HAL
open science

Diachronie de l'oral représenté

Céline Guillot-Barbance, Alexei Lavrentiev, Serge Heiden, Bénédicte Pincemin

To cite this version:

Céline Guillot-Barbance, Alexei Lavrentiev, Serge Heiden, Bénédicte Pincemin. Diachronie de l'oral représenté : délimitation et segmentation interne du dialogue (IXe-XVe siècle). Wendy Ayres-Benett; Anne Carlier; Julie Glikman; Thomas Rainsford; Gilles Siouffi; Carine Skupien Dekens. Nouvelles voies d'accès au changement linguistique. Actes du colloque de la SIDF, Classiques Garnier, pp.279-296, 2018, Nouvelles voies d'accès au changement linguistique. Actes du colloque de la SIDF, 10.15122/isbn.978-2-406-06946-1.p.0279 . halshs-01313822

HAL Id: halshs-01313822

https://shs.hal.science/halshs-01313822

Submitted on 10 May 2016

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.

Céline Guillot-Barbance, Alexei Lavrentiev, Serge Heiden & Bénédicte Pincemin
UMR 5191 ICAR / ENS de Lyon, CNRS, Université de Lyon

Diachronie de l'oral représenté : délimitation et segmentation interne du dialogue (IXe-XVe siècle)[1]

INTRODUCTION

Notre recherche s'insère dans une série de travaux sur ce qu'à la suite de Christiane Marchello-Nizia (2012) on a appelé l'« épisode d'oral représenté » et dont on cherche à étudier les marques linguistiques et formelles dans les textes médiévaux français. Une première série de travaux (Guillot *et al.* 2013) a porté sur la fréquence des unités linguistiques qui caractérisent ce type de discours et lui donnent des spécificités très fortes. On s'intéresse dans le cadre de cet article à la façon dont ces épisodes d'oral représenté se structurent et se délimitent du reste du texte, et plus largement à la manière dont l'écrit (manuscrit) en langue vernaculaire développe son propre système de balises avant qu'apparaissent l'imprimé et les marques typographiques modernes (guillemets, tirets).

L'étude a été menée sur un vaste ensemble d'éditions de référence de textes composés entre le IXe siècle et la mi XVe siècle, de genres discursifs diversifiés (4,7 millions de mots). Ce corpus est accessible sur le portail de la *Base de français médiéval* (http://txm.bfm-corpus.org) et peut être interrogé en ligne grâce à la plateforme d'analyse TXM. Il permet de se faire une idée assez précise de la délimitation et de la segmentation interne du dialogue et de son évolution du début à la fin de la période médiévale. Les données que nous avons exploitées, nos hypothèses et notre méthodologie de recherche sont exposées dans la section suivante. La seconde et dernière section est consacrée à l'analyse des observations et aux résultats de l'étude.

1. CADRE ET HYPOTHÈSES DE RECHERCHE

La méthodologie adoptée pour cette étude repose sur une démarche expérimentale et outillée. Les hypothèses sur lesquelles s'appuie notre recherche ont été élaborées à partir de la littérature existante. La constitution et l'exploitation d'un corpus adapté a ensuite permis de confronter aux textes et de valider ces hypothèses de départ. Nous commencerons par présenter ci-dessous notre définition de la notion d' « épisode d'oral représenté » et de quelques éléments connexes, avant de détailler la démarche suivie dans cette étude.

1.1. DÉFINITIONS : ÉPISODE D'ORAL REPRÉSENTÉ, ANNONCE, RAPPEL ET INCISE

L'épisode d'oral représenté (OR) se définit comme une séquence discursive se donnant comme une restitution à l'écrit de paroles prononcées oralement. Elle se compose d'une ou de plusieurs prises de parole (PP) qui s'enchaînent et qui constituent les unités internes de l'épisode. Dans la définition que nous adoptons ici, l'épisode d'oral représenté comprend également les éléments qui en précisent les limites externes et les articulations internes et que nous appellerons, à la suite de Marchello-Nizia, l'*annonce*, le *rappel* et l'*incise* :

> Aussi proposerons-nous de considérer comme une *unité linguistique* l'ensemble constitué par le discours direct (prise de parole unique, ou série de prises de parole enchaînées) et les trois éléments qui l'articulent au récit, de nommer cette unité « épisode d'Oral représenté » et de renommer le DD [discours direct] comme objet linguistique Oral représenté (Marchello-Nizia 2012, p. 248)[2].

[1] Les auteurs remercient le LABEX ASLAN (ANR-10-LABX-0081) de l'Université de Lyon pour son soutien financier dans le cadre du programme "Investissements d'Avenir" (ANR-11-IDEX-0007) de l'Etat Français géré par l'Agence Nationale de la Recherche (ANR).

[2] On voit que l'épisode d'oral représenté ne coïncide pas parfaitement avec le discours direct, c'est pourquoi nous utiliserons

L'annonce signale la limite initiale de l'épisode et comporte généralement un *verbum dicendi*. Le rappel en marque la limite finale et le retour au récit[3]. Le troisième élément qui sert également de marque de balisage est l'incise. Construite autour d'un *verbum dicendi*, l'incise partage avec l'annonce une double fonction : elle joue le rôle d'élément de structuration (on verra qu'elle marque soit le début de l'épisode, soit le changement de locuteur) et elle précise l'identité de l'émetteur des paroles prononcées.

D'autres unités linguistiques et graphiques (signes de ponctuation) jouent un rôle similaire et s'ajoutent souvent à ces trois éléments : termes d'adresse, interjections, marques d'exclamation, etc. Mais l'annonce, le rappel et l'incise semblent dès le départ spécialisés dans cette fonction organisationnelle et sont particulièrement fréquents dès les premiers textes (quoique dans des proportions qui varient au cours du temps). Ces trois éléments se distinguent également en ce qu'ils partagent avec le reste de l'épisode un certain nombre de caractéristiques communes : sur-représentation des verbes de parole et du présent de l'indicatif, fréquence de l'expression du sujet, nature sémantique du sujet (humain). C'est ce qui explique qu'on les intègre pleinement à l'épisode d'OR.

Notre étude se centre sur l'incise, ses caractéristiques formelles et fonctionnelles, les relations qu'elle entretient avec l'annonce et leur évolution conjointe au cours de la période médiévale. Comme on l'a indiqué plus haut, l'incise et l'annonce partagent dès le début une double fonction (délimitation d'une unité et identification du locuteur) et comportent toutes deux un *verbum dicendi*. Le tableau suivant fait toutefois apparaître les différences morphosyntaxiques qui permettent de les distinguer. Ce tableau rend compte de la physionomie majoritaire de l'annonce et de l'incise à partir de la mi XIIe siècle, c'est-à-dire à partir du moment où elles se dissocient plus nettement l'une de l'autre :

ANNONCE	INCISE
verbum dicendi (éventail large, absence du verbe *faire*)	*verbum dicendi* (éventail réduit, verbe *faire* possible)
expression du sujet facultative	expression du sujet obligatoire
ordre SV ou VS	ordre VS
désignation de l'allocutaire fréquente	désignation de l'allocutaire très rare
objet très rare	objet possible
place fixe	place variable
ponctuation possible dans le manuscrit	pas de ponctuation dans le manuscrit

TABLEAU 1 : Grandes tendances de l'incise et de l'annonce à partir de la mi *XIIe* siècle

Il ressort de ce tableau que l'annonce est plus libre dans sa composition (dans le choix des verbes, dans l'ordre relatif des éléments ou dans l'expression du sujet) mais plus contrainte dans sa position (placée au début de l'épisode). L'analyse d'un corpus diachronique montre cependant qu'avant la mi XIIe siècle la frontière entre annonce et incise est parfois très ténue. L'écrit en langue vernaculaire se développant au fil du temps, ces deux structures tendent à se distinguer formellement et à se spécialiser fonctionnellement.

1.2. TYPES D'ANNONCE ET D'INCISE ET POSITION DANS L'ÉPISODE D'ORAL REPRÉSENTÉ

L'observation attentive des données permet de détailler différentes sous-catégories d'annonces

toujours le terme d'oral représenté dans la suite de cet article.

[3] Tel est le rôle, par exemple, d'un grand nombre de subordonnées temporelles antéposées : « Aprés parlat ses filz envers Marsilies, / E dist al rei : "Guenes ad dit folie. / Tant ad erret nen est dreiz que plus vivet. / Livrez le mei, jo en ferai la justice." / *Quant l'oït Guenes*, l'espee en ad branlie (*Roland* v.49*5-499)* « Ensuite son fils s'adressa à lui et dit au roi : "Ganelon a dit des paroles insensées. Il en a tant fait qu'il ne doit pas vivre plus longtemps. Livrez-le-moi, j'en ferai justice." Quand Ganelon l'entendit, il brandit son épée.

et d'incises et l'on fera l'hypothèse ici que la répartition et la fréquence de ces types évoluent temporellement.

Le schéma suivant présente les différents éléments pouvant être distingués dans la structure d'un épisode d'OR composé de plusieurs prises de parole (PP_1 à PP_n). Il donne une vue d'ensemble de tous les éléments possibles, même si tous se rencontrent rarement simultanément dans un même épisode[4].

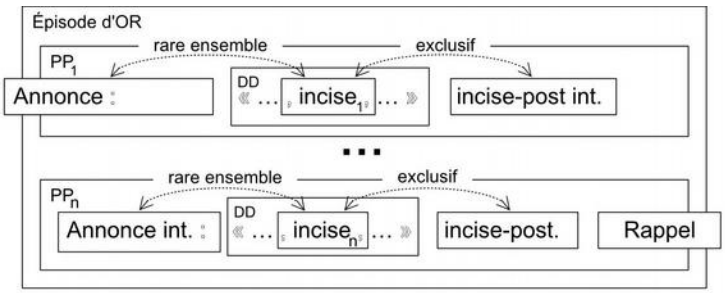

FIGURE 1 : Éléments structurant un épisode d'oral représenté constitué de plusieurs prises de parole.

Ce schéma représente l'*annonce,* l'*incise* et le *rappel*, ainsi que les PP et les passages au discours direct (DD). Il apporte également des précisions sur la typologie des incises et annonces, et sur la position que ces éléments peuvent occuper au sein de l'épisode. Un premier cas particulier est celui où l'incise se trouve à la fin d'une PP. Nous utilisons le terme *incise-post* pour désigner ce type d'incise (exemple 1, l'incise-post est en italiques).

(1) « Sire, sire, vos ne savez, / .II. ostes chevaliers avez. / – Dex an soit aorez », *fet il.* (*Charrette*, p. 36, v. 2549-2551)

« Seigneur, seigneur, vous ne savez pas, vous avez deux chevaliers pour hôtes. – Dieu soit loué », dit-il.

Dans les cas où une nouvelle PP suit immédiatement l'incise-post, cette incise est dans une position interne à l'épisode (elle ne le clôt pas). On parlera alors d'*incise-post interne* (exemple 2) :

(2) « [...] – Et don venez vos or ensi ? » / *fet Percevax.* « Sire, de ci, / d'un boen home, d'un saint hermite, / qui an ceste forest abite / qui ne vit, tant par est sainz hon, / se de la gloire de Deu non. » (*Perceval*, p. 384, v. 6090-6095)

« [...] – Et d'où venez-vous ainsi ? », dit Perceval. « Seigneur, d'ici, de chez un homme bon, un saint ermite, qui habite dans cette forêt, qui ne vit, tant il est saint, que de la gloire de Dieu. »

Une annonce peut également se trouver au début d'une PP au milieu d'un épisode d'OR (et non à la limite initiale de l'épisode). Nous utilisons dans ce cas le terme d'*annonce interne* (exemple 3)[5] :

(3) « [...] – Dans chevaliers, il est trop tart, / Fait li noirs chevaliers oscur ; / Mais soiés trestout aseür / Que je vendrai le matin ça. » / *Dist Gavains* : « Ce n'avendra ja, / Se Dieu plaist, que nos departons / Desi que je savrai ou non / Se je en merrai les cevals / Et la pucele et les vassaus, / Sans calonge et sans contredit ». (*Âtre périlleux*, p. 195, v. 6143-6148)

« [...] – Seigneur chevalier, il est trop tard, dit le chevalier noir et sombre ; mais soyez assuré que je viendrai ici demain matin. » Gauvain dit : « S'il plaît à Dieu, il n'arrivera pas que nous nous séparions

[4] Ce schéma n'est pas exhaustif. Il ne rend pas compte, notamment, de la possibilité de trouver plusieurs incises au sein du discours d'un même locuteur. On verra plus loin aussi que notre étude porte sur les incises qui se trouvent à proximité du début d'une PP, et pas celles qui figurent plus loin dans la PP. Ce choix a été motivé par des contraintes techniques (difficulté à repérer les incises éloignées), par le fait que ces dernières sont en nombre beaucoup plus réduit que celles qui nous intéressent ici, et par notre souci de corréler l'étude des incises à celle des annonces (placées à proximité, avant la première PP).

[5] Christiane Marchello-Nizia (2012) utilise dans ce cas le terme d'*annonce-2* et considère que celle-ci introduit un nouvel épisode d'OR. Nous préférons nous baser sur des critères purement formels et considérer qu'un épisode OR continue tant qu'il n'y a pas de reprise du récit entre les prises de parole successives.

avant que je sache oui ou non si j'emmènerai les chevaux, la jeune-fille et les chevaliers sans dispute ni contestation. »

Les incises ordinaires et les incises-post sont en principe mutuellement exclusives, mais l'annonce et l'incise peuvent se combiner (exemple 4), même si cela devient relativement rare à partir d'une certaine époque :

(4) Lors sel *comance a pleidoier* : / « Viax tu donc, *fet ele*, noier / Que par toi ne soit morz mes sire? / – Ce, *fet il*, ne puis je desdire, / Einz l'otroi bien. » (*Yvain*, p.86, v.1761-1765)

Elle commence alors à mener les débats : « Veux-tu donc nier, dit-elle, que mon mari soit mort par ta faute ? – Je ne peux dire le contraire, dit-il, je l'avoue sans réserve. »

Notre schéma inclut également des marques de ponctuation (le deux-points, les guillemets et les virgules autour des incises). Ce sont ces marques de ponctuation introduites par les éditeurs de textes qui nous ont permis, en combinaison avec les formes des verbes de parole, de repérer les occurrences pertinentes dans notre corpus. Nous verrons dans la suite de l'article que les types introduits dans ce schéma (incise-post, incise-post interne et annonce interne) sont relativement marginaux dans les textes. Mais leur succession temporelle nous renseignera sur l'évolution générale de l'incise et de l'annonce et sur la tendance à les distinguer de plus en plus nettement l'une de l'autre.

1.3. Hypothèses de recherche

Les études antérieures sur lesquelles nous nous fondons (Marchello-Nizia 2012 et à par ., Guillot *et al.* 2014) ont permis de dégager plusieurs hypothèses sur la diachronie de l'incise en français médiéval.

L'étude fouillée d'un corpus de six textes composés entre la mi XIIe et la mi XVe siècle a révélé une série de changements orientés, qui tendent à dissocier deux périodes chronologiques (ancien et moyen français) :

- du point de vue formel : l'incise se caractérise par la présence du verbe *faire* en ancien français et par d'autres verbes (notamment *dire*) en moyen français ; le verbe est d'abord fléchi au présent de l'indicatif puis au passé simple ; le sujet de l'incise, majoritairement pronominal (pronom personnel de 3e personne) en ancien français, cède la place au syntagme nominal (SN) et au nom propre (NP) en moyen français.
- du point de vue fonctionnel : l'incise s'utilise au XIIe siècle au début de l'épisode d'OR (elle se combine alors éventuellement avec l'annonce) aussi bien qu'au moment du changement de locuteur ; à la fin du Moyen Âge, elle semble s'être spécialisée dans le marquage de l'alternance des PP ; elle se trouve rarement en début d'épisode et ne se combine plus avec l'annonce ; l'expression du sujet sous forme nominale (SN ou NP) lève tout risque d'ambiguïté sur l'identité du locuteur des paroles rapportées et pourrait être corrélé avec le rôle acquis par l'incise dans l'organisation interne du dialogue représenté.

Ces différents critères de variation seront repris ici (à l'exception des temps verbaux[6]) et mis à l'épreuve d'un corpus chronologiquement plus étendu, en particulier pour la période la plus ancienne du français, quantitativement plus important et typologiquement plus varié (même s'il est globalement limité aux textes narratifs, qui sont pour l'essentiel ceux dans lesquels se trouvent les épisodes d'OR). L'approche empirique suivie dans cette recherche permettra *in fine* de vérifier les hypothèses de départ, de les préciser sur certains points et de proposer un

[6] L'étude comparative des temps verbaux n'a pu être menée sur tous les textes du corpus en raison de l'ambiguïté de certaines formes (notamment la graphie *dist*, qui peut s'analyser comme un présent ou un passé simple) et en l'absence d'étiquetage morphologique fiable du corpus.

modèle de l'évolution de l'incise en trois périodes successives.

1.4. Méthodologie et corpus

Le corpus utilisé pour cette recherche est le corpus principal de la Base de français médiéval dans son état de décembre 2013 (BFM2013[7]), composé de 142 textes intégraux (du IXe au XVe siècle), soit près de 4,7 millions de mots. La liste complète des textes disponibles dans la version courante de la BFM et leur répartition selon les différentes variables (date de composition, domaine, genre, dialecte, etc.) peuvent être obtenues en ligne grâce à l'outil « Création de sous-corpus de textes » du portail de la BFM (http://txm.bfm-corpus.org)[8].

Nous n'avons pas exploité le balisage du discours direct dont bénéficie la majorité des textes de la BFM, contrairement à ce que nous avions fait dans nos études précédentes, mais avons repéré les occurrences des annonces et des incises sur l'ensemble des textes en utilisant la requête de recherche CQL suivante :

```
[fropos="PON.*"] [word="[fd][aei].?t|resp[ou]n.*"%c]
```
[9]

Cette requête repose sur le principe que les incises et les annonces commencent par un verbe de parole (*dire, faire* ou *respondre*) précédé d'une ponctuation dans l'édition de référence. Cette ponctuation normalisée sert de support au repérage des incises mais n'est pas analysée en tant que telle. La requête a retourné 9 796 résultats provenant de 124 textes (sur 142)[10] de la BFM2013. Certains types d'incises nettement moins fréquents (notamment la formule *ce dit X*, environ 750 occ.) n'ont pas été pris en compte dans la requête et pourront faire l'objet d'une étude ultérieure.

Les résultats de la requête, sous forme de concordance, ont été exportés en format tabulaire, puis adaptés et pré-annotés à l'aide d'expressions régulières, avant d'être analysés et catégorisés manuellement à l'aide du tableur Libre Office Calc.

À titre d'exemple, l'expression régulière suivante a été utilisée pour pré-annoter les incises placées en début d'épisode et précédées d'une annonce :

```
(: ["'«»]( [^ \t]+){0,10}\t([.,!:;?"»«'\)\(-]+)\t.+)$
```

Cette expression indique qu'on recherche un deux-points suivi d'une des variantes typographiques du guillemet, suivi d'entre 0 et 10 mots ou ponctuations dans les contextes gauches de la concordance.

Les autres pré-annotations d'occurrences d'incises ont concerné les débuts d'épisode d'OR sans annonce et les changements de PP dans les dialogues.

La sélection des occurrences et leur pré-annotation ont permis de centrer l'étude sur les incises placées au début de l'épisode d'OR (dans un espace de 10 mots maximum après la limite initiale), ou au moment de l'alternance des PP. Comme on l'a signalé plus haut, les incises plus éloignées à l'intérieur d'une PP (pas plus de 500 occurrences, selon nos estimations) ont été éliminées suite à l'opération de pré-annotation.

Avant de procéder à la catégorisation de chaque occurrence d'annonce et d'incise, nous avons éliminé les occurrences non pertinentes, ou le « bruit » généré par la requête (par exemple, les verbes de parole qui n'étaient ni des annonces ni des incises). Nous avons également exclu les textes où le taux de « bruit » était trop important et ceux où le nombre

[7] http://bfm.ens-lyon.fr/article.php3?id_article=327.
[8] Voir notre documentation en ligne (Bertrand *et al.*, p. 13-19).
[9] Dans son état actuel le balisage du discours direct dans la base ne permet pas de repérer de façon sûre et homogène l'alternance des prises de parole. Nous avons donc centré la recherche sur les incises et annonces, et non sur les contextes d'occurrence (début d'épisode d'OR ou alternance de PP).
[10] Ces textes représentent 4,57 millions de mots, soit 97% de la BFM2013.

d'occurrences pertinentes était trop faible. Après cette opération de nettoyage, 7 100 occurrences provenant de 39 textes[11] ont été retenues pour la suite de l'analyse. Pour chaque occurrence, nous avons annoté le lemme du verbe de parole et la nature du sujet (SN, NP, pronom personnel, démonstratif ou autre type de sujet). Ces annotations ont permis de générer des tableaux d'analyse croisée que nous avons ensuite exploités pour la confrontation aux hypothèses initiales.

2. Observations et analyses

Nous reviendrons dans une première partie sur le typage des incises et annonces esquissé précédemment (section 1.2) et sur l'évolution des trois sous-types dont nous avons déjà signalé qu'ils étaient marginaux. Nous discuterons dans un second temps, critère par critère, l'évolution diachronique de l'incise « canonique ».

2.1. Sous-types d'incises et d'annonces en diachronie

L'examen détaillé du corpus permet de se rendre compte que les sous-types d'incises et d'annonces ne se rencontrent pas de manière uniforme à toutes les époques. Nous essaierons de montrer ici que leur succession n'est pas aléatoire, mais qu'elle répond à un principe général de dissociation formelle et fonctionnelle, dont on verra dans la section suivante qu'il peut également être mis en relation avec l'évolution de l'incise et de l'annonce « canoniques » (ni « post », ni « internes »).

Dans la période la plus ancienne du français (jusqu'à la mi XIIe), un petit groupe de textes comporte un nombre relativement élevé d'annonces internes (*Roland* 34 occ., *Couronnement de Louis* 41 occ., *Chanson de Guillaume* 39 occ.) :

(5) Ço dist Rollant : « Cornerai l'olifant, / Si l'orrat Carles, ki est as porz passant. / Jo vos plevis, ja returnerunt Franc. » / *Dist Oliver* : « Vergoigne sereit grant / E reprover a trestuz voz parenz ; » (*Roland*, p. 136, v. 1702-1706)

Roland dit : « Je cornerai l'olifant, Charles, qui passe les ports, l'entendra. Je vous le garantis, les Francs reviendront. » Olivier dit : « Ce serait un grand déshonneur et un blâme pour tous vos parents »

On observe que tous ces textes sont des chansons de geste versifiées, textes dont le caractère formulaire ne fait pas de doute. Ce que nous voudrions souligner ici, c'est que cette forme d'annonce présente un grand nombre de similarités avec l'incise : ordre VS (qui deviendra rapidement un ordre très marqué et qui est dès cette époque assez rare), expression du sujet obligatoire. Ces annonces ne se situent pas au début de l'épisode d'OR, comme c'est habituellement le cas, mais devant une nouvelle PP. Elles marquent donc le changement de locuteur, comme le feront peu après les incises. Deux textes composés à la fin du XIIe siècle en fournissent encore un nombre relativement élevé d'occurrences (*Quatre livre des rois*, 19 occ., *Escoufle* 29 occ.), mais les autres textes n'utilisent cette structure que de manière très sporadique (1 à 5 occ.).

Le sous-type de l'incise-post (interne) suit, quant à lui, l'évolution inverse. Très peu usité avant la fin du XIIe siècle, il fait sa véritable apparition dans deux textes de Chrétien de Troyes (*Charrette* 16 occ., *Perceval* 13 occ., *cf.* l'exemple 1 cité plus haut).

Il reste assez marginal pendant toute la période de l'ancien français : les trois textes qui l'utilisent le plus sont la *Queste del saint Graal* (11 occ. contre 518 incises canoniques), la

[11] Voici les sigles de ces textes : alexis, artu, atrper, becket, belinc, BenDuc1b, brut2, CharretteKu, ChronSMichelBo, clari, CligesKu, cnn, eneas1, eneas2, eracle, ErecKu, escoufle, floire_jl, froissart1, galeran, galeron, gcoin, gormont, jehpar, jouvencel, louis, menreims, merlin_suite_litt, moree, PercevalKu, qgraal_cm, qjm, roland, saintre, SGenPr1, slethgier, thebes, trispr, vergy, YvainKu. Les références complètes sont consultables sur le portail de la BFM : http://txm.bfm-corpus.org/files/guillot_al_sidf14_liste-des-textes.pdf.

Mort Artu (10 occ. contre 710 incises canoniques) et la *Suite du Merlin* (16 occ. contre 750 incises canoniques), et c'est dans le texte le plus récent du corpus, les *Cent nouvelles nouvelles* (daté de 1462), qu'il est relativement fréquent (40 occ.).

(6) « [...] – Creez, que si feray je », *dit elle*. Et a ces motz se seigne et s'en va. (*Cent nouvelles nouvelles*, p. 187)

« [...] – Croyez que je le ferai », dit-elle. Et à ces mots, elle se signe et s'en va.

Ces quelques éléments nous paraissent confirmer, d'une part, que l'annonce et l'incise connaissent d'importants changements au cours de la période médiévale, et d'autre part, que ces changements tendent à une plus nette dissociation de ces deux éléments. Nous allons voir à présent de quelle façon cette-même tendance se manifeste au travers de l'évolution fonctionnelle et formelle de l'incise « canonique ».

2.2. Évolution de l'incise canonique en diachronie

Notre analyse repose sur trois mesures caractéristiques de la variation des incises.

La première mesure nous renseigne sur la fonction principale de l'incise. Elle a été obtenue en calculant le rapport pour chaque texte entre le nombre d'incises placées au début d'une nouvelle PP (marquant le changement de locuteur) et celui des incises placées au début d'un épisode d'OR.

L'évolution diachronique de ce rapport est visualisée à travers un graphique synthétique (fig. 2), dans lequel chaque point représente un texte. Les textes symbolisés par un point plus petit et carré sont ceux dans lesquels le phénomène décompté est peu représenté (moins de 20 occurrences) : pour ces textes, l'interprétation doit être plus réservée, l'analyse se construisant d'abord à partir des autres points. L'axe horizontal du graphique est l'axe chronologique : les points à gauche représentent les textes les plus anciens, ceux à droite les plus récents. L'axe vertical représente quant à lui l'équilibre entre les deux grandeurs comparées : les points-textes en partie haute (au-dessus du trait placé à 1) sont ceux dans lesquels domine le numérateur du rapport, à savoir ici l'incise marquant le changement de locuteur, et ceux en partie basse sont ceux pour lesquels est majoritaire l'autre grandeur, ici l'incise placée au début d'un épisode d'OR. La concentration de points dans une zone haute ou basse du graphique peut donc se lire comme la dominance de l'une ou l'autre des deux fonctions de l'incise à une période donnée[12].

[12] Deux compléments techniques sur la construction des graphiques : 1) L'axe vertical utilise une échelle logarithmique, pour une double raison : cela optimise l'utilisation de l'espace, mais aussi et surtout, l'échelle logarithmique produit une représentation symétrique de part et d'autre de la ligne d'équilibre (au niveau du 1, où les deux grandeurs sont égales), ce qui permet de rendre compte du côté arbitraire de l'orientation du rapport. 2) Lorsque l'une des deux grandeurs mises en rapport vaut zéro, la position théorique du point dans le graphique est à l'infini. Pour pouvoir prendre en considération ces points dans notre analyse, nous avons remplacé les effectifs 0 par des 0,5 pour le calcul du rapport.

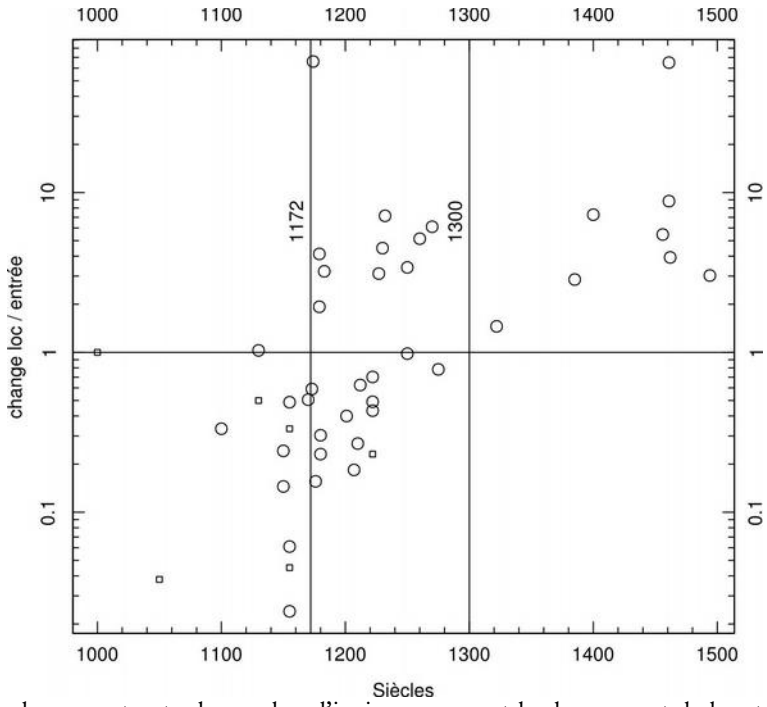

FIGURE 2 : Évolution du rapport entre le nombre d'incises marquant le changement de locuteur et le nombre d'incises placées au début d'un épisode d'OR.

Ainsi, on peut observer que dans la partie gauche du graphique (délimitée par une barre verticale) le rapport est favorable à l'entrée dans l'épisode d'OR, et que dans la partie droite (délimitée par une autre barre) il est favorable au changement de locuteur. Dans la partie intermédiaire, les textes se situent de part et d'autre de la zone centrale (ligne horizontale)[13]. Le graphique montre une évolution assez nette du rapport étudié, les points de bascule se situant aux alentours des années 1180 et 1300. Le graphique ne doit pas donner à penser qu'aux extrémités gauche (XIe et début XIIe siècle) et droite (XIVe-XVe) l'incise aurait une fonction unique. Mais les points étant relativement distants de la ligne horizontale (qui traduit un rapport de 1), la fonction principale de l'incise semble se dégager assez nettement dans ces deux zones. La partie intermédiaire représente la période de fluctuation ou de changement, la plupart des textes optant pour une fonction principale sans se conformer pour autant à une règle commune.

La seconde mesure (illustrée par la fig. 3) nous renseigne sur la nature du sujet de l'incise (tout type confondu). Elle établit le rapport entre les sujets pronominaux (pronoms personnels et autres) et les sujets nominaux (y compris le nom propre)[14]. Des barres verticales permettent à nouveau de distinguer trois grandes phases. La période la plus ancienne (antérieure à 1140) se caractérise par une majorité de sujets nominaux. La seule exception (*La Vie de Saint Alexis*) présente trop peu d'occurrences de l'incise (15) pour pouvoir chercher ici une généralisation. Une zone intermédiaire, qui court de 1140 à 1300, marque une préférence pour les sujets pronominaux. Mais la tendance n'est vraiment nette que pour les années 1140 à 1210 environ, la zone comprise entre 1210 et 1300 apparaissant comme une zone de fluctuation. Quant à la zone postérieure à 1300, elle penche nettement du côté du sujet nominal. Deux textes (*Les Quinze joies de mariage* (qjm dans le graphique) et *Les Cent nouvelles nouvelles* (cnn)) échappent à cette tendance. Cette particularité peut sans doute s'expliquer par le genre de ces textes : il s'agit de récits brefs comportant un nombre restreint de personnages, souvent nommés par leur type (mari / femme) et identifiables dans les

[13]On note que, du fait de nos données, les points sont par construction plus nombreux dans cette partie centrale (plus de textes pour la période des XIIe-XIVe siècles) que dans les marges gauche et droite.

[14] Le sujet nul en incise est extrêmement rare (7 occurrences identifiées dans la BFM2013).

dialogues par leur sexe, ce qui favorise l'emploi des pronoms personnels sujets au détriment des noms propres et des syntagmes nominaux.

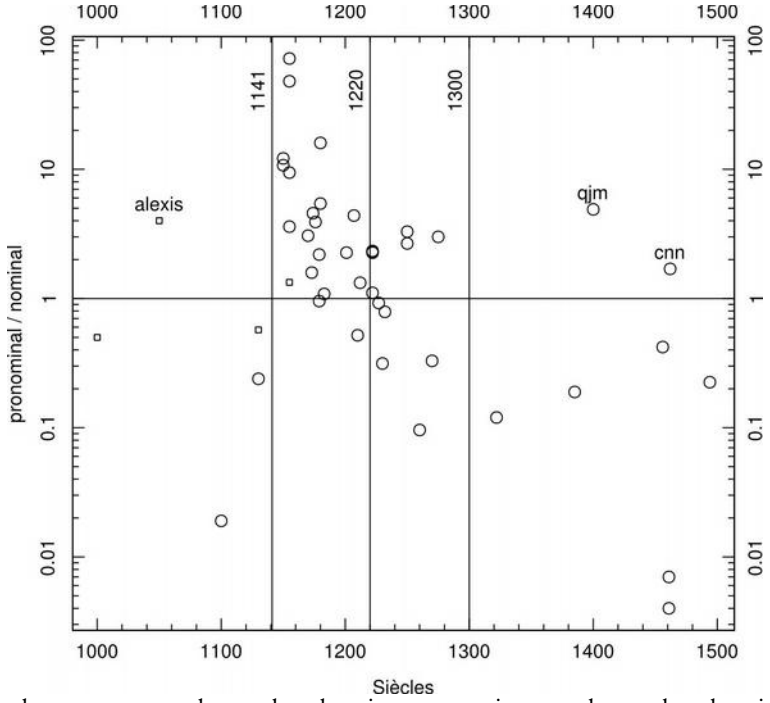

FIGURE 3 : Évolution du rapport entre le nombre de sujets pronominaux et le nombre de sujets nominaux dans les incises

La dernière mesure (illustrée par la fig. 4) nous renseigne sur la nature du verbe de l'incise. Elle calcule le rapport entre les usages du verbe *dire* et ceux du verbe *faire* dans chaque texte. Les trois zones observées précédemment apparaissent à nouveau. Avant 1150, c'est le verbe *dire* qui domine l'incise. Entre 1150 et 1310 *faire* est très nettement majoritaire, et à partir de 1310 le verbe *dire* reprend le dessus. Concernant les points qui s'écartent de ces tendances, nous n'avons pas pour le moment d'explication à la dominance de *faire* dans *Le Ménestrel de Reims* (menreims) et de *dire* dans *Les Quinze joies de mariage.* La *Vie de Saint Léger* (slethgier) ne compte que 3 incises et la position du *Brut* de Wace (brut2) dans le graphique s'explique par sa proximité temporelle avec la première zone, dont il adopte les caractéristiques.

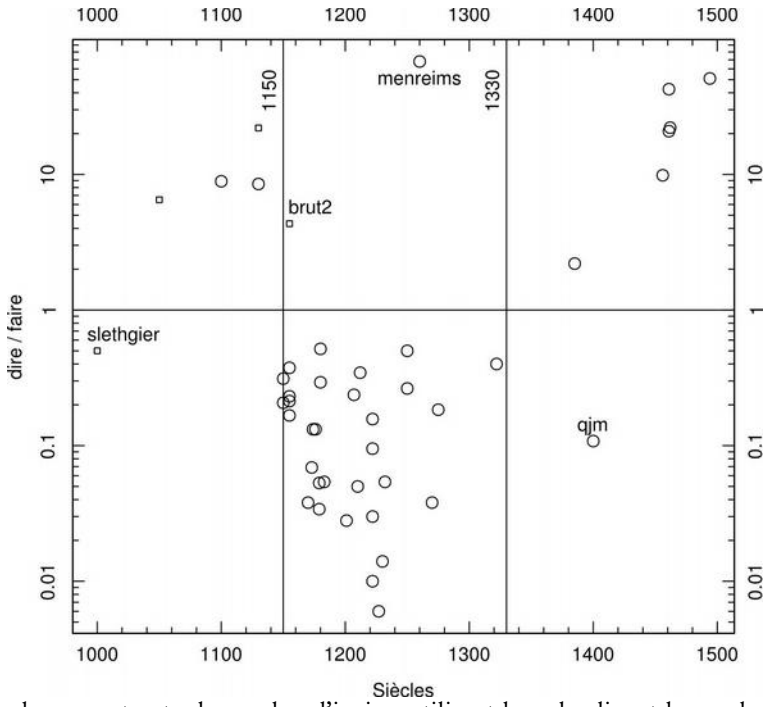

FIGURE 4 : Évolution du rapport entre le nombre d'incises utilisant le verbe dire et le nombre d'incises utilisant le verbe faire

Les trois mesures qui ont permis l'analyse diachronique de l'incise canonique permettent de confirmer les hypothèses exposées en 1.3. Elles permettent également de caractériser la période antérieure à la mi XIIe siècle (que nous appellerons « très ancien français »). Trois grandes périodes se dégagent ainsi, et nous récapitulons dans la section suivante ce qui fait leur spécificité.

2.3. SYNTHÈSE : UN MODÈLE DE L'ÉVOLUTION CONJOINTE DE L'INCISE ET DE L'ANNONCE EN TROIS PÉRIODES

L'analyse d'un grand nombre de textes s'échelonnant des origines du français (IXe siècle) à l'extrême fin du Moyen Âge (XVe siècle) nous permet de proposer une première version du modèle d'évolution de l'incise en français médiéval.

Phase n°1 : le très ancien français (IXe – mi XIIe siècle)
C'est au cours de cette période que la frontière formelle entre annonce et incise semble se préciser, du fait notamment de la raréfaction des annonces internes, formellement très proches de l'incise (ordre VS). L'incise se caractérise alors par la présence du verbe *dire* (qui est également le verbe majoritaire de l'annonce) et du sujet nominal. Du point de vue fonctionnel, incise et annonce semblent remplir un rôle assez proche, puisqu'elles indiquent majoritairement le début de l'épisode d'OR.

Phase n°2 : l'ancien français (mi XIIe – 1300)
La séparation formelle de l'incise et de l'annonce s'accentue. L'incise se caractérise très nettement par l'emploi du verbe *faire*, exclu de l'annonce. Le sujet pronominal devient majoritaire, au début de la période surtout (entre la mi XIIe et le début du XIIIe siècle). Du point de vue fonctionnel, l'incise se démarque peu à peu de l'annonce, puisqu'elle commence à s'employer dans certains textes plutôt au moment du changement de locuteur. Mais sa fonction reste globalement fluctuante, contrairement à celle de l'annonce, déjà spécialisée dans le marquage du début de l'épisode d'OR (disparition de l'annonce interne).

Phase n°3 : le moyen français (1300 – 1500)
L'incise s'est nettement spécialisée du point de vue fonctionnel. Elle indique préférentiellement le changement de locuteur et ne se combine plus à l'annonce. Son sujet est majoritairement nominal, ce qui permet une plus sûre identification du locuteur du tour de parole. Le verbe *dire* domine à nouveau.

Pour vérifier et étayer ce modèle d'évolution, il serait utile de disposer d'informations complémentaires (que nous n'avons pu obtenir faute de temps et pour des raisons techniques). En particulier, on aimerait pouvoir montrer que l'incise et l'annonce se généralisent au fil du temps, l'une dans le marquage du changement de locuteur, l'autre dans celui de l'entrée dans l'épisode d'OR. On aimerait ainsi connaître le rapport entre les changements de PP comportant une incise et les changements qui n'en comportent pas, et le rapport entre les débuts de séquence comportant une annonce et les débuts qui n'en comportent pas. Si les fonctions de l'incise et de l'annonce se spécialisent au cours du temps, on peut supposer que ces deux éléments deviennent de plus en plus nécessaires à l'organisation du dialogue représenté.

Pour clore cette section, nous ferons une remarque complémentaire. La part des dialogues semblant croître au fil du temps dans les textes narratifs (Marchello-Nizia 2012 et à par.), on peut supposer que le rapport entre incises marquant le changement de locuteur et incises marquant l'entrée dans l'épisode se trouve *de facto* modifié par cette évolution. Mais on peut imaginer aussi que c'est parce qu'il devient de plus en plus nécessaire de mettre en scène le dialogue oral au sein du texte écrit que s'élabore peu à peu un système de marques de balisage spécialisées.

CONCLUSION

Notre recherche visait à décrire aussi précisément que possible le système des marques linguistiques qui structurent les épisodes d'OR aux tout débuts du français et les transformations de ce système jusqu'à l'orée de l'imprimé. Dans cette optique, nous avons tenté de décrire de manière assez précise l'évolution que connaît l'incise en français médiéval, le rôle de premier plan qu'elle acquiert peu à peu dans le marquage de l'alternance des tours de parole et le rôle complémentaire de l'annonce.

Les résultats de cette étude vont à l'encontre de l'idée plus ou moins intuitive que la représentation du dialogue irait vers davantage de fluidité à mesure qu'on avance dans le temps. On observe en particulier que c'est dans la période antérieure aux années 1180 que le changement de locuteur n'est pas associé à une marque spécialisée et l'on a déjà constaté que le *Roman d'Eneas* (composé vers 1155), par exemple, enchaînait souvent plusieurs PP sans indicateur explicite d'alternance (ni incise, ni même terme d'adresse, etc.). Il est vrai qu'il s'agit d'un texte en vers (comme la plupart des œuvres composées à cette période), et que les débuts de PP coïncident souvent avec la structure métrique. Les critères de forme (vers/prose) et/ou génériques jouent certainement un rôle important dans l'évolution de l'épisode d'OR. L'apport principal de cette étude est de montrer que le développement de l'écrit en langue vernaculaire et l'essor de la prose semblent peu à peu faire émerger un système de marques linguistiques cohérentes et spécialisées dans la délimitation externe et la segmentation interne des épisodes d'OR. La description de ce système nous renseigne à la fois sur les caractéristiques linguistiques de cette représentation de l'oral et sur les moyens mis en œuvre par l'écrit pour se structurer en unités internes.

BIBLIOGRAPHIE

BERTRAND, Lauranne, GUILLOT, Céline, HEIDEN, Serge, LASCAR, Justine, LAVRENTIEV, Alexei, PINCEMIN, Bénédicte (2014) *Tutoriel TXM pour la BFM,* Version 2.0, Lyon, ENS de Lyon / UMR ICAR, [En ligne : http://txm.bfm-corpus.org/files/Tutoriel_TXM_BFM_V1.pdf].

GUILLOT, Céline, LAVRENTIEV, Alexei, PINCEMIN, Bénédicte, HEIDEN, Serge (2013) « Le discours direct au Moyen Âge : vers une définition et une méthodologie d'analyse », in D. LAGORGETTE et P. LARRIVÉE (édit.), *Représentations du sens linguistique 5*, Chambéry, Presses de l'université de Savoie, p. 17-41. [Version auteur disponible sur HAL-SHS : http://halshs.archives-ouvertes.fr/halshs-00820262]

GUILLOT, Céline, HEIDEN, Serge, LAVRENTIEV, Alexei, PINCEMIN, Bénédicte (2015) « L'oral représenté dans un corpus de français médiéval (9e-15e) : approche contrastive et outillée de la variation diasystémique, in K. JEPPESEN KRAGH et J. LINDSCHOUW (édit.), *Les variations diasystémiques et leurs interdépendances dans les langues romanes. Actes du Colloque DIA II à Copenhague (19-21 nov. 2012)*, Strasbourg, Editions de linguistique et de philologie, p. 15-27 [Version auteur disponible sur HAL-SHS : http://halshs.archives-ouvertes.fr/halshs-00760647].

GUILLOT, Céline, PRÉVOST, Sophie, LAVRENTIEV, Alexei (2014) « Oral représenté et diachronie : étude des incises en français médiéval », *Actes du CMLF 2014 – 4e Congrès mondial de linguistique française*, Berlin, 19-23 juillet 2014, EDP Sciences (www.linguistiquefrancaise.org), p. 259-276, [http://dx.doi.org/10.1051/shsconf/20140801284]

MARCHELLO-NIZIA, Christiane (2012) « L'oral représenté : un accès construit à une face cachée des langues 'mortes' », in C. GUILLOT, B. COMBETTES, A. LAVRENTIEV, É. OPPERMANN-MARSAUX et S. PRÉVOST (édit.), *Le changement en français. Etudes de linguistique diachronique,* Bern/Berlin/Bruxelles, Peter Lang, p. 247-264.

MARCHELLO-NIZIA, Christiane, « Les débuts de l'« oral représenté » en français : marquage du discours direct dans les plus anciens textes », *Mélanges Soutet*, à paraître.