

Nature des marqueurs des séries linéaires dans des articles scientifiques

Veronika Laippala

Département d'études françaises
Université de Turku, Finlande
veronika.laippala@utu.fi

1 Introduction

En adoptant une approche textuelle, cet article propose d'examiner les différents types de séries linéaires, c'est-à-dire de séquences énumératives consistant en éléments textuels au moins partiellement introduits par des marqueurs linguistiques. Une série linéaire prototypique, par exemple, peut consister en items introduits par des marqueurs d'intégration linéaire classiques (dorénavant MIL, cf. Turco & Coltier, 1988) : *Premièrement... Deuxièmement... Troisièmement...* Ces séries organisent le texte en séquences d'items équivalents et ordonnés.

Notre problématique dans cet article est double. D'abord, nous examinerons les combinaisons de marqueurs utilisées pour signaler l'organisation du texte en séries. Comment l'organisation séquentielle est-elle indiquée dans le texte ? Dans un deuxième temps, nous étudierons les contextes d'usage des différents types de marqueurs. Pourquoi l'organisation séquentielle est-elle signalée comme elle l'est ? Ainsi, nous visons à observer par exemple les contextes textuels dans lesquels on préfère l'omission du marqueur, i.e. un marquage implicite à l'usage d'un MIL ou un MIL à un corrélat anaphorique (tels que *le premier*). L'alternance des marquages est notamment intéressante dans les séries de longueurs différentes où les niveaux global et local du texte peuvent jouer un rôle ainsi que dans les séries *non-homogènes* dans lesquelles les items d'une même série sont introduits par de différents types de marqueurs : *Premièrement... Le second élément... 0... Le quatrième et dernier type...* (cf. Turco & Coltier, 1988) (Jackiewicz, 2002).

Les différents types de marqueurs peuvent avoir différents types de fonctions dans l'organisation du texte. Or, les recherches antérieures sur les séries linéaires et les marqueurs de linéarité traitent souvent tous les marqueurs de linéarité possibles comme des MIL ou se concentrent sur la description d'un groupe de marqueurs spécifiques (entre autres Jackiewicz, 2005) (Porhiel, 2007) (Schnedecker, 2002) (Luc et Virbel, 2001). Dans ce qui suit, en plus de décrire comment les items des séries linéaires peuvent être introduits, nous tenterons d'acquérir une vision plus englobante sur différents marqueurs. De plus, nous cherchons à étudier les contextes textuels dans lesquels les différents types de marqueurs sont utilisés.

Le corpus de cette étude consiste en 8 articles scientifiques publiés en linguistique française. Les revues utilisées sont *Cahiers de Grammaire* (2001, 2004), *Contemporary French* (2005), *French Language Studies* (2006), *Linguisticae Investigationes* (2003), *Marges Linguistiques* (2003, 2006) et *Revue Romane* (2005, 2006). Etant donné que le métatexte qui souvent se réalise dans les séries linéaires est typique des textes scientifiques (Hyland, 1998) (Dahl, 2004), ce corpus forme un contexte opportun pour l'étude des séries linéaires. Une autre raison qui motive ce choix du corpus est la structure et l'organisation des articles scientifiques qui permettent un usage varié de séries dans des contextes différents. On notera toutefois qu'en raison du travail de révision et du processus d'édition de ces articles, ce corpus ne permet pas une étude de séries dont l'acceptabilité puisse être remise en question (cf. par exemple Stein-Zinz, 2006).

Cet article rend compte d'une analyse pilote afférent à notre projet de thèse. Ainsi, nous commençons par examiner d'un point de vue théorique les différents types de marquage possibles dans l'organisation séquentielle du texte. Ensuite, après avoir défini les moyens de marquage possibles, nous observerons les

combinaisons de marqueurs présentes dans notre corpus. Enfin, pour terminer, nous examinerons les contextes d'usage des différents types de marqueurs afin de répondre aux questions posées ci-dessus.

2 Marquage de la série

Dans ce qui suit, nous examinerons d'abord les différentes façons de marquer l'organisation séquentielle du texte : les organisateurs textuels, les anaphoriques et le non-marquage. Ensuite, nous étudierons deux moyens alternatifs d'identifier une série, la prédiction et la rétro-évaluation. Par la prédiction, l'auteur s'engage à organiser son texte d'une façon spécifique (Tadros, 1994) : par ex. *Cette série consiste en deux parties...* La rétro-évaluation, par contre, résume le contenu sémantique et réfère aux items de la série comme un tout : *Ces deux parties...* (Jackiewicz, 2005). Selon des études antérieures (cf. ci-dessous), ces types de marquage ont des rôles différents dans l'organisation textuelle, ce qui peut au moins partiellement expliquer leurs usages dans des contextes textuels différents. Ainsi, avant d'examiner la façon dont ils sont utilisés dans le corpus, il est nécessaire de les étudier de plus près.

2.1 Connecteurs, organisateurs textuels et cadres

Les marques les plus étudiées comme introducteurs d'un item de la série linéaire sont les MIL (Turco & Coltier, 1988). Ces marqueurs, qui « perdent leur vocation primitive pour désigner la succession des constituants discursifs » (Turco & Coltier, 1988 : 58), fonctionnent soit comme les marqueurs d'ouverture signalant le premier item de la série soit comme marqueurs de relais et de clôture introduisant les items suivants. Turco & Coltier (1988) incluent dans ce groupe des marqueurs comme *d'abord, enfin et également, dans un troisième temps*. Ce groupe de MIL a été étendu par Adam & Revaz (1989) qui y insèrent également des expressions du type *le premier* et *le deuxième (+nom)*. Le même principe a été suivi par Jackiewicz (2002) (2005) et Jackiewicz & Minel (2003).

Pour notre recherche, le groupe des MIL se relève être problématique si l'on y inclut des expressions du type *le deuxième* et *le deuxième (+nom)*. Le problème résulte du fait que, dans la plupart des recherches, une distinction est établie entre *anaphoriques*, qui marquent des relations entre un référent antérieur dans le texte et une anaphorique, et *connecteurs* qui relient des contenus propositionnels et définissent la nature du lien entre eux (Charolles, 1997 : 3) (Charolles, 2005 : 13). Ces deux types de *relations de cohérence* sont des attestations linguistiques de la *cohérence*, d'un principe selon lequel un texte peut être interprété dans le contexte (Charolles, 1997 : 2-3) (Charolles, 2005 : 12) (cf. aussi Halliday & Hasan, 1976).

Comme connecteurs peuvent fonctionner plusieurs types d'adverbes, de conjonctions et de locutions conjonctives. Dans l'approche textuelle, la fonction de ces marqueurs est communément définie comme signalant la façon dont la relation de cohérence en question doit être interprétée (entre autres Busquets & al., 2001) (Charolles, 2005 : 13). Or, la fonction de ces marqueurs peut aussi être vue comme plus complexe. Par exemple, Schneuwly & al. (1989) opposent les connecteurs reliant des structures propositionnelles et les *organisateur textuels* structurant le texte de façon hiérarchique à différents niveaux et prenant ainsi en compte plusieurs dimensions du texte. Ils posent trois critères que les organisateurs textuels doivent remplir : (1) joindre des structures propositionnelles en précisant leur insertion dans le co-texte ou dans le contexte de production, (2) être placé hors la syntaxe de la phrase et (3) ne pas être soumis aux phénomènes d'accord (Schneuwly & al., 1989). Ils concluent (1989 : 43) que deux types d'organisateur textuels peuvent être distingués ; ceux qui le sont toujours (conjonctions de coordination et de subordination) et ceux qui le sont lorsqu'ils sont placés hors de la structure syntaxique de la phrase (*à sept heures, plus tard...*).

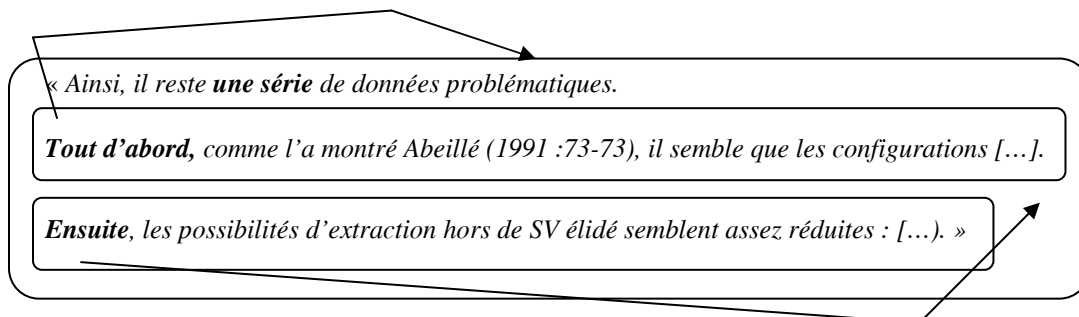
Une autre raison qui justifie la distinction entre les anaphoriques et les connecteurs d'une part et les cadres de l'autre est présentée par Charolles (1997) qui définit une fonction cohésive pour les adverbiaux détachés à gauche. Cette fonction, désignée *cadrage*, crée des relations de référence entre un adverbial détaché à gauche, le cadre, et les éléments qui le suivent dans le texte et qui doivent être interprétés selon le critère établi par le cadre. Grâce à sa position initiale, la portée du cadre, c'est-à-dire le champ

d'influence du critère sémantique spécifié, excède la phrase dont il fait syntaxiquement partie. (Péry-Woodley, 2005 : 5) (Charolles, 2005 : 1, 4)

La différence fondamentale entre les connecteurs et les cadres est la direction de référence ; tandis que les premières établissent des relations avec des éléments antérieurs dans le texte, le critère spécifié par le cadre crée des liens avec des éléments suivants dans le texte (Charolles, 1997 : 3) (Péry-Woodley, 2005 : 6). Le plus souvent, la fonction cadrative est assurée par des adverbiaux spatiaux et temporels. Pourtant, selon Charolles (2005 : 4), les MIL (en position initiale) peuvent aussi fonctionner comme cadres et sont en fait utilisés grâce à leur capacité de cadrage.

Bref, afin de garder la distinction entre les anaphoriques et les cadres, nous suivons Schneuwly & al. (1989) et distinguons les *organiseurs textuels* remplissant les critères présentés *supra* et les autres marqueurs. De plus, comme les marqueurs de linéarité fonctionnent souvent dans des structures hiérarchiques telles que les séries récursives et les structures énumératives à deux temps (Porhiel, 2007) (Jackiewicz, 2005) (cf. *infra*), nous nous en tenons également à ce terme. Enfin, conformément à Jackiewicz (2005), nous distinguons aussi les marqueurs *absolus* (*deuxièmement*, *finale*) signalant une position exacte dans la série et les marqueurs *relatifs* (*en outre*, *aussi*) introduisant seulement une relation additive. L'alternance de ces différents types de marqueur est illustrée par l'exemple suivant, où l'item 1 est introduit par un organisateur textuel absolu et l'item 2 par un marqueur relatif. Les organisateurs textuels comme cadres sont marqués par des cadres et les relations signalées par des organisateurs textuels par des flèches.

Graphique 1.



2.2 Anaphores

Contrairement aux organisateurs textuels, les anaphoriques ne remplissent pas les critères posés par Schneuwly & al. (1989) ; elles fonctionnent souvent par ex. comme sujets de la phrase. Ainsi, elles ne peuvent pas non plus fonctionner comme cadres. Dans la littérature linguistique, l'anaphore a été définie de nombreuses façons. En général, les travaux s'accordent sur le fait qu'elle représente « un rapport dissymétrique entre un terme, appelé l'anaphorique, et un antécédent » (Corblin, 1995 : 31). Ce qui importe dans ce rapport est son asymétrie ; pour être interprété, l'anaphorique a besoin d'un antécédent qui définit les conditions d'interprétation de l'anaphorique (Corblin, 1995 : 32) (pour une définition alternative, voir par ex. Cornish, 2005). Par exemple, les marqueurs de linéarité *le premier* – *le second* peuvent fonctionner comme anaphoriques ; afin de les interpréter, on a besoin de revenir sur ce qui a été précédemment dit dans le texte.

En ce qui concerne les marqueurs de linéarité anaphoriques, pour les besoins de cette étude nous y incluons toutes les expressions du type *le premier* – *le deuxième* (cf. Schnedecker, 2002). De plus, nous considérons comme anaphoriques les marques de linéarité du type *le premier* (+ nom) quand ils sont de nature non-autonome et renvoient à un élément antérieur dans le texte, c'est-à-dire quand ils sont coréférentiels (pour la problématique de la coréférentialité, cf. par ex. Charolles & Schnedecker, (1993). Vu la nature restreinte des anaphoriques possibles dans cette étude, cette restriction est logique ; les

numériques comme *le premier – le troisième élément* doivent avoir un antécédent dans le texte pour faire partie d'une série linéaire (cf. aussi Corblin, 1995 : 34-38).

Plus spécifiquement, les anaphoriques dans cet article sont dans une relation de coréférence partielle avec leur antécédent ; un item représente une partie de l'ensemble auquel les items réfèrent (Milner, 1978 : 51-54). Par exemple, dans l'exemple suivant, les anaphoriques *l'une* et *l'autre* réfèrent chacun subséquentement aux parties de la prédiction *deux parties* qu'ils représentent.

Tableau 1.

Prédiction	« Dans la même lignée que Webber (1979), H & S présupposent que le participant construit un modèle de discours en deux parties :
Item 1.	<i>l'une</i> correspond à la représentation propositionnelle du discours,
Item 2.	et <i>l'autre</i> au modèle du monde évoqué dans le discours [...].

2.3 Organismes textuels et anaphoriques dans l'organisation du texte

Ci-dessus nous venons de définir la façon dont nous comprenons et différencions les cadres et les organisateurs textuels d'une part et les anaphoriques de l'autre. Une distinction entre ces types de marqueurs est essentielle dans notre recherche puisque, selon des études antérieures, ils ont des capacités différentes d'organiser des segments textuels. Les cadres (et ainsi les organisateurs textuels) ont selon Charolles (1997) une portée en théorie illimitée. Jackiewicz & Minel constatent (2003 : 4) pourtant que la portée d'un MIL excède rarement deux paragraphes. Aussi Ho Dac (2007 : 227, 238-239) note que les organisateurs textuels ne sont pas utilisés au début de section et qu'également les connecteurs « pures » montrent une tendance très forte à organiser des segments intraparagraphiques. Ainsi, les organisateurs textuels ne devraient pas selon ces résultats introduire d'items très vastes dans notre corpus non plus.

De même que les organisateurs textuels, selon quelques études antérieures, ni les anaphoriques n'organisent de segments textuels vastes. Selon Schnedecker (2000 : 22), les zones d'antécédence des anaphoriques *le premier... le second...* ne s'écartent pas de plus de deux phrases des marqueurs. Aussi (Halliday, 1985 : 302) constate que les ellipses, qui ont la même forme de surface que les anaphoriques (*le premier.. le second...*) observés ici, réfèrent le plus souvent à la phrase précédente.

Il semble donc que ni les organisateurs textuels ni les anaphoriques n'organiseront de segments textuels très étendus. Ainsi, il est intéressant de savoir comment l'organisation des séries linéaires englobant plusieurs sections est marquée, si on assume qu'il y en a dans le corpus ? De plus, il est intéressant d'examiner quels types de segments organisent les séquences non-homogènes dans lesquelles tous les items ne sont pas introduits par le même type de marqueurs (cf. ci-dessous). Aussi, il se peut que les résultats obtenus dans d'autres études ne valent pas pour les articles scientifiques ; selon Ho Dac (2007 : 23) et une hypothèse posée par Péry-Woodley (2000 : 50, 111), le marquage de l'organisation textuelle dépend du genre du texte.

2.4 Répétition lexicale et items non marqués

Comme les marqueurs de linéarité, c'est-à-dire les organisateurs textuels et les anaphoriques, ne sont que des signes explicites de la cohérence, il n'est pas nécessaire que tous les items d'une série soient introduits par ces marqueurs pour être identifiés. Cependant, malgré l'absence de marqueur de linéarité, les items d'une série gardent souvent des relations lexicales entre eux grâce à leur homogénéité (cf. les définitions de la sériation par ex. chez Péry-Woodley (2000 : 138) et Jackiewicz (2005)). Par exemple, dans l'exemple suivant, l'item 1 peut être identifié grâce au mot *question* qui est répété dans l'item 2.

Tableau 2.

Item 1.	« <i>La question principale qui guidera notre investigation sera de déterminer dans quelle mesure des exemples tels que (7a) et (7b) peuvent être assimilés [...].</i> »
Item 2.	« <i>Une seconde question que l'on abordera dans cet article est celle du manque de productivité [...].</i> »

Cette répétition lexicale entre les items pourrait être interprétée comme anaphorique ; l'item répété est souvent anaphorique (cf. Apothéloz, 1995 : 36-38). De plus, les éléments textuels répétés sont dans une relation de coréférence partielle avec leur antécédent de la même façon que les marqueurs de linéarité anaphoriques présentés *supra*. Une deuxième possibilité serait de considérer cette relation lexicale comme des liens de répétition (*repetition links*) qui permettent à l'auteur d'ajouter quelque chose de nouveau au texte (cf. Hoey, 1991) ; cette approche soulignerait la structuration du texte par différents types de liens lexicaux. Pourtant, étant donné que notre but principal est d'étudier les marques de **linéarité**, nous ne nous concentrons pas sur l'identification détaillée de ces liens lexicaux. Par contraste, pour l'instant, nous définissons tous les items qui ne sont pas introduits par un organisateur textuel, par un anaphorique ou par un autre connecteur comme non marqués.

2.5 Organisation séquentielle par prédiction et rétro-évaluation

L'organisation séquentielle d'un segment textuel peut aussi s'effectuer par une prédiction ou une rétro-évaluation. Ces éléments créent des relations de cohérence entre les séries proprement dites et le texte et les relient ainsi plus fermement entre eux. De plus, la prédiction et la rétro-évaluation contribuent au marquage de l'organisation séquentielle surtout quand les items de la série sont non marqués.

Dans cet article, comme notre but est d'étudier le marquage des items de la série, le rôle de la prédiction et de la rétro-évaluation est souligné lorsque les items ou une partie des items ne sont pas marqués et lorsque les marqueurs ont une valeur relative (vs absolue, cf. 2.1 et Figure 1). L'exemple suivant témoigne de l'importance de ces éléments dans un tel cas : les items de la série sont non marqués et introduits seulement par la rétro-évaluation qui réfère à l'ensemble des deux items.

Tableau 3.

Item 1.	« [...] : si les cadres sont caractérisés par une portée vers l'avant, puisqu'ils indexent une ou plusieurs propositions qu'ils introduisent, »
Item 2.	« les anaphores regardent en général vers l'arrière, vu qu'elles renvoient à un élément du contexte précédent »
Rétro-évaluation	« Il est donc intéressant de se pencher sur des expressions qui partageraient les deux caractéristiques [...]. »

Dans l'analyse de la prédiction, nous suivons la typologie de prédiction de Tadros (1994). Elle distingue en tout sept catégories de prédiction, dont nous appliquons trois. La plus étudiée de ces catégories est l'énumération, qui consiste en un numéral et en un énuméré. Par exemple, la prédiction du Tableau 1 est énumérative. Plus spécifiquement, il s'agit d'une énumération exacte ; nous distinguons encore l'énumération exacte et l'énumération vague où le numéral n'est pas absolu (telle que *une série* dans la Figure 1). Nous prenons également en compte dans l'analyse comme prédictions les *questions* qui déclenchent la série et sont fréquentes surtout dans les séries argumentatives (cf. 3.3).

Enfin, la dernière catégorie de prédiction que nous adaptons diffère des autres en ce qu'elle n'a pas nécessairement de marque de surface saillante. Plus implicite que les autres catégories, l'*advance labelling* (par faute de traduction nous utilisons le terme anglais) consiste en « étiquetage et engagement de l'auteur à un acte de discours » (Tadros, 1994 : 73). Par exemple, l'*advance labelling* se réalise dans des prédictions telles que « ces résultats nous amènent à distinguer ». De plus, dans ce type de prédiction nous incluons les structures avec un élément cataphorique, telles que *dans ce qui suit* et *ci-dessous*.

Pour déterminer les types de rétro-évaluation, nous suivons Jackiewicz (2005), qui définit les marqueurs de rétro-évaluation comme reformulateurs (*bref, autrement dit*) ou de quantification totalisante (*les deux, l'ensemble*). Afin d'en savoir plus sur leur fonction dans le marquage de la linéarité, nous classons les marques d'une façon plus spécifique. En bref, comme types de rétro-évaluation, nous distinguons les organisateurs textuels conclusifs (*pour résumer, donc*) ainsi que les anaphores résomptives (*ces parties, les éléments ci-dessus*) et les anaphores résomptives avec un numéral (*ces deux parties, les trois éléments présentés*) (cf. Auricchio & al., 1995 sur l'anaphore résomptive).

3 Marquage de l'organisation séquentielle dans des articles scientifiques

Notre corpus consiste en 8 articles scientifiques. Dans ce matériel, il y a en tout 70 séries linéaires. Comme le montre l'analyse *infra*, les types de marqueurs ainsi que les longueurs des séries varient. Pour étudier les contextes textuels de la variété de marqueurs, nous examinerons les différentes séries en les classant selon le(s) type(s) de marquage identifié(s). Enfin, dans la dernière section, nous analyserons le marquage des séries à partir de la longueur des items, ce qui nous permettra d'étudier les marqueurs et leurs portées (pour les organisateurs textuels) aux niveaux global et local du texte.

3.1 Hypothèses

Ci-dessus, nous avons noté que selon des études antérieures, ni les anaphoriques ni les organisateurs textuels ne devraient marquer des séries linéaires très vastes. Notre étude nous permettra de tester ces résultats, que nous pouvons aussi prendre comme hypothèses de cet article, dans un corpus d'articles scientifiques ; comme nous avons également constaté, le marquage de l'organisation textuelle peut pourtant également dépendre du genre. Deuxièmement, un autre résultat de Jackiewicz (2002) que nous pouvons tester dans notre corpus concerne les séries avec des types de marqueurs différents (par ex. anaphoriques, autres connecteurs, items non marqués). Jackiewicz note que (2002 : 4), ces séries non-homogènes sont assez fréquentes. Il est intéressant d'examiner si le processus d'édition et de relecture que les articles scientifiques ont parcouru a un effet sur la fréquence de ces séries. Cette caractéristique de notre corpus peut aussi influencer l'usage atypique des marqueurs, tel que l'usage de *d'une part* sans *d'autre part* (cf. Stein-Zinz, 2006). Notre dernière hypothèse sera que ce type de série est rare dans notre corpus.

3.2 Séries homogènes

Le premier trait manifeste dans le marquage des séries étudiées est la variation des types de marqueurs dans une même série ; ainsi, nous pouvons constater que les résultats de Jackiewicz (2002) sont valides aussi pour notre corpus. Les séries *homogènes* où tous les items sont introduits par un même type de marqueur de linéarité (organisateur textuel ou anaphorique) sont plutôt minoritaires. En tout, le corpus contient 19 séries homogènes sur un nombre total de 70 séries. Dans le tableau 4, les séries sont classifiées selon le type de marqueurs. Comme *d'une part – d'autre part* et *l'un – l'autre* sont des paires de marqueurs fréquentes, nous les avons placées séparément.

Tableau 4.

types de marqueurs	marqueurs	nombre d'occurrences
organisateur textuels	d'une part – d'autre part,	5
autres organisateurs textuels	d'abord – ensuite – (enfin) premièrement – deuxièmement, etc.	5
anaphoriques	l'une(e) – l'autre	3
autres anaphoriques	le premier – le deuxième la première (nom) – la deuxième (nom)	5
autres	ou bien – ou bien	1

Le fait que nous n'y avons pas inclus de séries dans lesquelles l'ordre des marqueurs n'est pas ascendant aurait un effet (quoique minoritaire) sur la rareté des séries homogènes dans notre corpus. De plus, nous avons exclu les séries dont les marqueurs varient (*le premier... le second élément*). En pratique, l'application de ces critères nous a fait exclure 4 séries des séries homogènes ; trois d'entre elles sont du type (*la première + (nom) – la deuxième*) et une du type (*la seconde – la première – la troisième*). Ainsi, si on considérait ces séries comme homogènes, cette classe compterait 23 séries.

3.3 Séries avec marqueurs de linéarité mixtes

Les séries dont le marquage correspond le plus aux séries homogènes sont celles dont les items sont tous marqués, mais par des marqueurs de linéarité différents (organisateur textuel, anaphorique). Dans ce groupe, nous avons également inclus les séries dont un item est marqué par un autre type de connecteur, tel qu'un connecteur thématique *en ce qui concerne* ou *quant à* (cf. Porhiel, 2005). En tout, le corpus contient 14 séries de ce type ; elles sont donc aussi fréquentes que les séries homogènes.

L'exemple suivant est caractéristique de ce type de série ; chaque item est introduit par un type de marqueur différent. De plus, comme l'item 2 inclut une autre série linéaire (quoique très courte), cette série est aussi un exemple d'une série récursive. Une troisième caractéristique du marquage des séries linéaires illustrée par cet exemple est la reduplication des marqueurs ; les items 1 et 3 sont introduits par deux marqueurs.

Tableau 5.

Prédiction « advanced labelling »	« Notre article s'articulera <i>de la manière suivante</i> .
Item 1. Organisateur textuel absolu x 2	<i>En vue de caractériser davantage l'ellipse modale, nous explorerons d'abord, dans la section 2, les distinctions opérées par Hankamer & Sag (1976, 1984), dorénavant H&S, Chao (1988) et Gardent (1991) entre les deux types de processus anaphoriques : l'ellipse verbale d'une part, l'anaphore de complément nul d'autre part.</i>
Item 2. Organisateur textuel absolu	<i>Dans la section (3), nous nous pencherons alors, de manière plus spécifique, sur la distribution et les contraintes pesant sur l'ellipse modale. En particulier, nous montrerons, au vu de certains diagnostics, que ces constructions partagent beaucoup de caractéristiques avec la VPE de l'anglais.</i>
Item 3. Marqueur anaphorique + connecteur thématique	<i>La section (4) sera, quant à elle, consacrée à la question de la productivité des 59 modaux français selon leur interprétation déontique ou épistémique. Cette question nous amènera à discuter les interactions entre ellipse et quantification.</i>
Item 4. Marqueur anaphorique	<i>Une dernière section nous permettra d'offrir quelques remarques finales pour conclure. »</i>

Bien que l'exemple *supra* présente bien la majorité des séries linéaires avec des marqueurs mixtes, dans les 14 séries de ce type, une sous-catégorie peut être discernée. En fait, elle inclut 4 séries dont la forme de surface est similaire aux séries linéaires mais dont la fonction est différente : au lieu de structurer le texte en parties structurellement et fonctionnellement équivalentes (cf. définition de l'énumération Péry-Woodley, 2000), ces séries argumentatives (Jackiewicz, 2005) l'organisent en tant que séquence argumentative en thèse première, en données (prémisses) et enfin en nouvelle thèse (cf. Adam, 1992) (Toulmin, 1958). Ces séries, dans quelques cas précédés d'une question, commencent par un organisateur textuel (très souvent *d'abord*) qui introduit le premier item de la série et qui aussi est la thèse originale. L'item 2 est sans exception introduit par un connecteur adversatif, tel que *mais* : cet item présente les données (prémisses). Enfin, la nouvelle conclusion est présentée dans la rétro-évaluation, soit implicitement, soit suivie par un connecteur conclusif.

3.4 Séries partiellement non marquées

La plus fréquente des types de séries est la classe des séries non marquées. Nous commençons par analyser celles dont seulement une partie d'items sont non marqués, autrement dit introduits implicitement, sans organisateur textuel, anaphorique ou connecteur.

Le corpus contient 18 séries partiellement non marquées. La caractéristique distinctive de ces séries par rapport aux types examinés *supra* et *infra* est la nature vague du marquage des items. Malgré le fait que l'annonce du nombre d'items soit possible dans ce groupe par une prédiction énumérative exacte ou par une rétro-évaluation anaphorique avec un chiffre, ce n'est le cas que rarement, dans trois cas sur 26 prédictions exactes / rétro-évaluations avec un chiffre dans tout le corpus.

Un deuxième facteur contribuant à la nature imprécise des séries partiellement non marquées est la fréquence des marqueurs relatifs ; dans les 18 séries, on trouve 10 marqueurs qui ne spécifient pas la

place exacte de l'item introduit dans la série. Ainsi, en plus des items non marqués, ces séries incluent un nombre important d'items introduits par des marqueurs tels que connecteurs thématiques, organisateurs textuels relatifs et anaphoriques relatifs.

Le tableau suivant montre la répartition des prédictions / rétro-évaluations exactes entre les séries homogènes et les séries partiellement non marquées. Comme on peut le constater, les marqueurs absolus et les prédictions / rétro-évaluations exactes sont beaucoup plus fréquentes dans les séries homogènes, déjà explicitement marquées, que dans les séries partiellement non marquées. Ainsi, l'imprécision semble être une des caractéristiques de ces séries ; leur fonction n'est pas d'organiser le texte aussi strictement que celle des séries homogènes.

Tableau 6.

	Séries homogènes	Séries partiellement non marquées
Organisateurs textuels et anaphores relatives + autres connecteurs comme marqueurs d'items	3 / 23	10 / 18
Prédiction exacte / rétro-évaluation avec un chiffre	10 / 23	3 / 18

Le non-marquage et le marquage relatif posent aussi des problèmes d'identification pour les séries. Afin de les identifier dans le texte, nous avons défini des critères que les séries incluses dans l'analyse doivent remplir. Brièvement, nous exigeons que (1) au moins un item de la série soit marqué par un organisateur textuel absolu ou que, (2) s'il n'y a que des autres connecteurs, organisateurs textuels ou anaphoriques relatifs dans la série, la série soit précédée par une prédiction ou suivie par une rétro-évaluation. Ainsi, sont donc exclues de l'analyse les séries dont la linéarité peut être remise en question et les séries dont les items risquent d'être très hétérogènes (cf. la définition de l'énumération dans Péry-Woodley, 2000 : 138).

Dans l'exemple suivant, il y a deux marques de linéarité : la prédiction énumérative vague et l'adjectif *autre* qui a une valeur additive. Ainsi, malgré le fait que cette série remplisse les critères présentés ci-dessus, la linéarité n'y est pas très saillante. Enfin, on peut même se demander si ce type de série est réellement une structure énumérative, qui selon Péry-Woodley (2000 : 138) « consiste en entités structurellement & fonctionnellement équivalents, réalisés par la mise en forme identique ».

Tableau 7.

Prédiction énumérative vague	« Nous allons présenter ici quelques exemples de travaux cherchant les effets des [...]
Item 1. Non marqué.	<i>C'est à un éclairage à partir de l'histoire des langues que convie Banniard (1997) [...]</i>
Item 2. Non marqué	<i>Pour Lodge (1999), c'est le fait de considérer les formes vernaculaires de langue [...]</i>
Item 3.	<i>Un autre travail de Lodge (à paraître), concernant la grammaire et la prononciation [...]</i>
Item 4. Non marqué.	<i>Woolard (1985), discutant du modèle de Bourdieu autour du rôle de la langue dans [...] »</i>

3.5 Séries non marquées

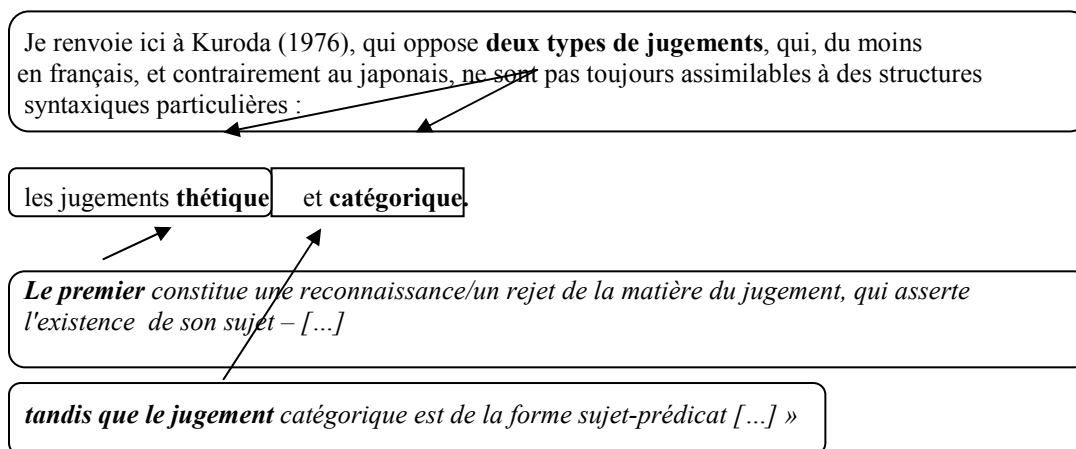
Enfin, le dernier type de série traité dans cet article est la série non marquée. Dans cette série, aucun item n'est introduit par un marqueur explicite ; le seul marqueur y est la conjonction *et* devant le dernier item.

Comme nous l'avons déjà signalé (cf. 2.5), les items et leur nombre sont indiqués dans ces séries par la prédiction ou par la rétro-évaluation. Afin d'exclure de l'analyse les séries dont la linéarité est plutôt questionable, nous exigeons que la prédiction ou la rétro-évaluation de ce type de série définisse le nombre exacte d'items. Ainsi, les séries dont la seule marque de surface serait la prédiction « *quelques (nom)* » ne seront pas prises en compte dans l'analyse.

Le corpus étudié contient 14 occurrences de ce type de série. Contrairement aux séries partiellement non marquées, l'imprécision n'est pas caractéristique de ces séries. Par contre, elles sont plutôt caractérisées par leur brièveté ; les longueurs des items de ces séries ne dépassent une phrase que rarement. La brièveté peut aussi expliquer le non-marquage des items ; comme les items sont déjà introduits ailleurs dans le texte, il n'est plus nécessaire de répéter le marquage dans la série proprement dite (cf. maxime de quantité, Grice (1975)).

Or, la brièveté risque également de résulter en complexité ; comme l'ont noté Jackiewicz & Minel, (2003 : 4), les séries courtes s'enchaînent souvent. Cette tendance est manifeste aussi dans notre corpus ; 10 sur 14 séries de ce type sont liées à une autre série pour former une série définie par Porhiel (2007) comme *une série à deux temps de nature organisationnelle*. La première série, non marquée, est suivie par une autre, dont les items sont introduits par des organisateurs textuels ou par des anaphoriques. Autrement dit, comme on peut le constater à partir de la Figure 2. ci-dessus, la première série fonctionne comme la zone d'antécédence à la deuxième, à laquelle les marqueurs de la deuxième série réfèrent (cf. aussi Schnedecker, 2002).

Graphique 2.



En fait, selon notre corpus, les séries à deux temps de nature organisationnelle semblent plus fréquentes que les autres types de série à deux temps définis par Porhiel (2007) ; celle de nature organisationnelle est le seul type présent dans notre corpus. Cependant, comme 10 séries à deux temps de nature organisationnelle du corpus sur 11 ont comme zone d'antécédence une série implicite et 10 séries non marquées sur 14 font partie d'une série à deux temps, il est possible de conclure que cette combinaison semble plutôt fréquente.

3.6 Marqueurs globaux et locaux – séries globales et locales

Une distinction entre les marqueurs organisant des éléments textuels étendus et ceux structurant des éléments courts a déjà été faite par Adam et Revaz (1989) qui différencient les marqueurs globaux fonctionnant sur le plan du texte et les marqueurs locaux qui relient des phrases ou des propositions. Une autre classification basée sur la longueur de la portée a été définie par Péry-Woodley (2000 : 50) qui place les relations de cohérence sur une échelle ; les relations du 1^{er} niveau relient des macro-segments, tandis que celles du dernier niveau des propositions.

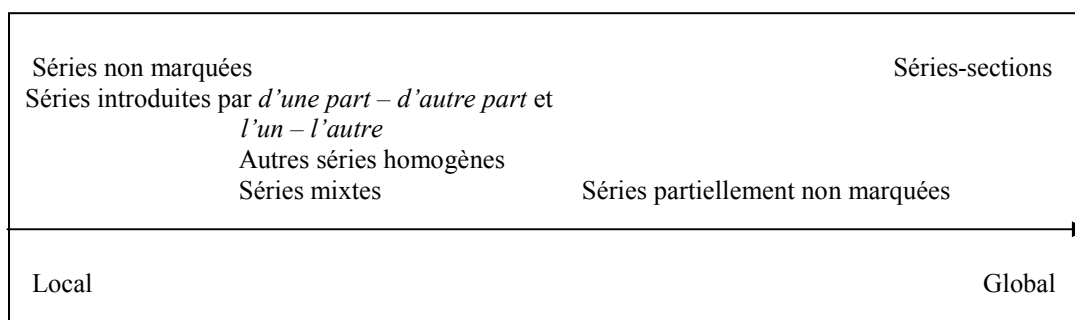
Dans le corpus, les longueurs des séries varient considérablement ; d'une part, il y a les séries courtes, locales selon Adam & Revaz (1989), et de l'autre les séries très longues, globales, qui peuvent même s'étendre sur plusieurs sections de l'article et ainsi organiser des segments textuels de plusieurs pages (dorénavant séries-sections). Nous avons déjà noté que les séries non marquées sont très courtes (cf. 3.5) ; la prédiction ou la rétro-évaluation peuvent fonctionner comme seul marqueur de la série dans les séries dont les longueurs des items n'excèdent pas une phrase. De plus, sont très courtes les séries introduites par les anaphoriques *l'un – l'autre* ; dans notre corpus, ceux-ci ne marquent jamais d'item dont la longueur dépasse une phrase.

Un autre résultat de Jackiewicz & Minel (2003) qui s'est avérée partiellement correcte dans notre corpus est la portée des organisateurs textuels. Car, comme le notent Jackiewicz & Minel (2003), elle ne dépasse quelques paragraphes que rarement. Il est pourtant nécessaire de spécifier que ce sont surtout les organisateurs textuels des séries homogènes ; ceux faisant partie des séries mixtes ou partiellement non marquées semblent relier également des segments plus étendus. La portée la plus restreinte est celle des séries homogènes introduites par *d'une part – d'autre part* ou *d'un côté – de l'autre* : dans notre corpus, la portée de ces marqueurs ne dépasse deux phrases que rarement.

Bien que les séries homogènes introduites par des organisateurs textuels ne relient pas d'éléments textuels longs, celles introduites par des anaphoriques du type *la première + (nom) – la deuxième + (nom)* peuvent le faire ; deux sur cinq séries de ce type fonctionnent à un niveau global où les items de la série forment une section. En fait, ces séries pourraient être classées dans leur propre sous-classe ; évidemment très longues, elles fonctionnent à un niveau différent, plus global que les autres séries. De plus, chaque item de ce type de série est précédé du titre de la section. Dans notre corpus, on trouve sept occurrences de ce type de série que nous appelons dorénavant série-section. Comme cinq de ces séries sont du type partiellement non marqué (et deux homogènes avec marqueurs anaphoriques), on pourrait penser que le titre de la section au début de l'item influence son marquage. Or, le rôle des titres n'est pas forcément décisif ; le marquage des séries-sections ressemble plutôt aux autres séries longues qui aussi sont souvent partiellement non marquées.

Pour conclure la discussion sur les marqueurs globaux et locaux, il semble que les séries globales et locales sont introduites dans le texte par des combinaisons de marqueurs différentes. Les séries non marquées ainsi que les séries homogènes introduites par *d'une part – d'autre part* et par *l'un – l'autre* n'organisent pas de segments très étendus. Cette tendance est manifeste dans notre corpus. Les séries-sections, par contre, souvent partiellement non marquées, fonctionnent à un niveau global du texte. Le schéma suivant illustre la répartition des différents types de séries sur un continuum local-global.

Graphique 3.



4 Conclusion

Le but de cet article a été d'examiner les différents moyens d'indiquer l'organisation séquentielle du texte ainsi que d'analyser les contextes d'usage des différents types de marqueurs. Les questions auxquelles nous avons cherché à répondre sont : Comment l'organisation séquentielle est-elle indiquée dans le texte ? Pourquoi l'organisation séquentielle est-elle signalée comme elle l'est ?

Afin d'examiner les différents moyens de signaler l'organisation du texte en séries linéaires, nous les avons d'abord classées en types de séries homogènes, mixtes, partiellement non marquées et non marquées. Ensuite, au cours de l'analyse nous avons également défini les classes distinctes pour les séries-sections et les séries argumentatives (aussi définies dans Jackiewicz, 2005). Les séries homogènes, dans lesquelles tous les items sont introduits par un même type de marqueur, sont dans notre corpus minoritaires. Ainsi, nous pouvons confirmer que les résultats de Jackiewicz (2002) sont valides pour les articles scientifiques de notre corpus aussi ; les séries non-homogènes, dans lesquelles les items d'une même série sont introduits par de différents types de marqueurs, sont majoritaires malgré le processus de révision parcouru par les articles du corpus. Le processus de révision n'est pourtant pas nécessairement sans effet sur le marquage. Car, comme nous avons assumé, le marquage atypique des paires de marqueurs, tel que l'usage de *d'une part* sans *d'autre part*, est absent dans notre corpus. Ainsi, il n'est pas exclu que ce type de série soit modifié par le processus de relecture.

Notre deuxième question de recherche portait donc sur l'alternance des différents types de marqueurs dans des contextes textuels différents. Le premier résultat surprenant concernant cette question est la distinction entre deux types de séries avec des marqueurs et fonctions textuelles différents. L'ordre et le nombre des items des séries homogènes sont clairs : les items y sont explicitement marqués et souvent précédés ou suivis par une prédiction exacte / rétro-évaluation avec un chiffre. Les séries partiellement non marquées, par contre, contiennent des items non marqués et une proportion d'items introduits par des marqueurs relatifs élevée. De plus, le nombre des items de ces séries est rarement spécifié par une prédiction / rétro-évaluation exacte bien que cela soit possible dans ce type de série. Par conséquent, le nombre et l'ordre des items de ce type de série ne sont pas nécessairement clairement indiqués. Ainsi, étant donné le marquage différent des items, les capacités organisatrices et les fonctions textuelles des séries homogènes et partiellement non marquées ne sont pas les mêmes. Tandis que les premières structurent les éléments en blocs dont les limites et l'ordre sont clairement indiqués, les secondes organisent le texte en éléments moins précis dont l'ordre ou le nombre absolu ne sont pas nécessairement primordiaux. Les séries linéaires ne visent donc pas toujours à organiser le texte de façon explicite en segments bien saillants.

Outre la fonction textuelle, le marquage des items est influencé par la longueur de la série en question : les séries étendues, globales, ne sont pas marquées de la même façon que les séries courtes. Ce résultat était partiellement attendu à la lumière des travaux antérieurs sur la portée des organisateurs textuels. Les séries non marquées et celles introduites par *d'une part – d'autre part* ou par *l'un – l'autre* fonctionnent selon nos résultats à un niveau local en organisant des segments textuels courts. Les séries partiellement non marquées, par contre, fonctionnent au niveau global et organisent des sections entières. Finalement, les séries mixtes et les autres séries homogènes organisent souvent des segments textuels de longueur moyenne. Ainsi, les combinaisons des différents types de marqueurs ne sont pas utilisées dans les mêmes contextes textuels mais elles ont leurs fonctions propres dans l'organisation du texte.

Enfin, il est essentiel de noter la taille modeste de notre corpus. Afin de s'assurer des résultats de cette étude, des analyses ultérieures avec un corpus plus étendu sont nécessaires.

Références

- Adam, J-M. (2001) [1992]. Les textes: types et prototypes. Récit, description, argumentation, explication et dialogue. Paris : Nathan.
- Adam, J-M. & Revaz, F. (1989). Aspects de la structuration du texte descriptif : les marques d'énumération et de reformulation. *Langue française*, 81, 59-98.
- Apothéloz, D. (1995). Rôle et fonctionnement de l'anaphore dans la dynamique textuelle. Genève : Librairie Droz.
- Auriccio, A. & al. (1995). L'anaphore démonstrative à fonction résomptive. *Pratiques*, 86, 27-52.
- Busquets, J. & al. (2001). La SDRT : Une approche de la cohérence du discours dans la tradition de la sémantique dynamique. *Verbum*, XXIII (1), 73-102.

- Charolles, M. (1997). Encadrement du discours: univers, champs, domaines et espaces. *Cahiers de Recherche Linguistique* 6, 1-73.
- Charolles, M. (2005). Framing adverbials and their role in discourse cohesion : from connection to forward labelling. *Papers of the Symposium on the Exploration and Modelling of Meaning (Sem 05)*, Biarritz, 14-15/11/05.
- Charolles, M. & Schnedecker, C. (1993). Coréférence et identité. Le problème des référents évolutifs. *Langages*, 112, 106-126.
- Corblin, F. (1995). Les formes de reprise dans le discours. *Anaphores et chaînes de référence*. Rennes : Presses Universitaires de Rennes.
- Cornish, F. (2002). Degrees of Indirectness: Two Types of Implicit Referents and their Retrieval via Unaccented Pronouns. In Branco, A., McEnery, T. Mitkov, R. (éds.), *Anaphora processing : linguistic, cognitive, and computational modelling*. Amsterdam : J. Benjamins, 199-220.
- Dahl, T. 2004. Textual metadiscourse in research articles: a marker of national culture or of academic discipline? *Journal of Pragmatics*, 36, 1807-1825.
- Grice, H.P. (1975). Logic and Conversation. In Cole, P. & Morgan, J.L (éds.), *Syntax and Semantics 3*. New York : Academic Press, 41-58.
- Halliday, M.A.K. & Hasan, R. (1976). *Cohesion in English*. London : Longman.
- Halliday, M.A.K. (1985). *An Introduction to Functional Grammar*. USA, Australia : Edward Arnold.
- Ho Duc, L.-M. (2007). La position initiale dans l'organisation du discours : une exploration en corpus. Thèse de doctorat. Université Toulouse-le Mirail.
- Hoey, M. (1991). *Patterns of Lexis in Text*. Hong Kong : Oxford University Press.
- Hyland, K. (1998). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics*, 30, 437-455.
- Jackiewicz, A. (2002). Répérage et délimitation des cadres organisationnels pour la segmentation automatique des textes. CIFT 02 : Colloque International sur la Fouille de Texte – Hammamet – Tunisie – 20-23 octobre 2002.
- Jackiewicz, A. (2005). Les séries linéaires dans le discours. *Langue française*, 148, 95-110.
- Jackiewicz, A. & Minel, J.-C. (2003). L'identification des structures discursives engendrées par les cadres organisationnels. Actes de la 10ème Conférence Traitement Automatique du Langage Naturel (TALN 2003) Batz-sur-Mer, 155-164.
- Luc, C. & Virbel, J. (2001). Le modèle d'architecture textuelle : fondements et expérimentation. *Verbum*, XXIII (1), 103-120.
- Martin, J.R. (1992). *English text : system and structure*. (1978). *De la syntaxe à l'interprétation : quantités, insultes, exclamations*. Paris : Seuil.
- Péry-Woodley, M.-P. (2000). Une pragmatique à fleur de texte : approche en corpus de l'organisation textuelle. Université de Toulouse-LeMirail : ERSS.
- Péry-Woodley, M.-P. (2005). Introduction. *Langue française*, 148, 3-8.
- Porhiel, S. (2005). Les séquences thématiques. *Langue française*, 148, 111-126.
- Porhiel, S. (2007). Les structures énumératives à deux temps. *Revue Romane*, 42(1), 103-135.
- Sinclair, J. (2004). *Trust the text*. London and New York : Routledge.
- Schnedecker, C. (2000). Ordres des ordinaux pronominaux. *Travaux de linguistique* 41(2000), 7-34.
- Schnedecker, C. (2002). Les corrélats anaphoriques l'un /l'autre, le premier/ le second : aspects cohésifs de la référence « en stéréo ». In Andersen, H.L. & Nolke, H. (éds.), *Macro-syntaxe et macro-sémantique*. Berne : Peter Lang, 257-283.
- Schneuwly, B. & al. (1989). Les organisateurs textuels dans quatre types de textes écrits. Etude chez des élèves de dix, douze et quatorze ans. *Langue française*, 81, 40-58.

- Stein-Zinz, S. (2006). De l'alterité spatiale à l'organisation textuelle : la locution d'une part... d'autre part. In Enjalbert, P. & al. (éds.), *Schédae. Colloque International Discours et Document*. Caen : Presses Universitaires de Caen, 26-34.
- Tadros, A. (1994). Predictive categories in expository text. In Coulhard, M. (éd.), *Advances in written text analysis.*: London and New York : Routledge, 69-82.
- Toulmin, S. (1958). *The Uses of Argument*. Cambridge : University Press.
- Turco, G. & Coltier, D. (1988). Des agents doubles de l'organisation textuelle, les marqueurs d'intégration linéaire. *Pratiques*, 57, 57-79.