



**HAL**  
open science

# Continu et discontinu dans l'œuvre poétique d'Apollinaire : la phrase apollinarienne au croisement de l'analyse stylistique et des statistiques textuelles

Clémence Jacquot

## ► To cite this version:

Clémence Jacquot. Continu et discontinu dans l'œuvre poétique d'Apollinaire : la phrase apollinarienne au croisement de l'analyse stylistique et des statistiques textuelles. Actes du CMLF 2012, 1, pp.1099-1116, 2012, 10.1051/shsconf/20120100131 . hal-02018372

**HAL Id: hal-02018372**

**<https://hal.science/hal-02018372>**

Submitted on 13 Feb 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ***Continu et discontinu* dans l'œuvre poétique d'Apollinaire : la phrase apollinarienne au croisement de l'analyse stylistique et des statistiques textuelles**

Jacquot, Clémence

Université Toulouse II le Mirail (laboratoire PLH/ELH)  
clemence.jacquot@gmail.com

## **Introduction**

« J'ai fait mon possible pour simplifier la syntaxe poétique » écrit Apollinaire à sa fiancée Madeleine en 1915<sup>1</sup>. Cette petite phrase d'apparence anodine est intéressante sous la plume du poète ; en effet, rares sont, chez lui, les commentaires et les prises de position métapoétiques. De fait, la volonté de « simplification » de la phrase sonne comme un programme d'écriture, qui fait d'ailleurs écho à nombre de réflexions esthétiques sur la langue et l'expression poétique dans un monde où l'avant-garde cherche à prendre ses distances du symbolisme finissant. Nous proposons dans l'étude présentée ici de prendre au mot Apollinaire et d'analyser comme un phénomène particulier de son écriture poétique la question de la « simplification » syntaxique. La simplification de la phrase apollinarienne implique que l'on s'intéresse à la manière dont le poète articule et relie les segments et les propositions. Elle invite dès lors à étudier précisément le type de liaisons syntaxiques utilisées par Apollinaire ainsi que l'armature logique sur laquelle elle repose : Cette dernière est-elle explicite ? Pourquoi ? Quels sont les moyens privilégiés de relier les phrases entre elles ? Cela correspond-il ou non à la volonté qu'exprime Apollinaire de simplifier la syntaxe ?

Lorsque l'on examine dans le détail le poème « Les Fenêtres », exemple du phénomène de simplification cité par le poète, on est frappé par l'abondance des phrases averbales apposées, par les effets de juxtaposition de prédicats et, il faut le dire, par la forte absence de liaisons syntaxiques (en particulier des conjonctions de subordination). Il semble qu'Apollinaire cherche précisément à donner ici une impression de *discontinuité*, c'est-à-dire qu'il joue sur ces effets de juxtaposition, de « collage »<sup>2</sup>, voire de mise en liste pour déconstruire mouvements et perceptions du monde extérieur et se faire, selon sa propre expression, le récepteur du « lyrisme ambiant »<sup>3</sup>. Nous parlerons donc de *discontinuité* pour définir ce type d'écriture qui tend vers la parataxe, et de *continuité* pour qualifier l'effet inverse, c'est-à-dire les moments où Apollinaire privilégie au contraire la complexité syntaxique de ses phrases en accentuant les marqueurs de liaison et les indices explicites de l'armature logique des propositions<sup>4</sup>.

L'étude que nous proposons s'inscrit dans le sillage de plusieurs travaux lexicographiques menés sur le corpus poétique d'Apollinaire<sup>5</sup>, en particulier ceux du collectif Hubert de Phalèse sur *Alcools*, mais cherche à élargir le champ d'investigation que ces précédents ouvrages ont pu dégager. Cet élargissement est nécessaire d'abord parce que ces travaux concentraient leur propos sur un seul recueil d'Apollinaire : *Alcools*, et ce pour des raisons techniques évidentes, telles que le comptage manuel des occurrences, dans le cas de P. Guiraud, qui impose, on le comprend, une réduction du corpus d'étude ; ou encore l'édition numérique du seul et unique recueil *Alcools* sur la base Frantext pour l'analyse de Hubert de Phalèse. On constate ensuite que ces deux ouvrages sont dépendants de l'évolution des techniques d'investigation lexicographique, en sorte qu'ils focalisent leurs remarques sur des classements du lexique en réseaux thématiques et « mots-clés ». Nous privilégions, quant à nous, la perspective stylistique et utilisons les ressources du Traitement Automatique des Langues (TAL) pour mettre en lumière et interroger de façon systématique les particularités syntaxiques du discours poétique d'Apollinaire. Pour analyser les effets de *continuité* et de *discontinuité*, nous avons choisi de nous concentrer sur l'analyse de la présence ou de l'absence des marqueurs de liaison syntaxique, en particulier les conjonctions, en comparant les corpus respectifs des œuvres poétiques et des œuvres en prose d'Apollinaire. Peut-on parler réellement d'une

simplification de la syntaxe poétique, en regard de l'écriture prosaïque ? C'est la question à laquelle nous essaierons de répondre.

## 1 Construction du corpus numérique apollinarien : méthodologie et enjeux théoriques

Utiliser des outils statistiques pour construire une analyse stylistique est encore parfois, en France, une entreprise délicate. On a généralement tendance à penser que l'informatique, en général, et les logiciels de comptage et rapatriement systématique des occurrences « aplatissent » le texte, le réduisent à de pures « données ». De fait, ils ne permettent pas, loin s'en faut, d'apprécier la structure globale d'une œuvre, ainsi que sa valeur esthétique ! Est-il pour autant judicieux de se couper de leurs services ? Ne serait-il pas préjudiciable, sous prétexte qu'il est impossible de réduire les chefs-d'œuvre littéraires à de simples systèmes linguistiques, d'affirmer que les outils de statistique textuelle n'offrent aucune méthode ni aucun point de vue intéressants ? Le préjugé esthétique (infondé ?) qui les poursuit va de pair avec celui qui veut que l'on ne puisse pas « décortiquer » ou « rationaliser » la poésie. Bien entendu, certaines de ces critiques sont fondées : la mauvaise utilisation des résultats statistiques mène à de purs sophismes voire à de véritables catastrophes logiques. Les conclusions hâtives, biaisées ou insuffisamment remises en contexte sont légion : elles induisent alors nombre de malentendus et de surinterprétations. Une méthode de travail rigoureuse est donc nécessaire, non seulement pour utiliser ces outils de manière appropriée, mais aussi pour poser des garde-fous et permettre d'interpréter correctement les résultats obtenus. Dès lors, les logiciels statistiques et les corpus annotés apparaissent comme des dispositifs particulièrement féconds pour éprouver les hypothèses de lecture littéraires.

En stylistique, par exemple, l'approche pragmatique et systématique qu'ils proposent élargit les horizons d'études : en effet, loin de se borner désormais à la seule analyse systématique du vocabulaire et des champs sémantiques spécifiques à un auteur, des fonctions telles que le calcul de spécificités de chaque corpus étudié ou encore les analyses factorielles des correspondances (AFC)<sup>6</sup> permettent d'identifier ou de préciser ce qui avait été intuitivement qualifié de « faits de langue », réactualisant ainsi, de façon opératoire, la notion d'« écart » stylistique. Autre fait d'importance : la numérisation et l'annotation des corpus étendent les frontières du champ de recherches. On peut désormais ne plus se restreindre à analyser seulement un ou deux recueils de l'auteur (*Alcools* et/ou *Calligrammes*) : il est maintenant possible d'embrasser l'œuvre entière d'Apollinaire, ou peu s'en faut.

### 1.1 Le corpus apollinarien

L'œuvre d'Apollinaire, soldat volontaire de la Première Guerre Mondiale et « mort pour la Patrie » en 1918, est toujours sous droits d'auteurs : ceux-ci ont été étendus jusqu'à 2018. Par conséquent, il n'y a pas (exception faite de quelques textes) de version numérique corrigée et amendée des œuvres apollinariennes que l'on puisse exporter et traiter grâce aux logiciels de statistique textuelle. Pour construire ce corpus, il nous a donc fallu copier, numériser la quasi-totalité des textes d'Apollinaire, puis l'annoter manuellement afin de le rendre apte à être lu et analysé par le logiciel que nous avons privilégié pour cette étude : TXM<sup>7</sup>.

Il est important, pour répondre à la question du *continu* et du *discontinu* dans la syntaxe poétique d'Apollinaire, d'avoir une vision globale de son œuvre littéraire, afin de pouvoir comparer les usages des conjonctions dans les différents genres employés par l'auteur. Nous avons donc travaillé sur l'ensemble, à quelques exceptions près<sup>8</sup>, de la production d'Apollinaire, composée d'œuvres en prose (contes et romans) et de recueils poétiques. L'ensemble constitue un corpus de quatorze ouvrages, soient 210 800 mots environ. Trois d'entre eux sont écrits en prose : *L'Enchanteur pourrissant* (conte, 1909), *L'Hérésiarque et Cie* et *Le Poète assassiné* (tous deux recueils de contes, datant respectivement de 1910 et de 1916). Les neuf autres sont des recueils de poèmes : *Bestiaire ou le Cortège d'Orphée* (1911), *Alcools* (1913), *Vitam impendere amori* (1917), *Calligrammes* (1918), *Il y a* (1925), *Poèmes à Lou* (1947), *Poèmes à La marraine de guerre* (1948), *Poèmes à Madeleine* (1949), *Le Guetteur mélancolique*

(1952), *Poèmes inédits* (1967) et *Poèmes retrouvés* (1967)<sup>9</sup>. Ce corpus, qui n'est pas totalement complet à ce jour, donne toutefois une vision d'ensemble relativement large de la production littéraire d'Apollinaire et permet d'appréhender la différence générique et l'éventuelle opposition syntaxique que l'on pourrait rencontrer entre les œuvres poétique et prosaïque d'Apollinaire.

### 1.1.1 Mise en page et agencement des textes : « vilisibilité »<sup>10</sup> du texte poétique

Grâce à la numérisation des œuvres d'Apollinaire, il est possible d'avoir un regard à la fois surplombant et plus précis sur l'ensemble de ses œuvres littéraires. Des problèmes se présentent alors à qui s'attache à l'annotation de ce corpus, et en particulier la question relative à la mise en page des poèmes. De fait, la numérisation des textes d'Apollinaire conduit effectivement, dans un premier temps, à les aplatir et à les réduire à une simple succession de mots. Or, la mise en forme du texte est une partie essentielle de la poétique d'Apollinaire en particulier, et de la poésie depuis la fin du XIXe siècle plus généralement. La page est alors partie intégrante de la « fabrique » du poème ; bien plus, elle devient l'un des éléments du dispositif poétique et croise même des problématiques esthétiques.

L'importance de la mise en page correspond d'abord à l'intérêt accru qu'ont manifesté de nombreux poètes et artistes (cubistes ou futuristes, par exemple) pour de nouvelles formes et de nouveaux supports susceptibles de représenter la réalité. D'après les vers d'Apollinaire, dans « Zone » – l'un des poèmes les plus originaux et les plus modernes d'*Alcools* :

Tu lis les prospectus les catalogues les affiches qui chantent tout haut  
Voilà la poésie ce matin et pour la prose il y a les journaux [...]

En somme, la poésie se « nourrit » des moyens d'expression populaires, aussi concis que puissants, que sont les publicités, catalogues, magazines, affiches, et autres « étiquettes de parfumeurs »<sup>11</sup>. L'affirmation de l'importance de la page poétique plante ses racines dans les réflexions esthétiques des écrivains et des artistes touchant au renouvellement des moyens langagiers et plastiques pour représenter le réel. À partir de la publication du poème de Mallarmé « Un coup de dés jamais n'abolira le hasard » – et surtout de sa redécouverte grâce à l'étude qu'en fait Thibaudet en 1911<sup>12</sup> –, de nombreux poètes du début du XXe siècle ont pris conscience de la matérialité tant de la page que du signe linguistique, réinvestissant dès lors sa valeur graphique et plastique.

Cette prise en compte du potentiel expressif et signifiant de la mise en page prend également place dans un contexte de réflexion sur ce qui définit le langage poétique, et ce qui réunit ou disjoint texte littéraire et arts visuels. C'est à cette époque que la littérature et les arts remettent en cause les paradigmes de la mimésis et cherchent à éprouver de nouveaux régimes de visibilité et de représentation du réel, en procédant notamment à une fragmentation des plans et des discours. La réalité apparaît alors comme une somme de perceptions complexes et discontinues : le lien – l'impression de *continuu* – est à recomposer *a posteriori*. C'est ce qui est à l'œuvre, à plus ou moins grande échelle, chez des poètes aussi divers qu'Apollinaire, Segalen, Cendrars, Marinetti ou encore, un peu plus tard chez les dadaïstes ou dans la poésie concrète de Pierre Garnier.

Dans la poésie apollinarienne, l'articulation entre *continuu* et *discontinuu* trouve un écho singulier et particulièrement opératoire dans l'entreprise calligrammatique. Les calligrammes, objets poétiques singuliers au carrefour de deux systèmes sémiotiques différents, permettent à Apollinaire d'appréhender la matérialité plastique du signe. Il mêle la dimension linguistique et la dimension iconique dans un même espace : celui du poème. Le calligramme fabrique ainsi du *continuu* en réinventant le lien entre les deux sémiologies et répond à « l'arbitraire du signe » en remotivant visuellement le langage poétique.

Le corpus apollinarien est donc un cas singulier et très intéressant pour étudier la notion de *continuu* et de *discontinuu* dans le langage poétique. L'importance de la mise en page, de la « vilisibilité » découlent-elles de la « simplification » de la syntaxe mentionnée précédemment ? Cette dernière est-elle au contraire une conséquence de la prise de conscience du potentiel expressif du signifiant ? On constate en effet dans un certain nombre de poèmes d'Apollinaire une dispersion, un éparpillement sur la page de syntagmes nominaux, de groupes de phrases et de vers libres en strophes irrégulières traduisant la vision

fragmentée du monde moderne. Les calligrammes, par leur disposition typographique atypique, vont jusqu'à bouleverser l'ordre de lecture et concurrencer la verticalité du texte poétique. Certains d'entre eux, comme « La Mandoline l'œillet et le bambou » ou encore « Lettre-Océan » (*Calligrammes*)<sup>13</sup> présentent des figures circulaires et imposent donc au lecteur de choisir, plus ou moins arbitrairement, un point – *focal* ? – de départ pour lire le poème. Dans « Lettre-Océan », le monde est perçu de façon topologique (cf. Longrée, 1981). Ce poème calligrammatique associe plusieurs représentations iconiques de la Tour Eiffel vue du dessus et le thème de la propagation acoustique des ondes télégraphiques, suggérée par le titre. Un tel poème offre donc plusieurs alternatives de lecture et exprime nettement l'impression de *discontinuité* qu'éprouve le poète face au monde, pris dans un échange vertigineux d'informations en un va et vient incessant des émissions et des réceptions.

### 1.1.2 La construction du corpus annoté apollinarien

Le *discontinuité* dans le langage poétique apollinarien s'articule à l'opposition traditionnelle entre le *linéaire* et le *tabulaire*. De fait, nombreux sont les poèmes d'Apollinaire qui invitent le critique à reconsidérer les modalités de lecture du texte poétique. À ce niveau de l'analyse, il est nécessaire de préciser les choses : chez Apollinaire, la notion de *discontinuité* engage d'abord à réfléchir à la rupture de la linéarité de la phrase, à la « simplification », voire à la « déconstruction » de la syntaxe, par le biais de la dispersion des syntagmes que l'on mentionnait ci-dessus. Dès lors, une analyse stylistique du *discontinuité* apollinarien va de pair avec l'examen quantitatif et qualitatif des figures de liaison syntaxique – les conjonctions par exemple. Pour réaliser une telle analyse, nous avons choisi de construire un corpus annoté des œuvres littéraires d'Apollinaire, afin de pouvoir comparer systématiquement entre eux de grands ensembles génériques (les œuvres en prose et les œuvres poétiques), mais aussi d'affiner les remarques en variant l'échelle des corpus d'études – par exemple en raisonnant séparément sur chaque recueil poétique. Mais le problème du *discontinuité* se greffe désormais à la construction du corpus elle-même : il s'agit alors de prendre en considération la dimension *tabulaire* que peuvent avoir nombre de textes poétiques d'Apollinaire. Comment traduire un calligramme dans un corpus annoté ? Plus généralement, comment ne pas « aplatir » des poèmes qui se caractérisent, comme l'écrivait à juste titre le Groupe  $\mu$ , par le dosage subtil d'une lecture *tabulaire* et d'une lecture *linéaire* ?

Tout l'enjeu de la construction du corpus annoté des textes apollinariens réside dans le respect d'un certain « relief » de l'écriture. Il semble effectivement que les poèmes d'Apollinaire, en particulier les calligrammes, invitent le critique à repenser les modalités de la lecture poétique, notamment grâce à la mise en scène du geste d'écriture. M.-L. Floréa (2009) résume bien la complexité de l'articulation des notions de discontinuité, de linéarité et de tabularité :

[...] Bien des aspects du texte viennent remettre en question la centralité de la linéarité. Un texte est en effet traversé par de nombreuses discontinuités qui viennent segmenter le fil linéaire de la lecture : ces discontinuités peuvent être marginales [...] : il s'agit du titre, du nom d'auteur, des notes, du sommaire... Elles peuvent aussi être constitutives de la textualité, dès lors qu'elles affectent la structure du corps du texte lui-même, qui se trouve subdivisé en plusieurs modules textuels. Cette distinction entre les discontinuités marginales et structurelles me permet de distinguer deux régimes de textualité différents, qui sont autant de façons d'agencer le matériau langagier pour former un texte : le régime de textualité linéaire et le régime de textualité tabulaire. (§ 11-12)

Une fois ces définitions posées, elle rappelle à juste titre que la dimension tabulaire des textes est souvent oubliée dans les corpus annotés :

L'interpénétration des concepts de texte et de corpus fait que la représentation de l'un influe fortement sur la représentation de l'autre. Ainsi, la conception actuelle du texte, fortement marquée [...] par le primat du linéaire, influence la conception du corpus, qui ne se conçoit que de façon linéaire, ce qui implique de linéariser, d'aplatir les textes tabulaires, de les faire entrer de force dans un moule qui n'est pas conçu pour eux, et de perdre ainsi une bonne partie de leur essence. L'analyse des textes tabulaires est par conséquent limitée, étant donné qu'ils sont traités par des corpus

linéaires : tous les composants du texte sont sur le même plan, les discontinuités textuelles sont gommées, la dimension plurisémiotique est évincée. (§ 41)

Ces remarques font écho aux interrogations spécifiques au texte apollinarien. On pense bien entendu aux calligrammes, que nous avons déjà évoqués précédemment, poèmes par définition « plurisémiotiques », mais on peut également élargir le propos et considérer la dimension tabulaire des textes induite par la suppression généralisée de la ponctuation dans l'œuvre poétique d'Apollinaire à partir de 1913. Le corpus apollinarien pose de fait des contraintes très spécifiques, qui touchent du doigt les limites, ou mettent du moins en lumière quels seraient les points méthodologiques à perfectionner dans la construction d'un corpus poétique. Comment donc traduire et intégrer la discontinuité ainsi définie dans les corpus annotés ?

## 1.2 La technologie XML, une façon de donner du « relief » au texte poétique ?

Nous nous sommes efforcés, pour respecter au mieux, le « relief » spécifique au texte apollinarien, de prendre en compte tout ce qui pouvait servir à lui conférer une structure et à mettre en lumière une hiérarchisation des informations qui le caractérise, en particulier ce que G. Genette (1987) appelle le « paratexte ». Pour répondre à ces exigences, nous avons valorisé la technologie XML.

Les logiciels de statistique textuelle comme Lexico<sup>3</sup><sup>14</sup> ou Hyperbase<sup>15</sup> permettent seulement une approche linéaire des corpus. Pour lutter contre l'aplatissement du texte littéraire, nous avons choisi le logiciel TXM, justement parce qu'il utilise la technologie XML. Cette dernière nous est apparue comme le moyen le plus approprié pour rendre compte de la structure des textes apollinariens. La hiérarchisation des informations paratextuelles se présente comme une succession de blocs textuels interdépendants. Chacun d'entre eux est introduit par une étiquette spécifique qui met en évidence sa fonction dans la structuration du texte. Ces étiquettes sont des balises XML que nous avons disposées manuellement pour la plupart, automatiquement pour certaines (la mention des « vers » par exemple) en entrée et fin de segments textuels.

Le premier niveau de hiérarchisation des informations, pour traiter la question du *continu* et du *discontinu*, est la distinction générique entre œuvres en prose et œuvres poétiques. Nous l'avons exprimé de la manière suivante : `<ouvrage type="[au choix poesie ou prose]">`. Nous avons ensuite voulu mettre en valeur le titre de chaque ouvrage, de manière à pouvoir, par la suite, les dater et les comparer les uns par rapport aux autres (`<titre niveau="recueil">`) ; puis le titre de chaque poème, conte ou nouvelle au sein de ce premier sous-ensemble (`<titre niveau="poeme">`). Nous avons également fait figurer, quand elles existent, les informations supplémentaires qui permettent d'individualiser chaque texte, comme les dates de rédaction (`<poeme anneeredaction="1906">`), de publication (`<poeme datepublication="1912">`) et d'édition des œuvres (`<poeme dateedition="1914">`), ainsi que les dédicaces et épigraphes (`epigraphe="à René Dalize"`). Enfin, nous avons restitué aux textes – spécialement aux poèmes – leur structuration de l'unité la plus grande à la plus réduite, en proposant les étiquettes suivantes : sous-titre (`<soustitre titre="Aubade chantée à Laetare l'an passé">`) ; partie (`<partie titreRecueil="Le poète assassiné" titre="I. Renommée">`) ; section (`<section titre="Stavelot 1899">`) ; corps du texte (`<texte/>`) ; paragraphe ou strophe (`<paragraphe/>` | `<strophe/>`) ; vers (`<vers/>`).

L'exemple suivant donne une idée de la hiérarchisation des informations permise par la technologie XML. Il est extrait du sous-corpus des œuvres poétiques d'Apollinaire :

```
<ouvrage type="poesieApollinaire">
  <recueil titre="Le Bestiaire ou Cortège d'Orphée" epigraphe="à Élémer Bourges">
    <titre niveau="recueil">Le Bestiaire ou Cortège d'Orphée</titre>
    <poeme recueil="BEST" titre="Orphée">
      <titre niveau="poeme">Orphée</titre>
      <texte>
        <strophe>
          <vers>Admirez le pouvoir insigne</vers>
          <vers>Et la noblesse de la ligne :</vers> [...]
        </strophe>
      </texte>
    </poeme>
```

L'exemple suivant est un aperçu du sous-corpus des œuvres en prose, qui offre évidemment quelques variations de structure :

```
<ouvrage type="proseApollinaire">
  <recueil name="L'HÉRÉSARQUE ET CIE" date="1910">
    <titre niveau="recueil">L'HÉRÉSARQUE ET CIE</titre>
    <texte titre="LE PASSANT DE PRAGUE" epigraphe="À Thadée Natanson ces philtres de
phantase.">
      <titre niveau="texte">LE PASSANT DE PRAGUE</titre>
      <section>
        <paragraphe>En mars 1902, je fus à Prague.</paragraphe>
        <paragraphe>J'arrivais de Dresde.</paragraphe>
        <paragraphe>Dès Bodenbach, où sont les douanes autrichiennes, les allures des
employés de chemin de fer m'avaient montré que la raideur allemande n'existe pas dans
l'empire des Habsbourg.</paragraphe> [...]
      </section>
    </texte> [...]
  </recueil>
```

La technologie XML permet donc, comme on le voit ici, d'offrir une structure hiérarchique aux corpus annotés. Il devient désormais possible d'analyser non plus seulement les fréquences des occurrences, mais également leur rôle et leur représentativité au sein de chacun de ces sous-ensembles que sont la section, la strophe ou encore le vers. Le balisage XML permet une certaine exhaustivité dans l'annotation de corpus : chaque texte peut dès lors être étudié dans une perspective générale et, sans que cela soit nécessairement contradictoire, individuellement grâce aux informations précises données par les différentes étiquettes, décrites ci-dessus.

### 1.3 L'apport du corpus annoté sous TXM à la méthodologie d'analyse stylistique

La linguistique de corpus n'a rien de nouveau, loin s'en faut ; néanmoins, les corpus annotés nous semblent renouveler la méthode de recherche en stylistique, car, outre le fait qu'ils permettent d'élargir considérablement l'échelle de l'analyse, ils offrent l'avantage de systématiser les remarques linguistiques et ouvrent de nouvelles perspectives pour les études de stylistique. Comme V. Magri-Mourgues (2006) le souligne fort justement, le perfectionnement des outils informatiques d'analyse textuelle invitent à s'interroger sur les enjeux épistémologiques du corpus stylistique :

L'interdépendance entre style et corpus est telle que l'on ne peut se définir sans l'autre, ou plutôt que la définition de l'un entraîne corollairement une évolution de l'autre. La variation du corpus d'étude induit des pratiques stylistiques différentes. Le corpus est un objet empirique et structuré selon les enjeux et les objectifs de la recherche. Il est par conséquent toujours contingent, déterminé par l'application que l'on veut en faire. [...] le chercheur oscille entre deux tendances complémentaires : une démarche déductive, lorsqu'une thèse préalable préside à l'établissement de ce corpus et une démarche inductive quand c'est l'observation de spécificités langagières qui sous-tend l'élaboration d'une théorie. (§ 2)

Il est également important de remarquer que travailler des corpus littéraires au moyen de logiciels de statistiques textuelles contribue à porter un regard différent sur des notions stylistiques canoniques, ou jugées obsolètes. L'idée d'« écart » stylistique acquiert par exemple une nouvelle pertinence, car le style ne se définit plus désormais comme une « catégorie stable », mais comme « le mouvement même de l'œuvre » (cf. A. Herschberg-Pierrot : 2006). Le style a une historicité, une « singularité ». Comme l'écrit V. Magri-Mourgues (2006 : § 7) :

La stylistique s'éloigne de la caractérisation du style d'un individu ou d'un auteur mais s'affirme comme l'étude de l'œuvre entendu comme un « continué » poursuivi depuis sa genèse jusqu'à l'acte de lecture. [...] C'est le style du corpus mouvant, avec ses manques, ses variations, qui est étudié et qui ne se confond nullement alors avec celui de l'écrivain.

La notion de « stylème » ou de « fait de langue » trouve un écho singulier dans l'analyse statistique des textes, en particulier lorsque l'on analyse ce qu'on appelle des « expressions régulières », mises en évidence par des listes de fréquences, calculées de façon systématique.

On peut alors rencontrer la notion de style d'auteur, qui n'est pas le style d'un individu. Ce style n'a pas d'existence hors de la ressemblance de famille des œuvres, qui se construit par estompement des spécificités des œuvres singulières, mais aussi comme un processus comprenant des évolutions, de ruptures, des polarités et des convergences d'une œuvre à l'autre. (A. Herschberg-Pierrot : 2006)

Les corpus annotés ont l'avantage de proposer également une grande souplesse pour le travail d'analyse. Dans le cas des œuvres d'Apollinaire, le logiciel TXM permet d'examiner l'ensemble des textes pêle-mêle, ou de façon plus structurée (grâce aux étiquettes que nous avons décrites, par exemple). Il offre également, comme Lexico3, la possibilité de décomposer le corpus en sous-corpus (poésie et prose) d'opérer des partitions au sein même de ces ensembles (comme une division du sous-corpus « poésie » en recueils poétiques). Ce qu'apporte TXM au chercheur en stylistique, c'est la possibilité de choisir le point de vue adopté, et, de fait, de varier les points de comparaison. C'est ainsi que l'on peut étudier le corpus apollinarien sous l'angle générique, mais aussi, à partir du même fichier, sous l'angle chronologique. Dans le premier cas, le critère de sélection des sous-ensembles sera le genre d'ouvrage étudié ; dans le second, la constitution de segmentations chronologiques des textes.

L'intérêt de ce partitionnement du corpus apollinarien, dans la perspective qui est la nôtre, est non seulement de comparer les deux grands ensembles génériques distincts – l'œuvre en prose et l'œuvre poétique d'Apollinaire (partition sanctionnée par les titres de l'édition Gallimard « Bibliothèque de la Pléiade ») –, mais aussi d'analyser différents recueils de poèmes ou de contes entre eux. Ce qui nous a paru important, c'est que l'on peut ainsi redynamiser l'opposition formelle et canonique de la prose et de la poésie en ne considérant plus la poésie comme une forme de « subversion des codes »<sup>16</sup> de la langue prosaïque. De plus, les études stylistiques menées par des spécialistes d'Apollinaire sont rares et se limitent généralement à l'analyse des formes du recueil *Alcools*<sup>17</sup>. Il est désormais possible, grâce au corpus annoté que nous présentons ici d'analyser l'ensemble des recueils poétiques d'Apollinaire et de pouvoir désormais élargir certaines remarques à la totalité de sa production poétique, exception faite de quelques poèmes encore inédits ou de quelques textes dont la transcription est toujours très délicate<sup>18</sup>.

## 2 Continu et discontinu dans l'écriture apollinarienne : rôle et distribution des conjonctions

Comment, dès lors, analyser l'articulation du *continu* et du *discontinu* chez Apollinaire ? Reprenons le commentaire métapoétique qu'Apollinaire adressait à Madeleine : « J'ai fait mon possible pour simplifier la syntaxe poétique et y ai réussi en certains cas, notamment un poème : les 'Fenêtres' ». Il poursuit en s'efforçant d'explicitier en quoi consiste cette « simplification syntaxique » : « [...] s'il s'agit de vitesse, de raccourci, le style télégraphique nous offre des ressources auxquelles l'ellipse donnera une force et une saveur merveilleusement lyriques »<sup>19</sup>. C'est surtout les mentions de « style télégraphique » et d'« ellipse » qui retiendront notre attention : le premier caractérise en effet les énoncés dont la syntaxe est délibérément réduite à sa trame la plus tenue, abandonnant souvent le groupe verbal, la ponctuation et privilégiant les notations resserrées sur les groupes nominaux ; la seconde identifie, sur le plan sémantique et syntaxique, le raccourci lui-même, la juxtaposition – inattendue ? – des segments de phrase.

Outre « Les Fenêtres »<sup>20</sup>, poème-conversation d'un genre nouveau publié pour la première fois en ouverture du catalogue de l'exposition de Robert Delaunay en 1913, un certain nombre de poèmes de *Calligrammes*, en particulier, correspondent à ce descriptif, à l'instar de « Lettre-Océan », premier poème calligrammatique donnant à voir plusieurs vues de la Tour Eiffel, grand émetteur-récepteur de la TSF. Bribes de conversations, câbles et divers messages s'y entremêlent pour en composer la figure étoilée. On pourrait citer le poème « Arbre »<sup>21</sup> qui met lui aussi ponctuellement en œuvre le principe de l'écriture discontinue, « simplifiée » que nous mentionnions précédemment :



[...]  
Un enfant  
Un veau dépouillé pendu à l'étal  
Un enfant  
Et cette banlieue de sable autour d'une pauvre ville au fond de l'est  
Un douanier se tenait là comme un ange  
À la porte d'un misérable paradis  
Et ce voyageur épileptique écumait dans la salle d'attente des premières  
Engoulevent Blaireau  
Et la Taupe-Ariane  
[...]

Cette énumération surprenante se construit comme on le voit par une juxtaposition d'énoncés apparemment sans lien. La trame décousue du texte est néanmoins structurée par des effets de répétitions plus ou moins affirmés – celle de « un enfant » qui encadre le second vers de l'extrait, celle du déterminant indéfini « un » qui actualise presque tous les groupes nominaux et fait ressortir de façon d'autant plus forte le démonstratif « ce voyageur épileptique »... L'ellipse va de pair avec le discontinu syntaxique : Apollinaire privilégie d'ailleurs les phrases averbales<sup>22</sup>, ainsi que des propositions indépendantes (« Un douanier se tenait là comme un ange / À la porte d'un misérable paradis », « Et ce voyageur épileptique écumait dans la salle d'attente des premières »). De plus, on constate que les mots de liaison sont très largement absents d'un tel énoncé : point de conjonction par exemple, outre « et », présente trois fois. Son rôle semble essentiellement rythmique, tandis que sa valeur sémantique est gommée. De fait, la « simplification syntaxique » mentionnée par Apollinaire à propos des « Fenêtres » nous semble donc, au vu de cet exemple caractéristique, correspondre à une réduction de l'énoncé à des phrases réduites au maximum, et à une disparition des marqueurs de lien syntaxique au profit de l'ellipse et du discontinu.

Afin de préciser notre investigation sur les ressorts de l'articulation entre *continu* et *discontinu*, nous nous sommes concentrés sur l'analyse quantitative d'une partie du langage caractéristique : les conjonctions. Ces dernières présentent l'avantage d'être représentées tant dans le sous-corpus poétique que dans le sous-corpus prosaïque. Elles se définissent également comme les marqueurs de liaison syntaxique par excellence.

## 2.1 Distribution comparée des conjonctions entre prose et poésie

À partir de la méthodologie que nous avons décrite ci-dessus, nous nous sommes attachés à quantifier et à comparer les usages et les variations d'utilisation des conjonctions dans le corpus apollinarien. Il s'agit d'analyser dans quelle mesure les conjonctions tendent à disparaître du corpus poétique d'Apollinaire, en comparaison avec ses recueils de conte.

Après un marquage morphosyntaxique du corpus apollinarien, exécuté par le logiciel *Treetagger*, TXM a pu générer le tableau suivant (voir ci-dessous). Ces résultats ont été obtenus par une recherche menée sur les différentes catégories du discours (ou *part of speech*) présentes dans l'ensemble du corpus d'Apollinaire<sup>23</sup>. Nous nous sommes concentrés sur le cas précis des conjonctions (que l'on retrouve dans TXM sous l'étiquette morphosyntaxique ["kon"]), en proposant la requête suivante : <[ttpos="kon"]>. Ce tableau présente donc le résultat global de la répartition des conjonctions, sans les différencier dans un premier temps. Pour établir les spécificités de la distribution des conjonctions, nous avons réalisé des partitions par recueils dans le corpus complet (voir la colonne de gauche). Une fois ces partitions opérées, il est alors possible de calculer précisément comment se répartissent les conjonctions en prose et en poésie et de présenter les spécificités propres à chaque recueil d'œuvres poétiques et prosaïques.

		PARTITION "RECUEIL"	FRÉQUENCE ABSOLUE	SPÉCIFICITÉS
Prose		<b>L'ENCHANTEUR ...</b>	1012	<b>+ 4</b>
		<b>L'HÉRÉSARQUE ET CIE</b>	2522	<b>- 13</b>
		<b>LE POÈTE ASSASSINÉ</b>	2343	<b>- 5</b>
Poésie		LE BESTIAIRE	43	0
		<b>ALCOOLS</b>	991	<b>+ 3</b>
		VITAM IMPENDERE...	23	0
		<b>POÈMES À LOU</b>	1015	<b>+ 2</b>
		POÈMES À MADELEINE	255	+ 1
		POÈMES À LA MARRAINE	18	+ 1
		<b>CALLIGRAMMES</b>	1017	<b>+ 3</b>
		<b>LE GUETTEUR ...</b>	654	<b>+ 4</b>
		IL Y A	450	+ 1
		POÈMES INÉDITS	148	0
		POÈMES RETROUVÉS	502	0
	TOTAL			10 993

Le calcul des spécificités (cf. Lafon et Muller, 1984) permet de déterminer, grâce à plusieurs formules opérant un constant va et vient entre le tout (le corpus complet) et la partie (chacun des sous-ensemble obtenu par la partition du corpus), le degré de sur- ou de sous-représentation d'une unité (ici, les conjonctions) dans un sous-corpus (les différents recueils d'Apollinaire) dans un rapport de proportionnalité relative<sup>24</sup>. Dès le premier coup d'œil, on remarque que les indices de sur ou sous-représentation stigmatise de façon criante la partition générique entre prose et poésie, à l'exception notable de *L'Enchanteur pourrissant*<sup>25</sup>.

Contrairement à ce que l'on pouvait attendre, les conjonctions (toutes formes confondues) sont sous-représentées dans les recueils de contes prosaïques et généralement sur-représentées en poésie. Un certain nombre de recueils (*Le Bestiaire*, *Vitam Impendere amori*, *Les Poèmes Inédits* et *les Poèmes retrouvés*, ainsi que, dans une moindre mesure, *Il y a*, *Poèmes à Madeleine* et *Poèmes à la Marraïne de guerre*) sont dans une moyenne d'emploi : leur écart-type positif étant inférieur à 2<sup>26</sup>, les résultats qu'ils offrent sont délicats à prendre en considération. Quatre grands recueils se détachent du lot – *Alcools*, *Calligrammes*, *Poèmes à Lou* et *Le Guetteur mélancolique* – et révèlent un usage des conjonctions supérieur à la moyenne<sup>27</sup>. D'emblée, ces résultats vont donc à l'encontre de la « simplicité syntaxique » que l'on évoquait précédemment : il semblerait, au contraire, que le corpus poétique apollinarien en général, au regard des deux grands recueils de conte, *L'Hérésiarque et cie* et *Le Poète assassiné*, se caractérisent par une présence accrue des conjonctions dans l'écriture.

Précisons ces résultats. Pour ce faire, nous avons distingué, en reprenant manuellement l'étiquetage morphosyntaxique de notre corpus effectué par *Treetagger*, les deux grands types de conjonctions : les conjonctions de coordination (["kon:coord"]) et les conjonctions de subordination (["kon:sub"]). Dans le tableau présenté ci-dessous, nous matérialisons la représentativité de chacune de ces sous-catégories, en donnant leur fréquence et leur indice de spécificité respectifs. Des remarques s'imposent immédiatement : on constate en effet que les résultats isolent à nouveau les deux grands recueils de conte, *L'Hérésiarque et cie* et *Le Poète assassiné*. Ces derniers présentent une sous-représentation manifeste des deux types de conjonctions, avec un taux plus élevé en ce qui concerne les conjonctions de coordination. *L'Enchanteur pourrissant* s'associe, quant à lui, avec deux recueils poétiques – *Alcools*, *Le Guetteur mélancolique* – par sa sur-représentation significative des conjonctions de coordination. *Calligrammes* et *Poèmes à Lou* (dont l'écriture respective est, dans une large part, contemporaine) révèlent en revanche une sur-représentation des conjonctions de subordination.

		CONJONCTIONS DE COORDINATION		CONJONCTIONS DE SUBORDINATION	
		FREQUENCE	INDICE	FREQUENCE	INDICE
Prose	<b>L'ENCHANTEUR...</b>	633	+ 3	379	+ 1
	<b>L'HÉRÉSARQUE...</b>	1566	- 8	956	- 6
	<b>LE POÈTE ASSASSINÉ</b>	1438	- 4	905	- 2
Poésie	LE BESTIAIRE	27	0	16	0
	<b>ALCOOLS</b>	630	+ 3	361	0
	VITAM IMPENDERE...	17	0	6	0
	<b>POÈMES À LOU</b>	555	0	460	+ 7
	POÈMES À MADELEINE	147	0	108	+ 1
	POÈMES À LA MARR...	9	0	9	+ 1
	<b>CALLIGRAMMES</b>	606	+ 1	411	+ 3
	<b>LE GUETTEUR...</b>	437	+ 6	217	0
	<b>IL Y A</b>	301	+ 2	149	0
	POÈMES INÉDITS	96	0	52	0
	POÈMES RETROUVÉS	336	0	166	- 1
		6798		4195	

Comment expliquer cette dernière remarque ? N'y a-t-il pas là un paradoxe d'autant plus grand que c'était dans les poèmes de *Calligrammes* que l'on rencontrait les exemples les plus frappants de l'usage du « style télégraphique » et de « l'ellipse » dont parlait Apollinaire ? Pourquoi trouve-t-on dans *Calligrammes*, dont l'esthétique, les images et les motifs semblent tout entier dévolus à l'impression de discontinuité, un si grand nombre de mots de liaison symptomatiques d'une certaine complexité syntaxique ?

Cette dernière constatation est d'autant plus paradoxale que c'est dans les poèmes de *Calligrammes* que l'on trouvait la plupart des exemples de simplification syntaxique dont nous parlions précédemment. Quelles sont précisément les conjonctions employées ? Comment expliquer ce phénomène ? Pour préciser davantage notre champ d'investigation, nous avons listé un certain nombre de conjonctions très fréquentes dans le corpus apollinarien pour pouvoir les quantifier précisément et apporter un éclairage sur le paradoxe que nous avons mis au jour.

Pour éviter les erreurs et les oublis provoqués par la casse, nous avons privilégié dans notre requête les lemmes (["tlemma"]) pour une recherche de « lemma » dans *TreeTagger*. Nous nous sommes d'abord intéressés aux conjonctions de subordination dites *simples* : « quand », « lorsque », « puisque » et « comme ». La requête pour les quantifier sur TXM est la suivante : [tlemma="quand"]| [tlemma="lorsque"]| [tlemma="puisque"]| [tlemma="comme"]. Grâce au calcul de spécificités, plus parlant et surtout plus probant que la fréquence relative et absolue, nous pouvons générer le tableau suivant :

	L'ENCHANTEUR	L'HERESARQUE	LE POETE ASSASSINE	LE BESTIAIRE	ALCOOLS	P. A LOU	P. A MADELEINE	CALLIGRAMMES	LE GUETTEUR	IL Y A	P. A MARRAINE	P. INEDITS	P. RETROUVES	VITAM
<b>quand</b>	0	- 11	- 1	0	+ 2	+ 1	0	+ 1	+ 5	+ 2	0	+ 1	0	0
<b>lorsque</b>	0	0	+ 4	0	- 1	0	- 1	- 2	0	0	0	0	- 1	0
<b>puisque</b>	+ 1	- 3	- 1	0	0	+ 2	+ 1	0	0	0	+ 1	0	0	0
<b>comme</b>	- 1	- 10	- 6	0	+ 3	+ 12	+ 4	+ 5	0	0	0	0	- 1	0

Nous avons ensuite quantifié un certain nombre de conjonctions de subordination composées : (« tandis que », « parce que », « pour que », « après que », « depuis que », « sans que », « pendant que », « alors que »<sup>28</sup>), grâce à la requête suivante :

[tlemma="tandis|parce|pour|après|depuis|sans|pendant|alors"] [word="qu['e']"]

Nous rendons compte dans le tableau ci-dessous de leur répartition par indice de spécificité :

	L'ENCHANTEUR	L'HERESIAIQUE	LE POETE ASSASSINE	LE BESTIAIRE	ALCOOLS	P. A LOU	P. A MADELEINE	CALLIGRAMMES	LE GUETTEUR	IL Y A	P. A MARRAINE	P. INEDITS	P. RETROUVES	VITAM
tandis que	0	0	+ 1	0	0	0	0	0	0	0	0	0	0	0
<b>parce que</b>	<b>+ 5</b>	0	- 1	0	<b>+ 3</b>	<b>- 2</b>	0	0	0	0	+ 1	0	0	0
pour que	0	0	0	+ 1	0	0	0	0	- 1	0	+ 1	0	0	0
après que	0	0	0	<b>+ 2</b>	0	0	+ 1	0	+ 1	+ 1	<b>+ 2</b>	+ 1	+ 1	<b>+ 2</b>
depuis que	0	+ 1	0	+ 1	0	0	+ 1	0	0	0	<b>+ 2</b>	+ 1	0	<b>+ 2</b>

Enfin, nous avons calculé le nombre d'occurrences des conjonctions de coordination, grâce à la requête suivante : [ ttlema="mais|ou|et|or|ni|car" ] et avons à nouveau calculé leur spécificité pour chaque recueil du corpus apollinarien :

	L'ENCHANTEUR	L'HERESIAIQUE	LE POETE ASSASSINE	LE BESTIAIRE	ALCOOLS	P. A LOU	P. A MADELEINE	CALLIGRAMMES	LE GUETTEUR	IL Y A	P. A MARRAINE	P. INEDITS	P. RETROUVES	VITAM
<b>mais</b>	<b>+ 2</b>	1	0	0	<b>- 5</b>	- 1	<b>- 2</b>	+ 1	0	0	0	0	- 1	0
<b>ou</b>	- 1	0	<b>- 3</b>	0	<b>+ 2</b>	- 1	0	<b>+ 3</b>	0	0	0	0	0	0
<b>et</b>	0	<b>- 8</b>	<b>- 3</b>	0	<b>+ 6</b>	0	0	0	<b>+ 5</b>	+ 1	0	0	+ 1	0
<b>or</b>	0	- 1	<b>- 3</b>	0	0	0	- 1	0	<b>+ 5</b>	+ 1	0	<b>+ 2</b>	<b>+ 2</b>	0
<b>ni</b>	<b>+ 3</b>	<b>- 3</b>	0	0	<b>+ 3</b>	<b>- 3</b>	0	0	+ 1	+ 1	0	0	0	0
<b>car</b>	<b>+ 8</b>	0	0	0	<b>- 2</b>	0	0	- 1	0	0	0	0	0	0

Les résultats de cette investigation menée au sein des œuvres poétiques d'Apollinaire sont significatifs. Ce qui est à nouveau remarquable ici, c'est la grande disparité dans les indices de représentation qui mettent au jour des différences d'écriture au sein des recueils. Prenons l'exemple de « et », sous-représenté en prose, il ne se distingue que dans deux recueils – *Alcools* et *Le Guetteur mélancolique* – par une présence très marquée. De façon assez symptomatique également, on remarque que dans le corpus apollinarien, la conjonction « or » est généralement préférée à « mais » (qui demeure dans une moyenne d'emplois, voire est largement sous-représenté dans *Alcools*), dans trois recueils. Entre ici en jeu la valeur sonore du mot et son homonymie avec le métal, la couleur « or », chers au lexique apollinarien.

De même, on constate que la plupart des conjonctions de subordination composées sont généralement dans une moyenne d'emplois, exception faite de « parce que » qui sort du lot avec une sur-représentation dans *Alcools* – de même que « puisque » dans les *Poèmes à Lou*. Certes, il faut apporter quelques réserves à ce type de répertoire de données et le confronter avec circonspection au relevé précis des occurrences dans le texte. Dans le cas de « parce que », par exemple, qui n'a pas, de prime abord, une connotation « poétique » très évidente, on s'aperçoit que sur les 35 mentions que présente le recueil *Alcools*, 11 d'entre elles ne figurent que dans un seul et même poème : le « Poème lu au Mariage d'André Salmon »<sup>29</sup> :

[...]

Réjouissons-nous non pas **parce que** notre amitié a été le fleuve qui nous a fertilisés  
 Terrains riverains dont l'abondance est la nourriture que tous espèrent  
 Ni **parce que** nos verres nous jettent encore une fois le regard d'Orphée mourant  
 Ni **parce que** nous avons tant grandi que beaucoup pourraient confondre nos yeux et  
 les étoiles  
 Ni **parce que** les drapeaux claquent aux fenêtres des citoyens qui sont contents depuis  
 cent ans d'avoir la vie et de menues choses à défendre  
 Ni **parce que** fondés en poésie nous avons des droits sur les paroles qui forment et  
 défont l'Univers

Ni **parce que** nous pouvons pleurer sans ridicule et que nous savons rire  
Ni **parce que** nous fumons et buvons comme autrefois  
Réjouissons-nous **parce que** directeur du feu et des poètes  
L'amour qui emplît ainsi que la lumière  
Tout le solide espace entre les étoiles et les planètes  
L'amour veut qu'aujourd'hui mon ami André Salmon se marie [...]

On peut faire la même remarque pour *Calligrammes* qui regroupe 4 occurrences des 5 « parce que » du recueil dans la seule ode « À l'Italie » :

[...]  
Je t'envoie mes amitiés Italie et m'apprête à applaudir aux hauts faits de ta bleusaille  
Non **parce que** j'imagine qu'il y aura jamais plus de bonheur ou de malheur en ce monde  
Mais **parce que** comme toi j'aime à penser seul et que les Boches m'en empêcheraient  
Mais **parce que** le goût naturel de la perfection que nous avons l'un et l'autre si on les laissait faire serait vite remplacé par je ne sais quelles commodités dont je n'ai que faire  
Et surtout **parce que** comme toi je sais je veux choisir et qu'eux voudraient nous forcer à ne plus choisir  
Une même destinée nous lie en cette occase [...]

Ce qui caractérise ces textes – en marge de la production poétique apollinienne – est la déclinaison de scènes et d'images qui détournent et déjouent une articulation syntaxique complexe et répétitive (le procédé valorisé dans ces deux poèmes est vraiment l'anaphore des conjonctions) qu'on aurait tendance à taxer de « lourdeur ». Le projet discursif de chacun de ces deux poèmes se distingue ici de la plupart des autres textes : le premier, hommage à un ami, est un poème de circonstance, le second, ode à l'Italie entrée en guerre, entre dans la tradition du poème épideictique. Ces poèmes mettent à l'honneur ces conjonctions de façon particulièrement saillante, pour traduire généralement l'enthousiasme du locuteur.

En revanche, les conjonctions de subordination simples attirent bel et bien l'œil en révélant de grandes disparités entre les recueils. Le calcul de spécificité fait notamment ressortir la sur-représentation de la conjonction « comme », attendue en poésie puisque c'est elle qui introduit la majorité des structures comparatives. C'est elle qui explique, en partie, l'importante distribution des conjonctions dans *Calligrammes* et les *Poèmes à Lou*, que nous avons soulignée. En effet, si l'on isole « comme » de l'ensemble des conjonctions de subordination, on obtient les spécificités suivantes :

	L'ENCHANTEUR	L'HERESIAIQUE	LE POETE ASSASSINE	LE BESTIAIRE	ALCOOLS	P. A LOU	P. A MADELEINE	CALLIGRAMMES	LE GUETTEUR	IL Y A	P. A MARRAINE	P. INEDITS	P. RETROUVES	VITAM
comme	0	- 8	- 3	0	+ 1	+ 11	+ 5	+ 3	0	0	0	0	0	0
[ttpos="KON:sub"]	+ 2	- 2	0	0	0	+ 2	0	+ 1	0	0	+ 1	0	0	0

La sur-représentation des conjonctions de subordination s'amenuise dans la plupart des cas – à l'exception des *Poèmes à Lou* – en sorte que l'on peut aisément affirmer la prépondérance de « comme » dans la distribution des conjonctions en poésie, mais aussi en prose. De fait, on constate que la sous-représentation des conjonctions diminue également en prose, lorsque l'on isole « comme ».

Autre différence significative : la répartition des conjonctions temporelles « quand » et « lorsque » en prose et en poésie.

## 2.2 L'exemple des conjonctions temporelles « quand » et « lorsque »

Si l'on calcule l'écart-type en opposant, de façon un peu schématique, les deux sous-corpus génériques, cette disparité est très frappante :

	fréquence	Sous-corpus "PROSE"	Sous-corpus "POESIE"
[tlemma= "quand"]	270	- 15	+ 15
[tlemma= "lorsque"]	123	+ 5	- 5

On s'aperçoit que la distribution s'inverse : largement sous-représentée en prose, la conjonction « quand » semble appartenir à ce qu'on pourrait appeler un « lexique poétique » d'Apollinaire, tandis que « lorsque » s'illustre en prose.

Les résultats s'affinent quand on considère le détail de chacun des recueils : « quand » est, en effet, sous-représenté dans *L'Hérésiarque et cie* (l'écart-type est de -11), mais sur-représentée de façon significative dans près de trois recueils poétiques : *Alcools, Il y a* (+2) et *Le Guetteur mélancolique* (+5). La conjonction « lorsque », si elle entre dans une moyenne d'usage – faible quantitativement : on compte seulement 123 occurrences dans tout le corpus apollinarien – est sur-représentée dans *Le Poète assassiné* (+4) et sous-représentée dans *Calligrammes* (-2). Une telle répartition est-elle signifiante ? Est-elle motivée sémantiquement ?

« Lorsque », dans le corpus poétique d'Apollinaire, apparaît dans le cadre très précis de la narration d'un fait isolé. La focalisation se fait alors, selon l'usage normal de la conjonction, sur la circonstance précise de l'événement raconté, que l'on met en valeur. L'emploi poétique de « quand » est, quant à lui, plus complexe. En effet, si l'on examine précisément chaque occurrence du mot, on remarque que cette conjonction présente ce que nous appellerions un « feuilleté » sémantique, une épaisseur du sens. Les différentes nuances du mot sont les suivantes :

- La focalisation sur l'instant précis où se passe l'action (« quand » = « au moment où »)
- La focalisation sur la récurrence cyclique de l'action (« quand » = « toutes les fois que »)

et plus marginales :

- L'opposition (« quand » = alors que)
- La focalisation sur l'antériorité de l'action (« quand » = « après que »)

Les deux premières nuances sont les plus représentées, mais ce qui importe, c'est finalement les cas limites, les moments d'indécision du sens, très fréquents dans le corpus poétique apollinarien, et dont on peut donner quelques exemples :

### 2.2.1 Concurrence entre l'aspect ponctuel et l'aspect cyclique du processus décrit :

Automne malade et adoré / Tu mourras **quand** l'ouragan soufflera dans les roseraies /  
**Quand** il aura neigé / Dans les vergers

(« Automne malade », *Alcools*)<sup>30</sup>

On serait plutôt tenté de gloser par « Tu mourras **au moment où** l'ouragan soufflera dans les roseraies... », mais le déroulement des saisons, l'aspect cyclique du temps – que l'on retrouve dans tant d'autres poèmes d'Apollinaire – provoque une indécision.

### 2.2.2 Le rôle ambigu des temps verbaux

Cortèges ô cortèges / C'est **quand** jadis le roi s'en allait à Vincennes / **Quand** les  
ambassadeurs arrivaient à Paris / **Quand** le maigre Suger se hâtait vers la Seine /  
**Quand** l'émeute mourait autour de Saint-Merry

(« Le Musicien de Saint-Merry », *Calligrammes*)<sup>31</sup>

Cet exemple est particulièrement intéressant pour décrire l'épaisseur sémantique de « quand ». Dans le premier vers : « C'est quand jadis le roi s'en allait à Vincennes », l'emploi de l'imparfait (itératif ?) « s'en allait », associé au présent relativement a-temporel de la principale réduite à son strict minimum « c'est »,

incite à interpréter « quand » comme un équivalent de « toutes les fois que ». Même analyse pour le second vers « Quand les ambassadeurs arrivaient à Paris ».

Or, le reste de l'énumération, structurée par l'anaphore en « quand », décrit des faits historiques précis, datables et identifiables, infirmant donc la première lecture que l'on fait de la conjonction et l'assimilant à « au moment où ». On se trouve donc ici face à un véritable télescopage temporel, une ambiguïté temporelle qui s'ajoute à l'ubiquité du poète qui, dans ce même poème, note et restitue les détails d'actions anecdotiques qui ont lieu simultanément dans divers endroits.

Il y aurait beaucoup d'autres exemples de ce type (nous estimons qu'ils constituent environ 20% des occurrences de ces conjonctions temporelles). Nous pouvons en tout cas avancer que, généralement, les occurrences de « quand » mettent en lumière, dans le corpus poétique, la prédilection d'Apollinaire pour l'expression de l'instantanéité. De plus, la valeur répétitive de « quand » permet une variation polysémique, de vers à vers, entre l'anecdotique et la récursivité.

Il est également évident que l'usage qu'Apollinaire fait respectivement de « quand » et de « lorsque » est aussi motivé par la prosodie et les effets d'anaphore, de rimes et d'assonances que chacun de ces vocables régit. Sans pouvoir entrer ici dans le détail de l'analyse, nous soulignerons cependant que « quand », monosyllabique et à la sonorité nasale trouvera facilement dans ce que l'on appellera le « système phonique » apollinarien des échos particulièrement fructueux. Un examen même rapide des cooccurrences du lemme « quand » (soit [ttlemma="quand"]) nous en convaincra aisément. En effet, outre les effets d'anaphores et de reprises en échos de cette même conjonction « quand », au sein de la strophe, on remarque la récurrence de mots en [ã] tels que « enterrements », « anges », « blancs », « amant », « danser », « souvent », « chant », « temps », etc., ainsi que les formes verbales de participes présents. De plus, il s'agit de prendre en compte les allitérations en [k], elles aussi nombreuses. À l'inverse, le même examen est beaucoup moins riche pour « lorsque », qui présente peu de cooccurrences à première vue significatives pour l'analyse phonique des phrases. Bien entendu, cet aspect mériterait un traitement particulier, et si possible, le recours à des techniques de transcriptions phonétiques des textes poétiques que nous ne pouvons développer ici, et que décrivent très bien J.-M. Viprey (2002) et M. Bernard (2006). Quoiqu'il en soit, cette indécision du sens, cette incertitude sur la valeur du temps dans les poèmes d'Apollinaire, synthétisé par les emplois de « quand » met en lumière une tension entre l'anecdotique pur et le « floutage » du référent, et peut-être plus largement l'ambition de créer un nouveau discours lyrique. La poésie d'Apollinaire se caractérise en effet, d'après l'analyse de ces quelques exemples, par un processus de densification du langage, y compris dans l'usage qui est fait de certaines conjonctions et formes de lien syntaxique.

### 3 Conclusion

Pour déterminer les caractéristiques de l'expression du « *discontinu* » dans le langage poétique d'Apollinaire, nous avons proposé une analyse comparée, grâce aux techniques d'annotation de corpus, des œuvres en prose et des recueils poétiques apollinariens. La question de l'articulation entre *continu* et *discontinu* dans le discours d'Apollinaire est complexe et met en lumière un paradoxe. En poésie, où le langage tend à se densifier, à privilégier l'ellipse productrice de merveilleux, où le *lien* sémantique entre les vers se crypte parfois, appelant le lecteur à son élucidation<sup>32</sup>. Si la « simplification de la syntaxe poétique » apparaît de façon évidente dans certains poèmes d'Apollinaire privilégiant un rétrécissement de l'énoncé en bribes de mots, en groupes nominaux et en construction de phrases simples, d'autres jouent parfois sur la démultiplication du lien syntaxique.

En effet, le sous-corpus poétique compte en proportion davantage de conjonctions que le sous-corpus des œuvres en prose, toutefois l'analyse que nous avons pu faire de la répartition de cette catégorie grammaticale met au jour plusieurs remarques. D'abord, l'importance de « comme », qui infléchit les résultats vers une sur-représentation des conjonctions en poésie ; ensuite, le choix des conjonctions elles-mêmes qui ancre le discours, par ses possibilités sonores et graphiques, dans un genre précis. La brève étude de deux conjonctions temporelles « quand » et « lorsque » montre que le marqueur de lien syntaxique est utilisé à dessein par le poète. Enchaînées les unes à la suite des autres et exhibées dans leur

ambiguïté sémantique ou au contraire utilisées avec parcimonie, l'une et l'autre sont à nouveau le témoin du procédé de densification du discours poétique, mis en œuvre par Apollinaire. La *continuité* que créent ici les conjonctions, mise généralement au service d'une narration, est fréquemment minée de l'intérieur par un subtil glissement du sens qui met à mal la logique. S'il y a bien lien, ce dernier est loin d'être évident ou plutôt, est bien souvent faussement explicite. Dès lors, elle corrobore certains critères de la poétique apollinaire : l'alliance subtile de l'anecdotique et du général, de l'infiniment petit et de l'infiniment grand, ainsi que la réflexion sur les enchaînements logiques et le rôle poétique du hasard et de la surprise.

Les deux genres – prose et poésie – sont-ils à opposer ? Répondent-ils à des modalités d'écriture si différentes ? Grâce aux logiciels de statistique textuelle, nous avons essayé de souligner la différence d'utilisation et de fréquence des conjonctions en prose et en poésie. Le calcul de spécificité notamment permet de mettre en lumière des phénomènes langagiers au sein même de l'œuvre d'un auteur, et non plus seulement par rapport à une « norme » extérieure que l'on fixerait arbitrairement. Le floutage des frontières que n'a cessé d'opérer Apollinaire entre les deux genres – prose et poésie – met en relief les limites de cette opposition. Néanmoins, nous espérons avoir montré, à travers cette étude, qu'une analyse systématique de certains marqueurs du discours – les conjonctions – réactualise la question d'un style poétique et d'un style prosaïque chez Apollinaire.

### Références bibliographiques

- Adam, J.-M. and Viprey J.-M. (2009). « Corpus de textes, textes en corpus. Problématique et présentation », *Corpus*, 8 [mis en ligne en juillet 2010]. URL : <http://corpus.revues.org/index1672.html>
- Anis, J. (1983). « Vilisibilité du texte poétique », *Langue française*, 59/1, 1983, p. 88-102.
- Apollinaire, G. et al. (1994). *Œuvres poétiques*. Paris: Gallimard (« Bibliothèque de la Pléiade »).
- Apollinaire, G., and Campa, L. (2005). *Lettres à Madeleine. Tendre comme le souvenir*. Paris : Gallimard.
- Bernard, M. (2006). « Transcription phonétique des grands corpus littéraires. Les règles du jeu », *Corpus*, 5.
- Brunet, E. (1989). « Hyperbase : logiciel documentaire et statistique pour l'exploitation des grands corpus », *Tools for humanists*. Toronto, p. 33-36.
- Cendrars, B. (1912). *Les Pâques*. Paris : Éd. des Hommes nouveaux.
- Cendrars, B. (1913). *La Prose du Transsibérien et de la petite Jehanne de France*. Paris : Éd. des Hommes nouveaux.
- Chevalier, J.-Cl. (1970). *Étude stylistique de quelques poèmes de Guillaume Apollinaire : essai d'analyse des formes poétiques*. Abbeville : F. Paillart (éd).
- Florea, M.-L. (2009, nov.). « Tabularité : des textes aux corpus », *Corpus*, 8 [mis en ligne en mai 2011]. URL: <http://corpus.revues.org/index1792.html>
- Gardes-Tamine, J. and Monte, M. (2007). « Linguistique et poésie: état des lieux et perspectives », *Semen*, 24 [mis en ligne en février 2008]. URL: <http://semen.revues.org/6583>
- Garnier, P. (1963, janv.). « "Poèmes à dire" et "Poèmes à voir" » (avec le "Manifeste pour une poésie nouvelle, visuelle et phonique", 30 sept. 1962). *Les Lettres*, 29. André Silvaire (éd).
- Genette, G. (1987). *Seuils*. Paris : Seuil (« Poétique »).
- Groupe  $\mu$  (1992). *Traité du signe visuel*. Paris : Seuil.
- Groupe  $\mu$  (1997). « Temporalité et image fixe », *Texte*, 21-22, p. 41-70.
- Guiraud, P. (1953). *Index des mots d'Alcools de G. Apollinaire*. Paris: Klincksieck.
- Habert, B. [et al.] (1997). *Les Linguistiques de corpus*. Paris : A. Colin (« Collection U »).
- Herschberg-Pierrot, A. (2006, déc.). « Style, corpus et genèse », *Corpus*, 5 [mis en ligne en juillet 2007]. URL : <http://corpus.revues.org/index445.html>



- Lafon, P. et Muller, Ch. (1984). *Dépouillements et statistiques en lexicométrie*, Genève : Slatkine, Paris : Champion (« Travaux de linguistique quantitative »).
- Lapacherie, J.-G. (1982, avril). « Écriture et lecture du calligramme », *Poétique*, 50, p. 194-202.
- Lapacherie, J.-G. (1984, sept.). « De la grammatextualité », *Poétique*, 59, p. 283-294.
- Longree, G. H.F. (1981, juillet-oct.). « Du poème symboliste au calligramme, "Lettre-Océan" et le défolement du topologique », *Que Vlo-ve ?*, 1, 29-30. Dresat : Jambes, Belgique, p. 1-19.
- Magri-Mourgues, V. (2006, déc.). « Corpus et Stylistique », *Corpus*, 5 [mis en ligne en juillet 2007]. URL : <http://corpus.revues.org/index440.html>
- Mann, W.C., and Thompson, S.A. (1988). « Rhetorical Structure Theory : Toward a functional theory of text organization », *Text*, 8 (3), p. 243-281.
- Marinetti, F. T. (1914a). *Dune* (parole in liberté). Milano : Edizioni Futuriste di Poesia.
- Marinetti, F. T. (1914b). *Zang Tumb Tumb*, Milano : Edizioni Futuriste di Poesia.
- Mallarmé, S. (1897). « Un Coup de dés jamais n'abolira le hasard », *Cosmopolis*.
- Morhange-Bégué, Cl. (1983). « *La Chanson du mal-aimé* » d'Apollinaire. *Essai d'analyse structurale et stylistique*. Paris : Minard.
- Neveu, F. (dir.), (2006, mars). « Approches de la discontinuité syntaxique et énonciative ». *L'information Grammaticale*, 109.
- Noailly, M. (2002). « L'ajout après le point n'est-il qu'un simple artifice graphique ? », in Authier-Revuz, J. et Lala, M.-C. (2002). *Figures d'ajout : phrase, texte, écriture [colloque du 4 et 5 juin 1999 à la Sorbonne dans le cadre de l'Université Paris 3]*. Paris : Presses de la Sorbonne nouvelle, p.133-146.
- Phalèse, H. (de) (1996). *Quintessence d'« Alcools » : le recueil d'Apollinaire à travers les nouvelles technologies*. Paris : Nizet.
- Rastier, F. (2005). « Enjeux épistémologiques de la linguistique de corpus », in Williams, G. (éd.) *La Linguistique de corpus*. Rennes : Presses Universitaires de Rennes, p. 31-46.
- Segalen, V. (1912). *Stèles*. Pékin: Presses du Pei-t'ang.
- Viprey, J.-M. (2002). *Analyses textuelles et hypertextuelles des Fleurs du Mal*. Paris: Champion.

---

<sup>1</sup> Apollinaire, G. et Campa L. *Lettres à Madeleine. Tendre comme le souvenir*, Paris, Gallimard, 2005, p. 77.

<sup>2</sup> On a en effet souvent parlé de « collage » pour caractériser le poème « Les Fenêtres », notamment parce qu'il s'associe à l'entreprise artistique « orphique » des Delaunay en 1912. Ce dernier paraît d'ailleurs en janvier 1913 en préface du catalogue de l'exposition Delaunay à la galerie Der Sturm.

<sup>3</sup> « Simultanisme-Librettisme », in G. Apollinaire [et al.], *Œuvres en prose complètes II*, Paris, Gallimard (« Pléiade »), 1991, p. 976.

<sup>4</sup> Nous insistons sur le fait que nous n'utiliserons pas l'expression « discontinuité syntaxique » uniquement dans son sens usuel de structure de détachement, de dislocation ou d'extraction (cf. Neveu F.) mais choisissons d'élargir les perspectives en nous intéressant au phénomène de liaison syntaxique. Nous lui préférons, pour ce faire, le terme de « discontinu ».

<sup>5</sup> On citera par exemple l'ouvrage de P. Guiraud (1953) et Hubert de Phalèse (1996).

<sup>6</sup> L'AFC permet la visualisation des résultats sur un axe graphique, fréquemment sous la forme d'un nuage de points. Il s'agit d'un « tableau croisé dynamique » permettant de représenter matériellement les « différences » ou « écarts » entre les différentes valeurs étudiées.

<sup>7</sup> Voir le site du projet Textométrie de l'ENS de Lyon qui a créé le logiciel TXM. URL : <http://textometrie.ens-lyon.fr/>

---

<sup>8</sup> Nous avons laissé de côté l'œuvre théâtrales d'Apollinaire, ainsi qu'un roman inachevé, *La Femme assise*, que nous n'avons pas pu, à ce stade de notre travail, intégrer à l'ensemble observé ici.

<sup>9</sup> Nous précisons que les dates données ici sont les dates d'édition et non celles de rédaction des différents recueils d'Apollinaire.

<sup>10</sup> Expression empruntée à J. Anis, (1983).

<sup>11</sup> Expression empruntée à Apollinaire dans « Le Musicien de Saint Merry » (*Calligrammes*). Cf. Apollinaire, G. [et al.] (1994). *Œuvres poétiques*. Paris : Gallimard (« Bibliothèque de la Pléiade »), p. 188-191.

<sup>12</sup> Thibaudet, A. [1911] (2006). *La poésie de Stéphane Mallarmé*, (préface de J.-Y. Tadié). Paris : Gallimard, (« Tel »).

<sup>13</sup> Apollinaire, G. [et al.] (1994). *Op. cit.* p. 209 et p. 183-185

<sup>14</sup> URL: <http://www.tal.univ-paris3.fr/lexico/>

<sup>15</sup> URL: <http://ancilla.unice.fr/~brunet/pub/hyperbase.html>

<sup>16</sup> cf. Gardes-Tamines J. et Monte, M. (2007). Cette expression est, selon les auteures, « une de ces formules incantatoires un peu creuses dont sont friands les esprits épris de certitudes » qui définirait de façon très imprécise la poéticité.

<sup>17</sup> Cf. Chevalier, J.-Cl. (1970) et Morhange Begué, C. (1983).

<sup>18</sup> Nous pensons notamment à quelques calligrammes qui ne sont reproduits que sous la forme de fac-similés difficilement lisibles.

<sup>19</sup> Apollinaire, G. et Campa L. (2005), p. 77.

<sup>20</sup> Apollinaire, G. [et al.] (1994), p. 168-169.

<sup>21</sup> *Ibid.* p. 178.

<sup>22</sup> Ou, en tout cas, ce qu'en l'absence de toute ponctuation, nous pouvons considérer comme telle.

<sup>23</sup> Cette recherche dans le logiciel TXM passe par un filtrage des données, selon la requête : [tptos] (*TreeTagger Part Of Speech*).

<sup>24</sup> La valeur de cette spécificité repose sur un calcul de probabilité, lui-même fondé sur une moyenne des occurrences, établie sur l'ensemble du corpus. Si la valeur calculée est sur-représentée, c'est-à-dire que sa représentation dans un sous-corpus est plus importante que la moyenne globale, le résultat est positif ; si elle est sous-représentée, c'est l'inverse. Le résultat est appelé « écart-type ». Il sert d'indicateur pour mesurer la dispersion, ou l'étalement, d'un ensemble de valeurs autour de leur moyenne. Ainsi, plus l'écart-type est faible, plus la population est homogène. Pour étudier des spécificités de manière pertinente, il faut que l'écart-type soit supérieur ou égal à -2 ou +2.

<sup>25</sup> Il n'est pas étonnant que *L'Enchanteur pourrissant* révèle une parenté formelle avec le corpus poétique d'Apollinaire, car c'est une œuvre à la frontière des genres, entre poésie et conte (cf. Jean Burgos, *Apollinaire et « L'Enchanteur pourrissant » : genèse d'une poétique*, Calliopées, 2009.)

« Œuvre-prototype », pour reprendre l'expression de Jean Burgos, *L'Enchanteur pourrissant* est en marge de l'œuvre prosaïque apollinarienne : son écriture participe davantage du processus de rédaction poétique que les recueils de conte. Son étiquette générique, incertaine, tout comme l'ajout en guise d'épilogue d'« Onirocritique » à l'ensemble du *conte* en font de toute façon un cas à part.

<sup>26</sup> L'« écart-type » est un indicateur qui sert à mesurer la dispersion, ou l'étalement, d'un ensemble de valeurs autour de leur moyenne. Ainsi, plus l'écart-type est faible, plus la population est homogène. Pour étudier des spécificités de manière pertinente, il faut que l'écart-type soit supérieur ou égal à -2 ou +2.

<sup>27</sup> Le calcul des spécificités présenté ci-dessus crée une « norme » d'emploi des conjonctions à partir de leur fréquence et proportionnellement au volume de chaque recueil. Ainsi, la « moyenne » dont il est question ici est bien celle du « langage apollinarien », tel qu'il apparaît dans le corpus que nous avons constitué.

<sup>28</sup> Nous avons délibérément laissé de côté, pour cette étude, le cas des comparatives, qui mériterait un traitement à part, tant par le nombre (très important) de leurs occurrences, que par leur fonction poétique évidente.

---

<sup>29</sup> Apollinaire, G. [et al.] (1994). *Op. cit.* p. 83-84.

<sup>30</sup> *Ibid.* p. 146.

<sup>31</sup> *Ibid.* p. 188-191.

<sup>32</sup> Nous pensons en particulier à l'entreprise calligrammatique d'Apollinaire. Dans les « calligrammes », le poète met en effet en scène une double sémiologie qui pousse le lecteur à chercher le lien entre les deux modes de lecture du texte, à construire en quelque sorte un réseau sémantique qui intriquerait texte et image.