1971-2014 |

Tableau 1 - Les volets du corpus multilingue.

Le volet allemand qui nous intéresse pour notre étude contient 203 585 803 tokens, issus de 65 671 textes. À partir de l'interface de requête de *CQPweb*, il est possible de retourner aux textes d'origine qui sont disponibles en format pdf.

3.2. Types d'annotation

Avant d'être déposées sur *CQPweb*, les données ont été annotées à l'aide des outils *TreeTagger* (Schiller, Thielen, Teufel & Stöckert, 1995/1999) et *RfTagger* (Schmid & Laws, 2008). Une documentation plus détaillée peut être consultée dans Elmiger (2015b).

3.3. Ensembles interrogeables

Nous avons créé deux types de sous-corpus. Le premier type correspond aux années de publication des données, le deuxième à d'autres périodes de temps : il est possible de restreindre les requêtes aux données du premier siècle, à celles du premier ou du deuxième demi-siècle, à celles d'une des décennies à partir de 1951, ainsi qu'à celles correspondant à la décennie la plus récente ou à celles des quatre années les plus récentes (2011-2014).

4. Description du *pattern*

Les formes doubles intégrales se composent essentiellement de deux noms de personne, l'un étant au féminin et l'autre au masculin. Tout de suite après le premier des deux se trouve soit une conjonction de coordination, comme *und* [et], *oder* [ou] ou *sowie* [ainsi que], soit une virgule ou une barre oblique. Ce *pattern* existe au singulier et au pluriel. Le nom féminin de la forme double se termine, à de rares exceptions près², par ".+in" au singulier et par ".+innen" au pluriel. Pour le nom masculin singulier, la langue allemande offre toute une gamme de terminaisons, comme en témoigne cette liste : "and, ant, ast, e, el, en, ent, er, eur, ich, ier, ig, ist, ling, or, ör". Les terminaisons possibles pour le nom masculin au pluriel sont "e, en, er, l, n, s, Ø", réductibles à "e, er, l, n, s". Elles varient en fonction de la forme de base des noms et

² Exemples : *Cousin - Cousine* [cousin - cousine] ; *Ehemann - Ehefrau* [époux - épouse].

en fonction des flexions des cas. Exemple : "*die Wählerinnen und Wähler*" [les électrices et électeurs] *versus* "*den Wählerinnen und Wählern*" [les / aux électrices et électeurs].

5. Méthodologie

Les méthodes utilisées doivent aider à identifier dans le *Corpus Bundesblatt* les formes doubles intégrales, désignant des personnes, à l'exclusion des noms propres. En raison du très grand nombre d'occurrences auquel on peut s'attendre dans certaines parties du corpus, ces méthodes doivent permettre d'évaluer avec une précision suffisante la qualité des résultats de requête, tout en captant un maximum d'occurrences pertinentes. Nous avons donc opté pour un procédé plus ou moins quantitatif ou qualitatif, selon les cas de figure.

5.1. Défis à surmonter

L'identification des formes doubles intégrales dans le *Corpus Bundesblatt* se heurte au fait que leurs constituants peuvent être des homonymes de noms propres (cf. le nom de famille *Schweizer*, forme qui peut aussi vouloir dire "Suisse" ou "suisse") ou de noms communs désignant autre chose que des personnes (cf. *Leiter* qui peut vouloir dire "responsable, chef, organisateur", mais aussi "échelle"). De plus, nous nous intéressons uniquement aux emplois généralisants, donc aux situations dans lesquelles on se réfère aussi bien à des femmes qu'à des hommes, sans pointer l'un ou l'autre sexe en particulier. Un tel pointage, donc un emploi particularisant, peut être constaté dans l'exemple suivant de 1968 : "*die erwachsenen Arbeitnehmerinnen und die jugendlichen Arbeitnehmer*" [les salariées adultes et les salariés adolescents]. Dans des situations de ce type, il faut s'aider du contexte plus ou moins large afin de se rendre compte si une forme double est utilisée de façon généralisante ou particularisante. Le même problème se pose pour les *patterns* contenant une virgule. À ces défis d'ordre sémantique s'ajoute un défi d'ordre morphologique : si les noms féminins sont facilement identifiables grâce à leur terminaison ".+in" au singulier et ".+innen" au pluriel, il n'en est pas de même pour les noms masculins. Surtout en ce qui concerne le singulier, on aurait toute une gamme de terminaisons à inclure dans l'étude. La prise en compte de tous les cas de figure dépasse toutefois le cadre de notre exposé³.

5.2. Délimitation de l'objet d'étude

Les arguments évoqués en section 5.1. nous incitent à limiter notre étude aux formes doubles au pluriel, plus fréquentes que celles au singulier, et plus particulièrement à celles contenant une conjonction de coordination ou une barre oblique. Nous ne traiterons pas ici les formes contenant une virgule qui sont assez rares et souvent non pertinentes. Nous avons décidé d'observer des situations dans lesquelles la conjonction est suivie de zéro, un ou deux éléments avant l'apparition du deuxième nom de personne. Cette distance permet d'identifier les occurrences comportant des expansions du nom, par exemple des adjectifs ou des adverbes intensificateurs, comme "*sogar*" [même] ou "*aber auch*" [mais aussi]. Ces derniers sont toutefois à considérer avec prudence ; ils peuvent signaler un emploi particularisant des éléments de la forme double intégrale. Une partie des occurrences qui les contiennent devraient donc en principe être écartées ce qui n'est possible qu'à l'aide d'un contrôle manuel du contexte élargi.

³ Une requête dans le corpus entier, tenant compte de toutes les terminaisons masculines au singulier, suivies par une conjonction de coordination, par zéro, un ou deux autres éléments et par un nom féminin au singulier fait trouver 11340 occurrences qui ne sont toutefois pas toutes pertinentes.

5.3. Étapes de travail

Nous avons d'abord répertorié dans le corpus entier les formes doubles dont les noms sont séparés par une barre oblique, en vérifiant manuellement la pertinence des occurrences. Ensuite, nous avons identifié une gamme large de types d'occurrences : nous avons cherché les noms féminins se terminant en ".+innen", suivis par un élément taggé comme conjonction et, à une distance entre zéro et deux éléments, par un nom qui ne se termine pas par ".+innen". Nous avons refait ce dernier type d'examen en inversant l'ordre des deux éléments. Nous avons vérifié manuellement, ligne par ligne, la pertinence des occurrences de ce *pattern* pour les périodes lors desquelles il est peu fréquent. Pour certaines périodes, la quantité très importante de données a empêché un contrôle minutieux avec observation du contexte élargi. Dès qu'une requête a généré plus de 700 occurrences, nous avons fait une estimation de la fréquence relative des résultats pertinents. Pour ce faire, nous nous sommes servie de la fonction *thinning* [élagage] de *CQPweb*. Le nombre de lignes vérifiées manuellement a dépendu du nombre total des occurrences de la période observée⁴. Pour finir, nous avons identifié à l'aide de la fonction *frequency breakdown* [liste de fréquence lemmatisée] les candidats les plus fréquents du *pattern* contenant une conjonction de coordination, et nous avons ensuite examiné l'évolution quantitative des deux têtes de liste dans le corpus entier.

6. Analyse des données

6.1. Pattern au pluriel contenant une barre oblique

Nous avons recherché, dans le corpus entier, toutes les occurrences des formes doubles au pluriel comportant une barre oblique.

Requêtes : féminin avant masculin : [word="[A-Z].+innen\|.+"] ; masculin avant féminin : [(word="[A-Z].+\|.+innen")&(word!=".\|.?.innen")]

| Période | féminin avant masculin | masculin avant féminin |
|-------------|------------------------|------------------------|
| Années 1980 | 0,0 | 0,24 |
| Années 1990 | 0,82 | 3,44 |
| Années 2000 | 1,08 | 3,08 |
| 2011-2014 | 2,7 | 2,07 |

Tableau 2 - Fréquence relative pour les formes doubles avec barre oblique.

Nous avons d'abord recensé les formes féminines au pluriel (".+innen"), suivies par une barre oblique, et ensuite, nous avons recherché le *pattern* "barre oblique suivie par ".+innen", en supprimant les mots comportant une barre suivie directement de "innen" ("Mitarbeiter/innen" [collaborateurs/trices]) ou d'un trait d'union avant "innen", comme dans "Mitarbeiter/-innen" [collaborateurs/-trices] ; il se trouve que ces mots sont des formes abrégées et non des formes intégrales. Par contre, nous avons comptabilisé les formes semi-intégrales avec ou sans trait d'union, comme on en trouve dans les exemples suivants : *Tourismus-Experten/-Expertinnen* [experts en tourisme/-expertes] ; *Spezialärzte/ärztinnen* [médecins hommes spécialisés/médecins femmes]. Le deuxième nom comporte, de façon allusive, la spécification exprimée littéralement pour le premier.

⁴ Nous nous sommes servie du tableau suivant pour déterminer la taille d'un échantillon tolérant jusqu'à 5 % d'erreurs : <https://fr.surveymonkey.com/mp/sample-size/>

Les résultats ont été rectifiés après contrôle manuel. Nous n'avons pas trouvé d'occurrences avant les années 80. Le Tableau 2 permet de constater qu'il s'agit ici d'un phénomène très récent et marginal, avec une distribution assez équilibrée entre les deux ordres syntaxiques possibles.

6.2. Pattern au pluriel contenant une conjonction de coordination

6.2.1. Ordre "féminin – masculin"

Nous avons recensé les occurrences pour lesquelles la forme ".+innen", restreinte aux noms féminins au pluriel, est suivie par une conjonction de coordination, par zéro, une ou deux formes libres et ensuite par un nom masculin au pluriel. Le Tableau 3 contient des fréquences contrôlées jusqu'aux années 80 incluses et ensuite des résultats estimés à l'aide de la fonction "*thinning, reproducible*" [élagage, reproductible]. Nous avons minoré les résultats en fonction du pourcentage d'occurrences non pertinentes constaté.

Requête : [(word=".+innen")&(rfpos="N.*Pl.Fem.*")][pos="KON"][] {0,2} [rfpos="N.*Pl.Masc.*"]

| Période | Fréquence | | | Nombre de lignes observées | Occurrences non pertinentes |
|-------------|-----------|-----------|---------|----------------------------|-----------------------------|
| | relative | contrôlée | estimée | | |
| 1849-1950 | 0,028 | x | | | |
| Années 1950 | 1,639 | x | | | |
| Années 1960 | 1,95 | x | | | |
| Années 1970 | 5,4 | x | | | |
| Années 1980 | 12,82 | x | | | |
| Années 1990 | 306,95 | | x | 300 | 0,66 % |
| Années 2000 | 569,44 | | x | 370 | 0,37 % |
| 2011-2014 | 682,72 | | x | 300 | 1 % |

Tableau 3 - Pattern "féminin - conjonction - {0,2} - masculin".

Après contrôle manuel, la plupart des formes doubles du premier siècle ont dû être éliminées en raison de leur emploi particularisant. La première forme pertinente, *Arbeiterinnen und Arbeiter* [ouvrières et ouvriers], apparaît en 1909. Les années 70 sont marquées par le droit de vote enfin accordé aux femmes suisses à partir de 1971. Il est donc normal de trouver des occurrences qui citent nommément les électrices. Nous pouvons constater que les deux formes les plus fréquentes des années 70, *Kolleginnen und Kollegen* [collègues femmes et collègues hommes] et *Mitbürgerinnen und Mitbürger* [concitoyennes et concitoyens], ont un caractère discursif particulier : elles apparaissent de façon privilégiée dans des scripts de discours oraux. Les années 90 marquent un tournant quantitatif pour les formes doubles suivant l'ordre "féminin – masculin" et contenant une conjonction de coordination. On passe de 154 occurrences pour les années 80 à 6316 occurrences. Les trois formes les plus fréquentes des années 90 ont des fréquences absolues qui sont environ trente à quarante fois plus importantes que celles des années 80 (554, 391 et 285 *versus* 13, 12 et 11). Pendant les années 2000, l'évolution des formes doubles progresse encore : par rapport aux années 90, la fréquence relative est presque doublée. Les têtes de liste ont des fréquences absolues encore plus importantes que celles des années 90.

6.2.2. *Ordre "masculin – féminin"*

Pour examiner l'ordre "masculin – féminin", nous avons recensé les occurrences de formes doubles intégrales dans lesquelles un nom masculin au pluriel est suivi par une conjonction de coordination, zéro, une ou deux formes libres et ensuite par un nom féminin au pluriel, se terminant en ".+innen".

Requête : [rfpos="N.*Pl.Masc.*"][pos="KON"][] {0,2} [(word=".+innen")&(rfpos="N.*Pl.Fem.*")]

| Période | Fréquence | | | Nombre de lignes observées | Occurrences non pertinentes |
|-------------|-----------|-----------|---------|----------------------------|-----------------------------|
| | relative | contrôlée | estimée | | |
| 1849-1900 | 3,37 | x | | | |
| 1901-1950 | 10,27 | x | | | |
| Années 1950 | 21,72 | x | | | |
| Années 1960 | 17,14 | x | | | |
| Années 1970 | 10,98 | x | | | |
| Années 1980 | 6,07 | x | | | |
| Années 1990 | 81,31 | | x | 200 | 7 % |
| Années 2000 | 128,61 | | x | 300 | 9 % |
| 2011-2014 | 82,64 | | x | 300 | 23,33 % |

Tableau 4 - Pattern "masculin - conjonction - {0,2} - féminin".

On peut constater que l'ordre masculin – féminin est représenté pendant toute la période couverte par le *Corpus Bundesblatt*, avec des fréquences relatives modestes. Après un pic assez significatif pendant les années 2000, la fréquence relative constatée pour la période 2011-2014 retombe au niveau des années 90. Les données observées pour l'ordre "masculin – féminin" ont deux particularités : premièrement, 3,95 % des formes doubles identifiées apparaissent dans des noms d'associations, donc des noms propres. Nous avons fait le choix de les prendre en compte (cf. 8.4.). Deuxièmement, pour les périodes dont les fréquences ont été estimées, la requête a mené à un résultat contenant un taux élevé d'occurrences non pertinentes, en particulier pour la période la plus récente. Il s'agit de pivots qui contiennent deux noms différents ou des noms désignant autre chose que des personnes. On constate également qu'un bon nombre de formes doubles se trouvent à cheval sur leur contexte de gauche et / ou de droite⁵, comme dans cet exemple de 2011 [des femmes juges et **des hommes juges ainsi que des avocates** et avocats].

| | | |
|------------------|----------------------------------|-------------|
| Richterinnen und | Richter sowie Anwältinnen | und Anwälte |
|------------------|----------------------------------|-------------|

⁵ Nous avons recherché les formes doubles à cheval sur leurs contextes de gauche et / ou de droite, séparées par la conjonction *sowie* [ainsi que]. Il y a 3,65 fois plus d'occurrences de ce type dans lesquelles les deux noms du pivot suivent l'ordre "masculin – féminin" (1949-2014) que d'occurrences suivant l'ordre "féminin – masculin" (1996-1984).

Ces éléments font partie de formes doubles suivant l'ordre "féminin – masculin" et n'ont donc pas été prises en compte.

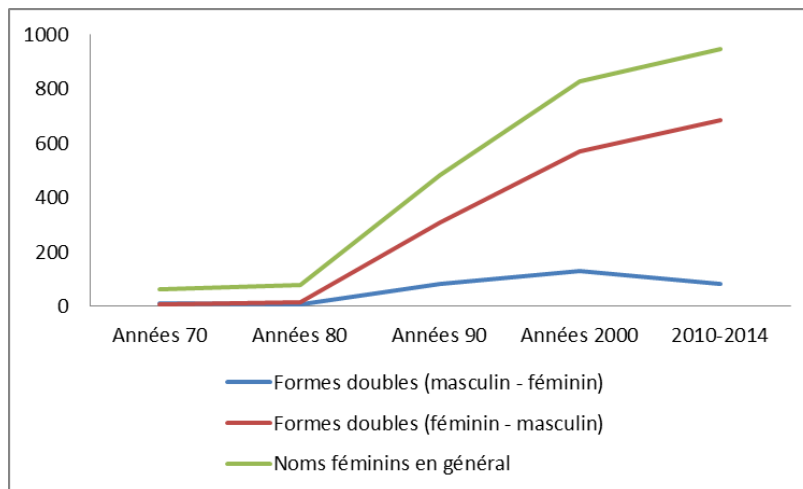


Figure 1 - Évolution des noms féminins au pluriel.

Le taux d'occurrences non pertinentes augmente avec le temps. Il semble légitime de dresser un parallèle avec l'évolution fulgurante des noms féminins pendant ce laps de temps, qu'ils fassent ou qu'ils ne fassent pas partie d'une forme double. Cette évolution générale dépasse clairement celle des formes doubles attestée en 6.2.1. et 6.2.2., même les fréquences des deux ordres syntaxiques réunies.

6.3. Formes les plus fréquentes

Après l'étude des divers types de *pattern* dans leur globalité, nous nous intéressons à la forme double intégrale la plus fréquente pour l'ordre "féminin – masculin", à savoir *Arbeitnehmerinnen und Arbeitnehmer* [salariées et salariés] et à la forme la plus fréquente pour l'ordre "masculin – féminin", à savoir "*Richter und Richterinnen*" [juges hommes et juges femmes]. Nous avons également observé les pendants syntaxiques de ces deux groupes de noms. Aucune occurrence avec barre oblique n'a été trouvée, et celle avec virgule est un phénomène très marginal.

Word= "*Arbeitnehmerinnen*" & lemma= "*Arbeitnehmer*" et vice versa

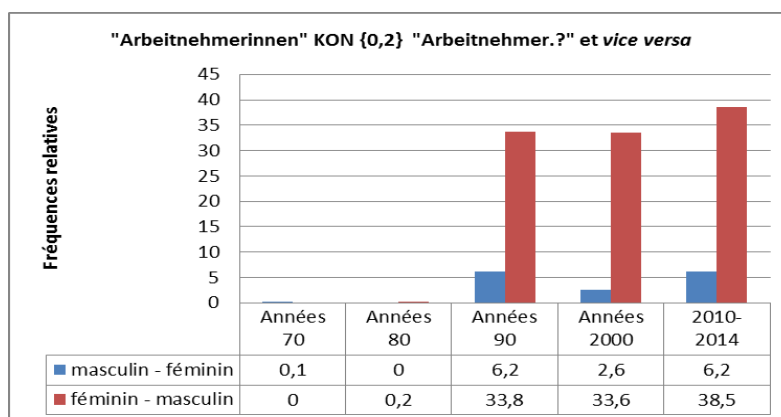
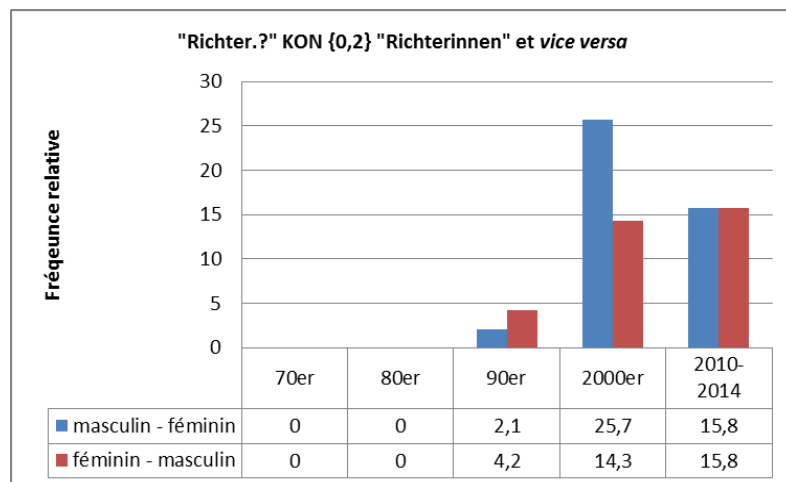


Figure 2 - Évolution de "*Arbeitnehmerinnen & Arbeitnehmer.?*".

conjonction de coordination – {0,2} – nom". Nous n'avons trouvé ni adjectifs, ni adverbes, ni chiffres après la conjonction. On pourrait presque considérer ces formes comme semi-figées, d'autant plus que sur les quatre types de réalisation observés, les deux qui comportent la conjonction *und* [et] se taillent la part du lion, à savoir *Arbeitnehmerinnen und Arbeitnehmer* [salariées et salariés] et *Arbeitnehmerinnen und Arbeitnehmern* [aux salariées et aux salariés].

La Figure 2 illustre le fait qu'il s'agit avec *Arbeitnehmerinnen & Arbeitnehmer* [salariées & salariés] d'une forme récente et en évolution continue. L'ordre "féminin – masculin" domine aussi du point de vue quantitatif. Aussi bien pour l'ordre "féminin – masculin" que pour l'ordre inverse, nous avons identifié peu de variations pour la réalisation du *pattern* de base "nom –

Lemma="Richter" & word="Richterinnen" et vice versa



L'évolution du *pattern* contenant *Richter* [juges hommes] et *Richterinnen* [juges femmes] diffère de celle des autres têtes des listes de fréquence de chaque décennie. L'ordre "masculin – féminin" s'impose davantage ici que pour les autres formes, et il finit à égalité avec l'ordre "féminin – masculin".

Figure 3 - Évolution de "Richter.?" & Richterinnen".

7. Résultats

Rappelons les défis à surmonter : il s'agissait d'identifier le plus de formes pertinentes possibles répondant aux critères formels et sémantiques des formes doubles intégrales au pluriel, en formulant des requêtes suffisamment ouvertes, sans négliger les vérifications locales en cas de doute par rapport à l'usage généralisant ou particularisant d'une occurrence.

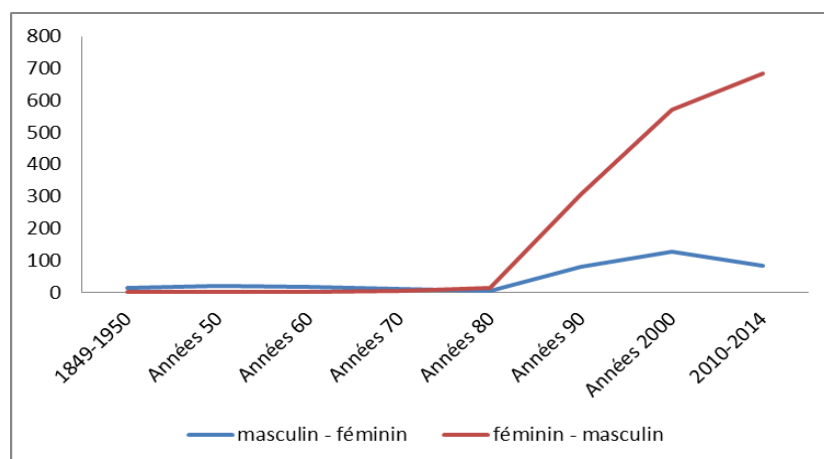


Figure 4 - Formes doubles au pluriel avec conjonction.

"masculin – féminin" dépasse légèrement celle de l'ordre "féminin – masculin". Après avoir atteint un score assez significatif pendant les années 2000, la fréquence de ce *pattern* régresse lors de la période la plus récente (2011-2014). À partir des années 90, période qui nous intéresse particulièrement, les différentes valeurs attestées pour l'ordre "masculin – féminin" restent bien en-dessous de celles que l'on constate pour l'ordre syntaxique "féminin – masculin". Depuis les années 90, ce *pattern* est en progression spectaculaire et continue.

Nos résultats font apparaître une distribution inégale entre l'ordre syntaxique "forme féminine suivie de la forme masculine" et l'ordre inverse. La Figure 4, retraçant l'évolution historique intégrale du *pattern* avec conjonction de coordination, permet de voir que jusqu'à la fin des années 70, la fréquence relative de l'ordre syntaxique "masculin –

Nous nous sommes attendue à ce que les adjectifs, adverbes et virgules entravent nos résultats. Toutefois, les têtes de listes de fréquence des formes doubles en sont démunies. De plus, le caractère quasiment figé de ces têtes de liste permet de relativiser la question des valeurs des fréquences qui sont moins contrôlées que les autres. En effet, quand les données

étaient trop nombreuses, nous avons fait des estimations de pertinence. Nous pensons que ces estimations reflètent bien la réalité, d'autant plus que les écarts entre les deux ordres syntaxiques sont tellement importants que les valeurs restent parlantes malgré les adaptations nécessitées par le pourcentage parfois très élevé d'occurrences non pertinentes des requêtes.

8. Les cas limites

Nous nous intéressons uniquement aux noms de personnes représentant un emploi généralisant, désignant des hommes et des femmes sans pointer l'un des deux sexes en particulier. Cet emploi n'est pas toujours garanti dans les données du *Bundesblatt*. Nous allons revenir ici sur les cas de figure observés dans notre présentation.

8.1. Adjectifs et adverbess

Il faut que l'élément spécifiant les noms du *pattern* soit le même pour les deux noms. Autrement, on risque d'avoir affaire à des situations dans lesquelles le rédacteur ou la rédactrice fait référence à l'un des deux sexes en particulier. Cela est le cas dans l'exemple suivant : "*Fr. 1.14 pro Stunde für ledige Arbeiter und alle Arbeiterinnen, die das 20. Altersjahr erreicht haben*" [1.14 fr pour les ouvriers célibataires et toutes les ouvrières qui ont atteint l'âge de 20 ans]. Cet exemple de 1960 reflète d'ailleurs la différence faite entre hommes et femmes en matière de salaire. Le quantifieur *alle* [toutes les], placé après la conjonction, est un signal clair pour un emploi particularisant ; sans sa présence, notre exemple serait un réel cas limite. Une autre situation est à prendre en compte, à savoir quand un nom est précédé d'un adjectif soulignant l'identité féminine ou masculine de la personne. Exemple de 1964 : "*bei weiblichen Angestellten und Arbeiterinnen fällt überdies die Heirat als zusätzliche Austrittsursache stark ins Gewicht*" [De plus, chez les employés féminins et les ouvrières, le mariage est une raison de départ qui pèse lourd].

8.2. Chiffres précédant les noms

Dans certaines occurrences, les noms sont précédés de chiffres. Exemple de 2002 : "*An der EAWAG lehren (...) 179 Mitarbeiterinnen und 225 Mitarbeiter*" [179 collaboratrices et 225 collaborateurs enseignent (...) à la EAWAG]. S'il s'agit de chiffres différents pour les deux noms, l'emploi de ces occurrences est particularisant puisqu'on pointe les deux sexes séparément.

8.3. Formes semi-intégrales

Les formes semi-intégrales, telles que "*Tourismus-Experten/-Expertinnen*" [experts en tourisme/expertes], ont-elles droit de cité à côté des formes pleinement intégrales ? En raison de leur ressemblance formelle, nous avons fait le choix de les inclure dans notre étude.

8.4. Noms propres

En principe, nous avons eu l'intention de considérer les noms propres comme non pertinentes. Il se trouve que 329 occurrences du corpus (1,61 mots par million d'occurrences) sont des noms d'associations contenant une forme double intégrale : "*Verband abstinenter Lehrer und Lehrerinnen*" [Association d'enseignants et enseignantes abstinentes]. Cette situation ne peut pas être assimilée au cas des noms propres qui sont des homophones de noms communs, désignant des personnes. C'est la raison pour laquelle nous les avons pris en compte.

9. Conclusion

Notre étude permet d'attester des tendances relativement récentes d'un usage de plus en plus fréquent de formes doubles intégrales, donc d'un type de formulation non sexiste parmi plusieurs, dans le domaine de la langue administrative suisse. Nous avons constaté que la barre oblique est rarement utilisée. Ces résultats sont compatibles avec les recommandations exprimées dans le *Leitfaden* (Schweizerische Bundeskanzlei, 2009).

Depuis les années 90, la prédominance de l'ordre syntaxique "féminin – masculin" est forte. Quant au pouvoir signifiant du choix de l'ordre des deux éléments d'une forme double, les avis divergent : à côté du principe "*ladies first*", des arguments prosodiques, topologiques ou autres sont avancés pour défendre l'un ou l'autre ordre, voire les deux à la fois.

Il reste encore du progrès à faire en matière législative et dans l'évolution de la société concernant la présence des femmes "dans les milieux économiques, scientifiques, politiques, l'administration et la vie publique" et "dans les positions d'influence" (Commission fédérale pour les questions féminines CFQF, nd ; voir aussi l'exemple des juges femmes et juges hommes en 6.3.). Toutefois, selon nos résultats, le gouvernement suisse a réussi le pari de rendre les femmes visibles dans bon nombre de situations où elles sont présentes autant que les hommes, du moins dans les textes diffusés en allemand.

10. Recherches futures

Nous continuerons nos recherches sur les formes doubles en nous intéressant au volet français, donc au *Corpus Feuille fédérale*. Nous aimerions alléger notre méthode, jusqu'à présent très prudente et marquée par des vérifications manuelles chronophages. Pour ce faire, nous comptons nous appuyer davantage sur des approches pratiquées en statistique textuelle.

11. Remerciements

Nous remercions chaleureusement Klaus Rothenhäusler pour son aide précieuse lors de la phase de création du corpus *Bundesblatt/Feuille fédérale/Foglio federale*. Un grand merci aussi à Kirsten Adamzik, Alessandra Alghisi, Françoise Demaizière, Daniel Elmiger, Verena Tunger et les relecteurs ou relectrices anonymes pour avoir commenté une version précédente de ce texte.

Références

- Chancellerie fédérale (1849-2014). *Bundesblatt* [feuille fédérale]. <https://www.admin.ch/gov/de/start/bundesrecht/bundesblatt.html>
- Chancellerie fédérale (2007). *Loi fédérale sur les langues nationales et la compréhension entre les communautés linguistiques (Loi sur les langues, LLC)*. <https://www.admin.ch/opc/fr/classified-compilation/20062545/index.html>
- Chancellerie fédérale (2015). *Explications concernant la feuille fédérale*. <https://www.admin.ch/gov/fr/accueil/droit-federal/feuille-federale/explications-concernant-la-feuille-federale.html>
- Commission fédérale pour les questions féminines CFQF (nd). "Femmes et égalité en général". <http://www.ekf.admin.ch/themen/00503/index.html?lang=fr>
- Commission fédérale pour les questions féminines CFQF (2009). "Femmes Pouvoir Histoire 1848-2000". <http://www.ekf.admin.ch/dokumentation/00444/00517/index.html?lang=fr>

- Elmiger D. (2015a). "La répétition de noms communs de personnes pour éviter le masculin à valeur générique". *Le discours et la langue*, vol. 7(2) : 97-112.
- Hardie A. (2012). CQPweb - combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics*, 17/3: 380-409.
- Schweizerische Bundeskanzlei (2009). [Geschlechtergerechte Sprache - Leitfaden zum geschlechtergerechten Formulieren im Deutschen](https://www.bk.admin.ch/dokumentation/sprachen/04915/05313/index.html?lang=de) [Langage non sexiste - Guide de formulation non sexiste en allemand]. 2^{ème} édition, revue et corrigée. 1^{ère} édition : 1996. <https://www.bk.admin.ch/dokumentation/sprachen/04915/05313/index.html?lang=de>

Corpus et outils

- Elmiger D. (2015b). *Les corpus Feuille fédérale / Bundesblatt / Foglio federale*. V. 1.2. (430/830). Genève: Retrieved from <http://archive-ouverte.unige.ch/unige:80593>
- Evert S. (2010). *CQPweb*. Système de gestion de corpus. <http://cwb.sourceforge.net/cqpweb.php>
- Schiller, A., Thielen, C., Teufel, S. & Stöckert, C. (1995/1999). *STTS (Stuttgart-Tübingen Tagset)*. <http://www.ims.uni-stuttgart.de/projekte/corplex/TagSets/stts-table.html>
- Schmid, H. & Laws, F. (2008). "Estimation of Conditional Probabilities with Decision Trees and an Application to Fine-Grained POS Tagging". *COLING 2008*, Manchester, Angleterre. <http://www.cis.uni-muenchen.de/~schmid/papers/Schmid-Laws.pdf>