

Organisation prosodique et style de parole en français parlé

*Michel WAUTHION*¹

Università degli Studi di Pavia, Dipartimento di Studi umanistici. 27100 PAVIA, Italie

Résumé. L'attention portée aux propriétés prosodiques du langage courant dans le discours porte les locuteurs à développer un savoir épilinguistique sur la signification et l'interprétation de l'intonation, de l'accentuation et du rythme dans le discours. Ce savoir intuitif accompagne une perception holistique de la prosodie dont il convient de conserver la mémoire comme élément de signification au moment de mesurer par le détail des analyses de corpus le fonctionnement de la prosodie. Les contraintes contextuelles et stylistiques, l'étude des genres du discours oral, l'examen des correspondances relatives entre la structure syntaxique et l'organisation prosodique replacent dans une perspective d'ensemble les éléments matériels qui identifient la voix d'un locuteur. Différents exemples sont proposés dans le cadre de la simulation théâtrale et de la voix singulière de Roland Barthes.

Abstract. Phonostylistics and the prosodic chart in current French. Attention to the prosodic properties operating through everyday language discourse leads ordinary speakers to develop an epilinguistic knowledge of the meaning and interpretation of suprasegmental intonation, accentuation and rhythm into discourse. This intuitive knowledge accompanies a holistic perception of prosody, the remembering of which should be retained as a relevant element of meaning when measuring the functioning of prosody through detailed corpus analyses. Contextual and stylistic constraints, the study of the genres of oral discourse, the examination of relative correspondences between syntactic structure and prosodic organization place in an overall perspective the material elements that identify a speaker's voice. This paper introduces examples taken from a corpus made both of actors on the record and also from the singular voice of Roland Barthes.

1. Introduction : interpréter, désinterpréter et réinterpréter

La prosodie est un élément de structuration du discours qu'on qualifie de suprasegmental, pour indiquer qu'il s'agit d'un degré de perception qui transcende le plan de l'organisation grammaticale ou de la construction lexicale. Chacun peut en faire l'expérience à différents

degrés, par exemple à l'écoute des langues étrangères. L'interprétation spontanée d'une parole dont on ignore le code s'appuie sur les éléments de reconnaissance non verbaux. C'est un principe mis en avant dans l'intercompréhension entre les langues de même famille. La perception de la courbe intonative, l'identification de la *clausure* post-verbale permettent de hasarder le découpage d'un discours en unités intonatives dont on ignore le contenu lexical et, partant, la signification. La situation de la parole dans un cadre discursif lui-même codifié détermine également une qualification de la parole entendue dans un type d'énoncé. Ainsi, il semble relativement aisé de ranger par la seule écoute un discours radiophonique dans un ou plusieurs genres correspondants suggérés par le cadre énonciatif : un bulletin d'informations, un entretien, une dramatisation, un discours politique. Cette expérience d'interprétation spontanée de la voix située est une expérience familière, souvent complétée par le repérage du lexique international qui permet de donner des éléments de contenu à une approche sémasiologique de la forme du discours, par exemple dans la didactique de l'intercompréhension entre les langues².

De manière générale, l'interprétation de l'élément suprasegmental perçu est une activité permanente du sujet parlant. Elle opère spontanément en réaction à l'émission des messages verbaux et non verbaux les plus divers, que ceux-ci lui soient ou non directement adressés et quelle qu'en soit l'intention communicative. Le mécanisme interprétatif est de type assimilateur, c'est-à-dire qu'il rapporte l'élément perçu à une expérience antérieure semblable, et négligera le plus souvent comme inintelligibles les éléments nouveaux qui ne peuvent être rapportés à une expérience quelconque. Notre perception de la voix est tout à la fois de nature synthétique et analytique, sans que nous soyons nécessairement en mesure de séparer l'une de l'autre. L'élément synthétique renvoie à l'*ethos* de la voix, fruit d'une expérience globale qui comprend les indicateurs de la personnalité, de l'identité vocale du locuteur : il y a certes là des éléments acoustiques analysables mais, comme l'indique Auchlin (2008), la voix possède en soi une dynamique d'appréhension globale qui reste, par essence, inanalysable. Telle nous fait impression de manière globale une couleur étendue sur la toile, dont on peut aussi étudier par ailleurs la pigmentation. Le composant analytique renvoie précisément aux paramètres décomposables de la prosodie : rythme, accentuation, mouvement mélodique, durée, intensité. Aussi peut-on dire que, dans une certaine mesure, l'analyse prosodique du message verbal nécessite de *désinterpréter* d'abord, pour *réinterpréter* ensuite. Désinterpréter la voix humaine signifie se démarquer d'une impression globale (une voix chaude, épicée, sèche, riche) relative à l'émission sonore sans la renier. L'interlocuteur aura soin de ne pas interpréter une voix en tant qu'elle représente le marqueur identitaire de celui qui parle. En revanche, les paramètres intonatifs et rythmiques du discours constituent l'interprétable du locuteur, au même titre que l'information grammaticale ou lexicale du discours. Il existe donc un paradoxe de l'*ethos* du locuteur, souligné par Auchlin (2008), selon lequel une perception globale et non médiante de la voix accompagne directement une expérience de discours mais ne se soumet pas à une analyse sémiotique. Le « grain de la voix » existe, sans aucun doute ; il marque l'auditeur mais n'appartient pas à un registre construit et décomposable.

Il faut encore préciser que la nature non intentionnelle de l'*ethos* de l'émission verbale n'empêche pas son analyse. Notre objectif consiste précisément à identifier ce qui constitue l'identité vocale de l'énonciateur, à analyser les paramètres de la voix afin de remplacer des descriptions perceptives par des éléments d'expression identifiables. Car si la voix est perçue globalement dans l'événement énonciatif, rien n'empêche d'isoler le mécanisme menant à telle ou telle perception. En toutes choses, il convient de mettre en relation des données objectives (comme les variations sonores de l'émission vocale dans l'interaction) avec des éléments subjectifs de perception (les effets de ces émissions ressentis par les interlocuteurs en présence, associés à un degré de convergence).

Disons enfin que la prosodie possède aussi une fonction interprétative très puissante, liée cette fois à l'analyse du discours : l'intonation et le rythme structurent l'expression et contribuent à la construction du sens dans la parole, au même titre que l'organisation syntaxique et pragmatique. L'analyse sémantique ne reconnaît pas toujours la juste part prise par l'information prosodique dans la construction du sens.

En définitive, notre proposition consiste à présenter deux niveaux de signification de la prosodie dans l'expression de la voix et à les mettre en relation en fonction de contextes discursifs également décrits : d'une part, l'*ethos* du locuteur ou la présentation de soi, qui relèvent de l'idiosyncrasie ; d'autre part, l'intonation et le rythme comme parties du sens de l'expression, qui sont des éléments objectifs assignables. Nous pouvons d'ores et déjà en structurer le projet dans quatre dimensions qui se recourent partiellement :

1.1. Syntaxe et prosodie en dialogue

Il est certes possible de rapporter dans l'écriture des marques expressives, de favoriser la dislocation et de se démarquer de l'archétype de la phrase verbale. Mais la mise en concurrence entre les marques expressives et la syntaxe est bien plus vive dans la langue parlée, où il s'avère possible de classer et comparer les genres du discours selon qu'ils soient segmentés plutôt en unités groupées par la syntaxe, par la prosodie ou par les deux (Degand & Simon, 2011). Car si l'intonation est un système qui sert à « linéariser la hiérarchie des structures syntaxiques, (...) [elle] n'a pas vocation à être seulement un indicateur de la structure syntaxique » (Rossi, 1999 : 7-8).

1.2. Phonogenres et phonostylistique

Les genres de l'oral ne sont pas fixés comme les genres littéraires et ils se prêtent à une division en espèces et sous-espèces adaptables selon les besoins de l'enquête. Le modèle *conceptionnel* de Koch et Oesterreicher (2001) permet de « caractériser le comportement communicatif des interlocuteurs par rapport aux déterminants situationnels et contextuels ». Il en résulte pour la langue parlée un tableau des situations où entrent en jeu les principaux traits qui déterminent les styles expressifs.

1.3. Stylistique et sémantique du discours

Autant que les circonstances, le contenu des propos détermine une adaptation peu connue des intonations à la teneur de ce qui se dit. Cela se traduit par des registres de langue accordés aux dictionnaires, à ce que Gumperz appelle les *contextualization cues* (1992b : 231), indices de contextualisation disséminés dans le discours mais aussi à l'emprise de la dimension esthétique. Les circonstances commandent un vocabulaire approprié à la situation de communication, mais les mots ou les thèmes traités possèdent aussi en eux-mêmes la faculté de dicter un discours de circonstance, comme l'observait l'ancienne rhétorique en soulignant l'importance de la relation entre *actio* et *elocutio*.

1.4. Psychologie et sciences cognitives

Dans la communication émotive, les émotions ressenties affectent ou sont susceptibles d'affecter le discours et en particulier l'expression de la voix. Comment la variable émotive est-elle perçue et décodée par l'interlocuteur, d'une façon plus intime que les indices de contextualisation de Gumperz ? L'analyse fera la part des choses entre l'idiosyncrasie de l'expression et les éléments neutralisables et assignables à un dérèglement passager de la maîtrise verbale. En revanche, les pathologies psychiatriques comme l'autisme dérèglent

l'expression prosodique par une mauvaise transmission des sensibilités émotives au comportement observé.

De ces quatre grands facteurs nous ne ferons dans la présente étude qu'introduire brièvement les trois premiers. Nous désirons montrer à travers des exemples précis comment la lecture combinée d'éléments de parole à l'aide de ces approches conduit à une approche de la prosodie qui conserve le bénéfice d'une analyse décomposée sans laisser pour compte une approche holistique qui correspondrait davantage à la perception spontanée de l'*ethos*.

2. Prosodie : quelques généralités

La prosodie désigne en linguistique le domaine de la phonétique qui étudie l'accentuation, le rythme, l'intonation et la durée de la réalisation des phonèmes. *L'accentuation* vise la manière dont la parole détache certaines syllabes dans l'énoncé. Les syllabes accentuées sont dites proéminentes. A strictement parler, *l'intonation* désigne la hauteur musicale (tonale) de la voix dans l'enchaînement des syllabes prononcées dans le discours. L'intonation peut être ascendante, descendante ou stable. Les variations de hauteur des syllabes enchaînées établissent ensemble un contour mélodique que l'analyse prosodique segmente en unités intonatives. Pour ce faire, on examine les variations de hauteur et de rythme de la courbe intonative. Le rythme, qu'on peut définir de façon simple comme la mesure temporelle qui sépare deux syllabes accentuées à l'intérieur d'une suite constituée, construit donc des unités prosodiques inscrites dans la durée, là où l'intonation leur confère une unité mélodique.

La frontière entre unités intonatives est marquée par une syllabe proéminente, c'est-à-dire une syllabe accentuée dans la suite sonore et perçue comme telle pour des raisons de hauteur, d'intensité, de durée et souvent pour l'ensemble de ces facteurs simultanément. Les syllabes proéminentes sont donc les balises qui marquent les frontières entre unités intonatives dans le discours. Elles ponctuent l'unité mélodique. Un groupe accentuel désigne un ensemble de syllabes ponctuées par une syllabe proéminente. Le degré de proéminence de la syllabe sert à mesurer l'importance et l'ampleur du groupe intonatif que celle-ci ponctue.

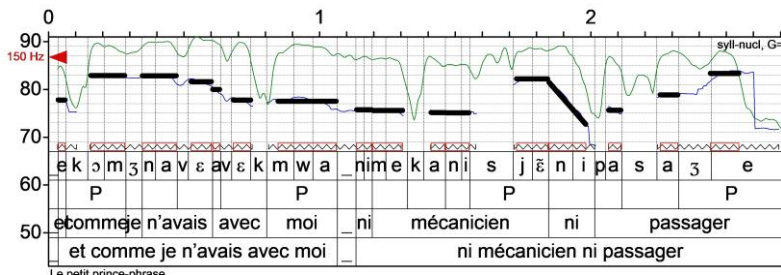
La fragmentation prosodique du discours est une question particulièrement complexe car il s'agit à la fois de déterminer les frontières entre unités intonatives sur des critères prosodiques et non grammaticaux, et de découper ces unités en groupes accentuels selon des critères prosodiques et non lexicaux. Pour ce faire, on convient d'une complétude ou saturation prosodique d'une part, et d'unités prosodiques minimales d'autre part. La complétude prosodique provient de l'examen des courbes intonatives et de la qualité des proéminences accentuelles détectées. Les éléments prosodiques minimaux ou mots prosodiques sont nécessairement associés à un mot lexical plein, possédant une syllabe accentuable. Mais l'accent lexical du français, s'il définit des voyelles accentuables, n'empêche que le groupe accentuel peut excéder le noyau lexical et inclure une quantité allant jusqu'à sept à onze syllabes, selon la vitesse d'élocution du locuteur (Martin, 2009 : 98). Une unité intonative majeure compte un ou plusieurs groupes accentuels et se ponctue par une syllabe proéminente remarquable³.

La contrainte prosodique d'éviter la collision d'accents (suite de deux syllabes accentuées sans pause) oblige le locuteur à introduire des pauses ou à déplacer l'accent prosodique sur la syllabe précédente. Un principe d'eurythmie veut qu'il soit possible de

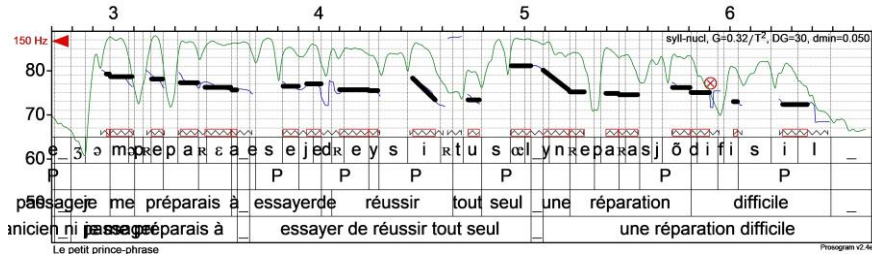
rééquilibrer le nombre de syllabes de chaque groupe prosodique à l'intérieur d'une unité de rang prosodique majeur. (Martin, 2009 : 100-102)

2.1. Détection de frontières et anomalies mélodiques déclamatoires^{4 5}

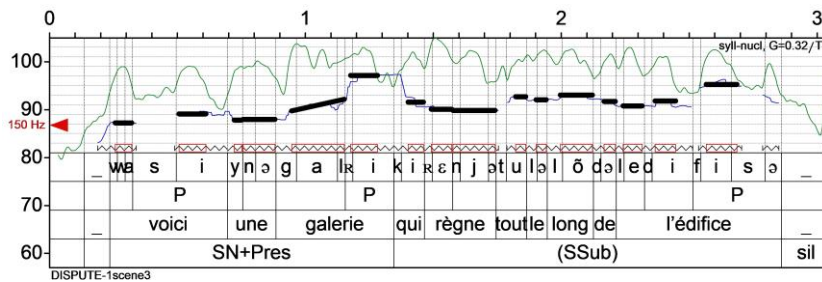
1. Et comme je n'avais avec moi ni mécanicien ni passager, //



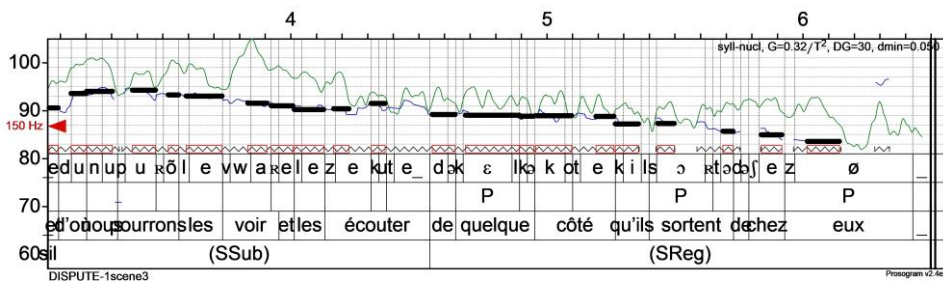
2. je me préparais à essayer de réussir tout seul une réparation difficile



3. Voici une galerie qui règne tout le long de l'édifice //



4. et d'où nous pourrons les voir et les écouter de quelque côté qu'ils sortent de chez eux



Comparons brièvement deux échantillons déclamatoires enregistrés. Extraits d'une tirade et d'un monologue, 1-2 est lu par un comédien chevronné et 3-4 déclamé par un jeune acteur. L'un comme l'autre possèdent les capacités techniques pour produire en un temps court et de façon articulée des ensembles syntaxiques complexes d'environ quarante syllabes, qui se segmentent en deux parties inégales dont la seconde est la plus longue et à laquelle nous allons nous intéresser. Située dans un registre tonal beaucoup plus bas – inférieur à 110Hz, la voix de Gérard Philipe possède dans l'échantillon de phrase analysée une amplitude mélodique de 11,7 demi-tons, légèrement moindre que le second acteur (13,6 demi-tons). Les deux productions possèdent une durée semblable, la déclamation de Philipe possède un débit et une phonation moindre mais l'écart ne paraît pas significatif (respectivement 5,89 et 6,07 syll/sec pour un taux articulatoire de 92,7 et 95,3%).

Qu'est-ce qui distingue alors les deux productions ? Le *Prosogramme* (Mertens, 2004) met en évidence le traitement distinct de l'agitation mélodique et de la construction rythmique. Dans la déclamation de la phrase de Marivaux (3-4), l'acteur semble faire de la seconde unité intonative (4) un seul groupe de souffle de vingt-trois syllabes. En revanche et malgré la contenance tonale dont il fait preuve, Gérard Philipe réalise pour lire Saint-Exupéry trois *glissando* dans le second segment (2), ce qui permet d'organiser un découpage rythmique des groupes que le débit rapide empêche de percevoir directement. En outre, la construction rythmique s'appuie sur une série de sifflantes sourdes pour détacher des syllabes détectées comme proéminentes par le script *Prosoprom*⁶. Nous trouvons dans le jeu de Philipe une architecture tonale structurante qu'on cherchera en vain dans le phrasé de Marivaux, qu'on peut décrire comme un *diminuendo* plat épuisant la capacité interprétative de l'auditeur. On en conclut que le second interprète a reproduit sur le plan intonatif une unité syntaxique trop complexe, qu'il fallait découper en unités plus petites. Martin ne dit rien d'autre lorsqu'il considère les effets maîtrisés de l'obligation du contraste des pentes dans les phrases courtes ou longues : « Ce mécanisme permet de [...] comprendre pourquoi les phrases longues présentent des contours prosodiques manifestant entre eux des contrastes de plusieurs natures, durée, variation mélodique, etc., alors que les phrases courtes voient la plupart du temps leurs marques prosodiques neutralisées et donc révélant peu ou pas de différences prosodiques avec les autres syllabes. » (Martin, 2009 :109)

3. Structuration prosodique et syntaxique du discours parlé

Le travail d'analyse et de structuration prosodique se complète par la comparaison avec l'organisation grammaticale du discours oral. Sur la grammaire du français parlé, les travaux de Blanche-Benveniste (1990) ont permis de travailler sur la macro-syntaxe et de vérifier combien la notion moderne de phrase est peu adaptée pour décrire les unités de rection maximales à l'œuvre dans l'exercice de la parole.

Plusieurs études mentionnent l'écart entre la syntaxe et la prosodie du français, par exemple Martin (2011), qui met en évidence un mécanisme de stockage-concaténation et de *contraste des pentes*⁷ qui préside à la perception de complétude des unités prosodiques hiérarchisées. En proposant « une analyse hypothético-déductive » des mécanismes prosodiques, Martin décrit les contours prosodiques selon des traits pertinents qu'il réduit à trois variables : montant, ample, convexe. Le contraste des pentes voudrait qu'un groupe prosodique de mouvement donné s'insère dans un groupe prosodique supérieur d'orientation inverse. De manière générale, cette approche met en évidence l'autonomie des groupements prosodiques par rapport aux contraintes d'organisation syntaxique du discours. Selon Martin en effet, en dehors de l'évitement des *collisions syntaxiques*, c'est-à-dire « l'impossibilité de regrouper des groupes accentuels qui seraient dominés par des nœuds distincts dans la structure syntaxique » (Martin, 2011 : 46), les structures syntaxiques et prosodiques évoluent de manière relativement autonome. Fort de ce constat, il devient raisonnable de décrire les typologies phonostylistiques selon la contrainte

exercée par la structuration prosodique ou l'organisation syntaxique. La comparaison entre structuration prosodique et organisation syntaxique mène en effet Degand et Simon (2011), par une série d'études successives, à faire émerger une structuration supérieure du discours oral, qu'elles appellent « unités de base du discours » ou *Basic Discourse Unit* (BDU). L'analyse d'un corpus multigenres annoté était l'hypothèse selon laquelle les discours de l'oral, classés le long de l'axe proximité/distance selon les critères établis par Koch et Oesterreicher (2001), possèdent une distribution récurrente de BDU. En analysant successivement les unités (syntaxiques) de rection et les unités (prosodiques) intonatives majeures d'un même corpus, les auteurs proposent de classer les BDU en unités congruentes, groupées par la syntaxe, groupées par la prosodie ou régulatrices. Des éléments de calcul quantitatif permettent ensuite de décrire les styles de parole selon la segmentation privilégiée dans le discours. On observe en effet, selon les auteurs, que les discours dans lesquels prévalent les BDU congruentes renvoient à une présentation directe et neutre de l'information ; que les BDU groupées par la syntaxe produisent un style emphatique, didactique et préparé, que les BDU groupées par l'intonation produisent une macro-unité informationnelle et que les BDU de régulation ont un effet métadiscursif manifeste. Nos auteurs constatent sans surprise le regroupement par la syntaxe des discours académiques, des discours politiques et des homélies. À l'inverse, les genres caractérisés par un plus grand degré de spontanéité, tels que la narration conversationnelle et l'interview libre, se distinguent par une proportion élevée de BDU groupées par l'intonation. Enfin, les entretiens libres fournissent également un bon nombre de BDU congruentes (de neutralité) et les discours académiques sont riches en BDU régulatrices, qui sont favorables au propos d'explication.

4. Phonogenres et phonostylistique

L'étude phonostylistique de la prosodie vise à proposer une typologie des discours comme autant de genres de l'oral et à les caractériser du point de vue expressif. Le point de départ de cette classification naît d'une réflexion portée par Koch & Oesterreicher (2001) sur l'ancrage des discours sur un axe double dont l'un varie entre proximité et distance communicatives, et l'autre entre oralité et scripturalité. Les auteurs élaborent ainsi un modèle « conceptionnel » à l'intérieur duquel il est possible de « caractériser le comportement communicatif des interlocuteurs par rapport aux déterminants situationnels et contextuels ». Il en résulte pour la langue parlée un tableau des situations où entrent en jeu quatre principaux traits qui influencent les styles expressifs : caractère médiatique, degré de préparation, type d'audience et degré d'interactivité. On y ajoutera aussi la prise en compte de l'orientation thématique, ainsi que le fait Labov (1972), qui considère que l'interaction sociale est une idiosyncrasie née du croisement de trois facteurs : la relation entre les interlocuteurs, le contexte social et la thématique de la communication. Une activité orale de discours sera organisée en types dès lors qu'à travers une étude de corpus possédant des traits communs, des invariants pourront être observés. Ces observables existent bien entendu sur plusieurs niveaux, en partie explorés : linguistique (contraintes grammaticales et lexicales), rhétorique et expressif, pragmatique. L'analyse prosodique cherche à établir les paramètres de la variation stylistique en utilisant les éléments factuels dont elle dispose : la mesure temporelle et rythmique (débit de parole et taux d'articulation), l'amplitude mélodique et tonale, l'accentuation (degré de prééminences) et les pauses.

Pour Pierre Léon, les phonostyles sont des styles sonores « tels qu'ils sont perçus en tant que caractéristiques d'un individu (jeune, vieux, homme, femme) d'un groupe social (bourgeois, prolétaire) ou d'une circonstance particulière (discours politique, sermon) » (1993 :3). Une activité orale de discours sera organisée en types dès lors qu'à travers une étude de corpus possédant des traits communs, des invariants pourront être observés. Ces observables existent bien entendu sur plusieurs niveaux, en partie explorés : linguistique (contraintes grammaticales

et lexicales), rhétorique et expressif, pragmatique. La difficulté d'identification des genres de l'oral comme objets didactiques consiste à en isoler les traits prosodiques et rhétoriques spécifiques tout en s'assurant qu'ils ne constituent pas des idiosyncrasies et qu'ils soient reproductibles dans une progression qui, à partir de traits stylistiques saillants, permettent de mieux percevoir et reproduire l'intonation du français.

5. Prosodie et structure syntaxique chez Barthes

« (1) *Il y a là une, euh, une espèce de//de jouissance* (2) – *je répète le mot- //* (3) *dont j'ai effectivement de plus en plus envie //* (4) *de faire sinon vertu tout au moins de faire théorie-* (5) *c'est peut-être la même chose⁸ ».*

L'écoute de ce fragment extrait d'entretien radiodiffusé permet d'observer, chez Barthes, une production orale grammaticalement complexe et structurée comme un écrit oralisé, à peine perturbé par la reprise de deux morphèmes grammaticaux. La phrase complexe est augmentée par deux incises en corps (2) et en fin (5) de texte. Du point de vue prosodique en revanche, on perçoit très bien la présence de 3 pauses (//) qui s'expliquent à la fois par des contraintes syntaxiques (agencement de syntagmes) et psychologiques (recherche expressive). La première, d'abord articulée puis silencieuse (1,83sec), est en effet précédée d'un allongement de la voyelle médiane marquant l'hésitation et suivie par une reprise de la préposition. La seconde (0,6sec) est strictement syntaxique et la troisième (articulée de 0,5sec) est plutôt une recherche expressive placée sur un chainage des unités syntagmatiques. On se trouve de la sorte dans un cas typique d'écrit oralisé, ou de parole obéissant à une organisation grammaticale assez stricte. La phrase est complexe et comprend deux incises ou incidentes, une subordonnée relative à l'intérieur de laquelle l'auteur reformule et nuance son propos.

L'utilitaire ProsoBox (Goldman & Simon, 2020), développé à partir de Praat (P. Boersma & D. Weenink, 2018) permet de disposer d'une détection semi-automatique des proéminences, qui servent nommément à identifier à tout le moins les groupes intonatifs. Ainsi, dans cet extrait l'algorithme identifie quatorze syllabes saillantes⁹, dont la discussion serait la suivante :

La première unité intonative est une courbe de continuité correspondant à une assertion : « *Il y a là une* (P1) */une espèce* (P2) *de* (P3) *jouissance*(P4) ». Des quatre syllabes proéminentes détectées, P2 et P4 correspondent à une finale de groupe intonatif ; les deux autres sont détectées par l'allongement ou la longue pause expressive qui les suivent. Il s'agit bien d'une unité de continuation (mineure et majeure) car la voyelle accentuée en P1 et P2 est haute, ce qui signifie que la proposition est incomplète.

La deuxième unité intonative est une parenthèse de cinq syllabes dont deux sont accentuées ; le premier accent est secondaire et rythmique, le second reporte par glissando ascendant la hauteur tonale à celle qui précède la parenthèse.

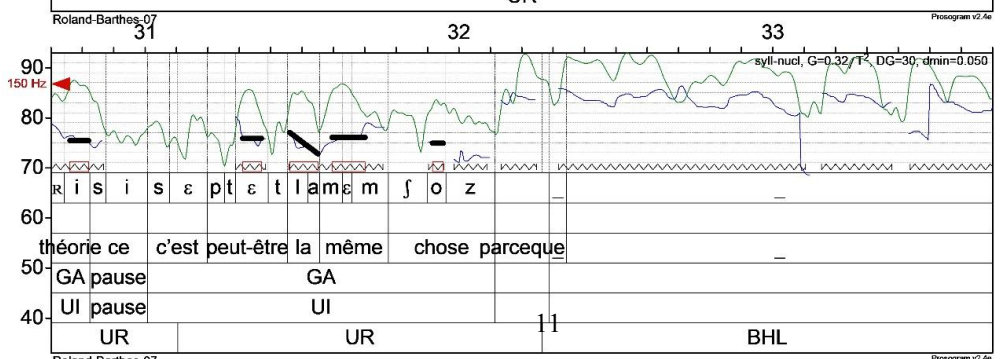
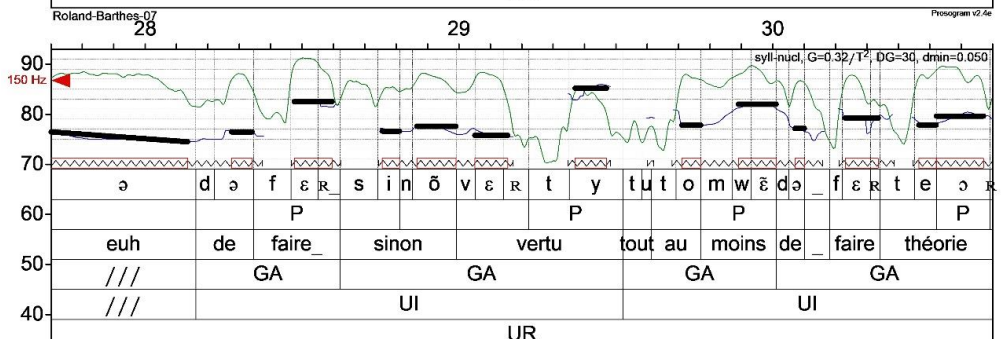
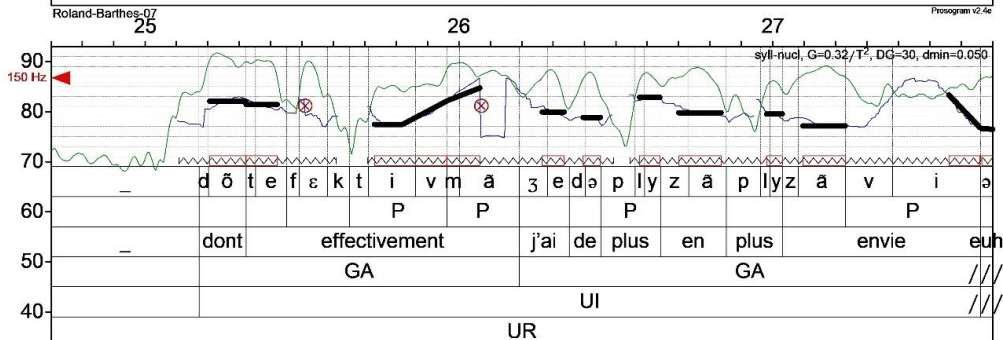
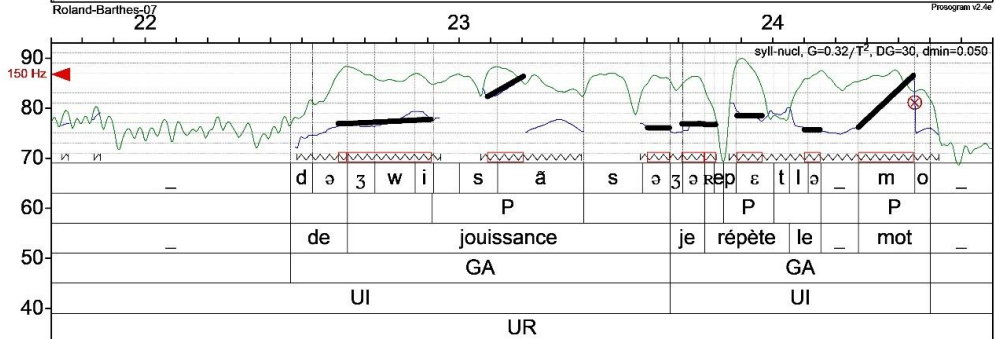
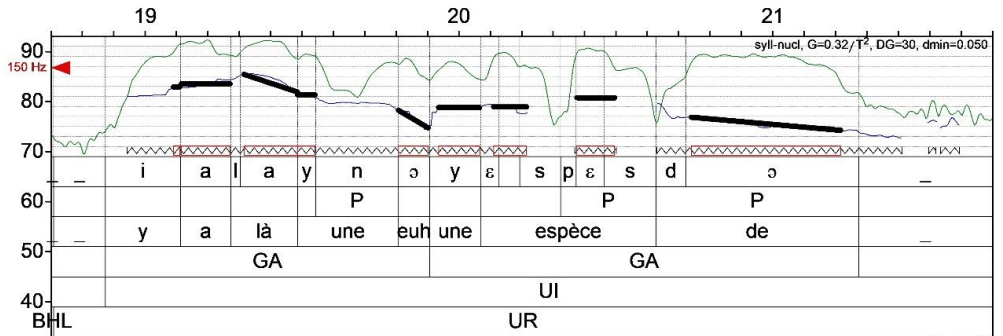
La troisième unité intonative comprend à nouveau quatre proéminences détectées pour deux groupes intonatifs, le premier et le troisième représentant des accentuations secondaires davantage rythmiques. Le niveau de l'accent final de l'unité décrit une longue et progressive baisse de hauteur correspondant à la recherche expressive de l'énonciateur. Cet allongement ne doit pas s'interpréter comme une finalité mais un suspens.

« *De faire* (P1) *sinon vertu* (P2) *tout au moins* (P3) *de faire théorie* (P4) » La dernière unité intonative se termine par un niveau BB- de finalité (Mertens, 2008). On pourrait la découper en deux sous-unités dont la première est ascendante, conformément à la courbe de continuité, et la seconde descendante. Chacune de ces sous-unités commence par le groupe bisyllabique « de faire » dont la dernière syllabe est accentuée à l'initiale (P1) mais pas à la reprise, qui est

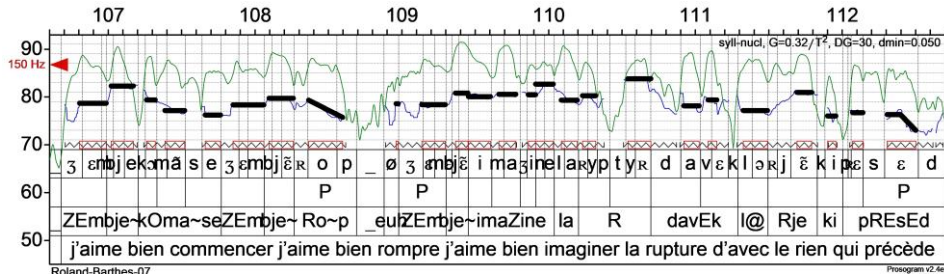
précédée d'une locution adverbiale modérant le propos. Comme l'accent est posé sur l'adverbe, il ne peut se reporter à l'identique sur le groupe suivant sous peine de collision. Le contraste des pentes opère entre le niveau haut du premier groupe accentuel (P2) et le niveau infra-bas du second (P4). Une seconde incise est privée de toute proéminence car elle opère entièrement en deçà du niveau accentuel précédent, comme un écho.

En résumé, l'analyse segmentale de ce premier fragment fait apparaître une grande richesse expressive. Les pauses marquées par le locuteur lui permettent à l'évidence de ramasser sa pensée mais on observe que leur point d'ancrage correspond à des césures syntagmatiques, ce qui permet de confirmer l'impression globale d'écrit oralisé.

Un peu plus loin dans la même partie de cet entretien, on peut observer que la pensée est cette fois-ci organisée par la structure rythmique de l'expression verbale.



« J'aime bien commencer j'aime bien rompre j'aime bien imaginer la rupture d'avec le rien qui précède. »



Il y a bien deux unités intonatives distinctes mais on observe que le glissando décroissant de la première unité reste de registre supérieur à celui de la seconde unité, dont la courbe mélodique décroissante est aussi progressive pour accompagner le groupe beaucoup plus important de syllabes : 8 syllabes dans l'UI1 et 16 dans la seconde (le double). Il en résulte un dégradé perceptif que renforce le lissage de la pente intonative :

« (G11) *J'aime bien commencer* (G12) *j'aime bien rompre* »

Le premier groupe intonatif est borné par une courbe de continuation, le second établit une finalité. La reprise de l'énoncé déclaratif construit un effet rhétorique remarquable. Le produit est ici un contraste expressif et un diminuendo. L'outil ProsoProm a détecté une seule syllabe prééminente, sans doute à cause du faible intervalle entre les deux. L'effet de répétition (la reprise adverbiale marque le sommet intonatif des deux propositions sans qu'une insistance soit directement perceptible à l'oral) confère une unité de rythme à l'ensemble prosodique.

« *J'aime (P1) bien imaginer la rupture avec le lien qui précède (P2)* »

La suite identifie un accent initial et un accent final. Entre les deux, le *prosogramme* détache les trois accents finaux de mot : *imaginer*, *rupture* et *liens*, sans les marquer comme prééminents. Par cet exemple on voit que la structure intonative est effectivement construite et décomposable en unités diversement hiérarchisées : le rythme expressif est chez Barthes un élément de structuration essentiel dans l'intelligibilité du propos, autant que l'accentuation. C'est ainsi que, dans une partie successive du même entretien, on trouve encore :

« *J'aimerais beaucoup des fragments qui auraient un certain calibre, une certaine dimension, une certaine régularité.* »

La reprise du groupe indéfini « *un certain* » souligne par une scansion les propriétés du fragment recherchées par Barthes. Opère ici une unité rythmique manifeste, qui explique peut-être pourquoi l'outil ProsoProm ne détecte qu'une prééminence sur la syllabe finale de *dimension*. En tout état de cause, on note des ruptures tonales avec glissando sur les finales de *fragments*, *dimension*, *calibre*, *régularité*. La dernière lexie, *régularité* possède un glissando plus marqué puisqu'elle agit en clôture de l'unité intonative finale. La dernière syllabe est aussi porteuse d'une intonation infra-basse correspondant à la clôture intonative, formant contraste avec l'intonation montante des deux lexies précédentes.

6. Stylistique et sémantique du discours

La voix de Barthes a laissé un souvenir chez ceux qui l'ont connu ou qui ont été les auditeurs des cours qu'il donnait au Collège de France ou à l'École pratique des Hautes Études. Julia Kristeva :

« Ce timbre d'une fragilité ferme confère à sa communication immédiate, malgré la discrétion de la conversation, malgré la distance, la force d'un contact physique. L'homme qui s'adresse à vous livre une parole au-delà du sens. (...) [U]ne courbe chantée ironique, comme pour marquer l'inanité du propos commun et de sa propre demande; (...) Tout cela qui résonne toujours, encore, au présent, est inscrit dans l'étoffe du son et dans les inflexions de la mélodie, qui vous atteignent avant la signification et au-delà d'elle. Etablissant une complicité sonore, intemporelle, inconsciente, cette voix durable devient ainsi le support incontournable sur lequel s'inscrit un enseignement fluctuant, mobile, radicalement a-didactique. » (Kristeva, 1982 : 119-120 – nous soulignons)

Dans la notice d'accompagnement des enregistrements, Renaud Machart décrit plus concrètement la voix de Barthes dans les termes suivants (nos italiques) : « une voix faite d'une *fondamentale grave, sombre, mate dont la richesse génère des harmoniques naturelles*. Des harmoniques en l'occurrence françaises, avec cette manière qu'avaient les gens de sa génération et de sa culture de parler un français aujourd'hui disparu, avec des *nasales longues*, « épiciées ». (...) Barthes parlait un *français patiné*, en rien *hystérique* et *surligné*. »

On ne retiendra de ce florilège que les considérations prosodiques. Le « français patiné, en rien hystérique et surligné », la « fondamentale grave » sont effectivement confirmés par l'analyse sonore, où les écarts mélodiques sont très contenus. Dans la séquence d'enregistrement où ont été pris les exemples qui précèdent (188 secondes), l'étendue de la voix est de 12,4 demitons et les tonalités oscillent entre 62 et 148 Hz, c'est –à-dire qu'elles restent en dessous du seuil moyen de la voix masculine, conventionnellement fixé à 150 Hz. Ceci donne une impression de voix grave et lissée, car l'agitation mélodique y est également contenue. Elle n'est donc pas hystérique¹⁰ dans la mesure où elle reste contenue dans cette amplitude et que le taux de remplissage sonore de la parole est élevé : 78,21% du temps de parole est en effet occupé par l'expression, dont on peut ensuite mesurer la durée des noyaux syllabiques¹¹. Le débit moyen de la parole est de 5,25 syll/sec, ce qui est relativement lent mais surtout posé puisqu'il comporte peu de silences.

On constate à l'aide du *prosogramme* que le taux de glissando est important chez Barthes, de 10 à 13% selon les indices de segmentation choisis (descentes mélodiques à 70% et montées à 30%). A cela s'ajoute que les intervalles de variation tonale sont aussi importants à l'intérieur des noyaux syllabiques qu'entre les syllabes. Cela explique concrètement l'impression de patine ressentie à l'audition : les voyelles sont longues et accompagnent la variation mélodique par ces effets de descente en finale mais aussi de montée, comme nous l'avons esquissé. En d'autres termes, la voix de Barthes reste contenue dans un registre bas à l'intérieur duquel la variation mélodique est progressive et répartie équitablement dans le noyau syllabique et en dehors de celui-ci. On peut contraster un tel profil prosodique avec celui de ces voix qui, usant d'un registre beaucoup plus étendu, privilégient les écarts mélodiques opérés en dehors des noyaux syllabiques : peu ou pas de glissando rendront la perception de l'intonation difficilement chaleureuse. Enfin, l'analyse suivie de la voix vérifie que le schéma intonatif de Barthes fait somme toute peu usage de l'accent initial de mot, qu'on appelle parfois l'accent pédagogique, très présent aussi dans le discours politique ou journalistique. Les quelques exemples relevés concernent précisément les concepts que l'écrivain veut mettre en exergue : insistance dans « absolument » (34-35) ; « commencer » (95-96). Les cas d'accents finaux sont en revanche très

nombreux (« formes brèves » (49-50) ; « japonais » (68-69)) mais on relèvera aussi les renversements sur des groupes accentués plus longs comme « l'attaque d'un fragment, l'attaque d'un paragraphe » (103-106) où l'accent du mot répété marque la cadence, le complément du nom ne bénéficiant que d'un accent secondaire dans le groupe rythmique.

7. Conclusions

« Le « grain », ce serait cela : la matérialité du corps parlant sa langue maternelle (...) Le « grain », c'est le corps dans la voix qui chante, dans la main qui écrit, dans le membre qui exécute. »

L'expression forgée par Barthes a fait long feu pour décrire la corporéité sonore. Le « grain de la voix » renvoyait pour lui à un type d'interprétation du lied chanté dont la qualité tient à une diction corporelle plutôt qu'à un phrasé articulé dominé par la respiration. Barthes met en avant dans le grain de la voix la présence du corps dans le chant, dont le meilleur indice est selon lui la priorité que l'interprète donnera aux voyelles sur les consonnes, dont l'auteur recommande d'en polir les aspérités. Pour mieux se faire comprendre, il oppose une génération d'artistes à une autre et insiste sur la qualité de celle qui savait placer son interprétation au service des noyaux vocaliques et non l'inverse : « Les consonnes dont on pense trop facilement qu'elles forment l'armature de notre langue (...) et que l'on impose toujours « d'articuler », de détacher; d'emphatiser, pour satisfaire la clarté du sens, Panzéra recommandait au contraire, dans bien des cas, de les patiner, (...) d'en faire le simple tremplin de la voyelle admirable. » (Barthes, 2002 :1439) On ne peut manquer de constater que l'analyse de la propre voix du sémiologue reflète certains des traits qu'il cherche à saisir à travers la métaphore du grain. D'abord, nous avons constaté que la trajectoire moyenne des intervalles dans le fragment analysé est de 17,99 demi-tons, plus proche de l'agitation mélodique des noyaux syllabiques que des intervalles moyens parcourus entre les noyaux. Si l'on compare la voix de Barthes à celle de Céline ou Queneau, on est frappé par l'importance chez ces derniers auteurs des intervalles intervocaliques pour construire le registre tonal des paroles. Cela signifie en clair que la parole de Barthes est davantage centrée sur l'importance du registre des noyaux vocaliques dans la construction de son identité sonore. Et le constat dressé de l'importance des voyelles pour la construction du registre tonal peut être reproduit dans l'analyse des longueurs vocaliques. Les noyaux vocaliques occupent la moitié du temps de phonation chez Barthes et 26% chez un Queneau. On peut raisonnablement conclure à une production sonore où la couleur des voyelles est très présente. Cela laisse à penser que le *grain* renseigné comme l'indice de présence du corps dans la parole ne serait rien d'autre qu'une vibration plus constante des cordes vocales, associée à une part plus grande des voyelles dans l'agitation mélodique.

Cette analyse stylistique adossée à une étude préalable de la relation entre prosodie et syntaxe dans la structuration du discours permet d'établir une première description de l'identité sonore du locuteur en contexte. Une lecture plus complète tiendra compte de l'empreinte émotive et éventuellement pathologique dans l'architecture de la voix.

Références bibliographiques

- Auchlin A. (2000). Ethos et expérience du discours, dans Wauthion, M. et Simon, A.-C. (dir.), *Politesse et idéologie*, Louvain-la-Neuve, Peeters, pages 75-93.
- Auchelin A. (2008). Du phonostyle à l'éthos, les prosodies comme interfaces entre sens et corps. Pour repenser l'acte de langage à partir de l'activité de parole, *Actes du III Symposium sur 'Analyse du Discours*, Belo Horizonte, Brésil, 19 pages.

- Barthes R. (2002) [1972], Le grain de la voix, *Œuvres complètes*, t.2, p. 1436-1442.
- Boersma, P. & D. Weenink (2018). *Praat: doing phonetics by computer [Computer program]*. Version 6.0.37, consultée le 14 mars 2018 sur <http://www.praat.org>
- Goldman J.-Ph., M. Avanzi, A. Lacheret-Dujour, A.-C. Simon & A. Auchlin (2007), A Methodology for the Automatic Detection of Perceived Prominent Syllables in French, *Proceedings of Interspeech07*, Anvers, 27-31 aout.
- Goldman J.-PH, A.-C. Simon (2020). ProsoBox, a Praat Plugin for Analysing Prosody, *Acts of the 10th Speech Prosody Symposium*, Tokyo.
- Gumperz J. J. (1992a). Contextualization revisited In P. Auer & A. di Luzio (dir.), *The contextualization of language*, Amsterdam, John Benjamins, p. 39-53.
- Gumperz J. J., (1992b). Contextualization and understanding, in A. Duranti & Ch. Goodwin (dir.), *Rethinking context*, Cambridge University Press, p. 229-253.
- Kristeva J. (1982). La voix de Barthes. *Communications*, 36, p. 119-123.
- Koch, P. et W. Oesterreicher. (2001). Langage parlé et langage écrit. In Günter Holtus (ed). *Lexikon der romanistischen Linguistik (LRL)*, 1,2, Tübingen: Narr, p. 584-627.
- Léon P. (1993). *Précis de Phonostylistique*, Paris, Nathan.
- Martin Ph. (2009). *L'intonation du français*, Paris, A. Colin.
- Mertens P. (2004). Un outil pour la transcription de la prosodie dans les corpus oraux. *Traitement Automatique des langues* 45 (2), 109-130.
- Mertens P. (2008), Syntaxe, prosodie et structure informationnelle : une approche prédictive pour l'analyse de l'intonation dans le discours. *Travaux de linguistique*, 56, p. 97-124.
- Rossi M. (1999), *L'intonation : le système du français*, Paris, Ophrys.
- Simon A.-C., L. Degand (2011). L'analyse en unités discursives de base : pourquoi et comment?. *Langue française* n°170, p. 45-59.

¹ Courriel: michelflorentgeorges.wauthion@unipv.it

² Meissner F.J., *EuroComRom : Les sept tamis : Lire les langues romanes dès le départ*, Shaker Verlag, 2004.

³ Pour Degand et Simon (2011 : 49) « toute syllabe proéminente en position finale de mot lexical (plein) est potentiellement une frontière d'UIM si elle répond à une des contraintes suivantes : (a) elle est suivie par une pause silencieuse d'une durée minimale de 200 ms ; (b) elle est allongée (trois fois plus longue que la durée moyenne des syllabes qui précèdent ; (c) elle porte un contour mélodique montant particulièrement saillant (entre 5 et 10 demi-tons).

⁴ Fragment (1-2) tiré de l'adaptation par Gérard Philipe de *Le Petit Prince* de Saint-Exupéry, Publié par le label Festival en 1954 et diffusé aux éditions Gallimard.

⁵ Fragment (3-4) extrait de *La Dispute* de Marivaux par le Centre dramatique national de Tours, mise en scène de Jacques Vincey, le 3 juin 2016.

⁶ Le script Prosoprom (Goldman et al., 2007) permet la détection automatique des proéminences à partir de paramétrages incluant l'identification du noyau syllabique, sa durée relative, la hauteur du fondamental et le mouvement mélodique, Voir aussi Goldman et Simon (2020)

⁷ Le contraste des pentes désigne une régularité d'organisation prosodique selon laquelle un mot prosodique terminé par un contraste scalaire montant appartient à un groupe prosodique majeur terminé par un contraste inverse. Cette hypothèse permet à

Martin d'opérer le regroupement des mots prosodiques en relation de dépendance avec un élément de contour de pente inverse situé à droite (Martin 2009 : 86)

⁸ Roland Barthes, *Fragments de voix*, INA, Les grandes heures de Radio-France, 2004. Entretiens de février 1977 avec Jean-Marie Benoist et Bernard-Henri Levy diffusés sur France-Culture.

⁹ « the syllable is considered a primitive unit for prosodic analysis, not only regarding temporal variables (such as speech rate), but also regarding rhythmic and accentuation variables (relying on prominent syllables distribution and features), and even intonation (apprehended as a sequence of tonal targets combining into contours, rather than as a continuous curve) » (Goldman & Simon, 2020)

¹⁰ Léon rappelle que les troubles vocaux renvoient à des cas de surtension ou de sous-tension vocale. « Dans les cas de surtension, la voix est *rauque*, parfois nasalisée, au timbre criard » (Léon, 1993 :270). C'est ce que nous avons observé par endroit dans la voix de Queneau et plus notablement chez Artaud.

¹¹ Les mesures du taux de phonation calculées par le *prosogramme* varient considérablement en fonction des critères de segmentation choisis. Le découpage calculé sur les noyaux syllabiques donne un taux de phonation de 57,5% ; le critère de syllabation automatique donne le pourcentage plus élevé mentionné en corps de texte. Toutefois, les deux critères de segmentation convergent dans la mesure du débit de la voix de Barthes et du registre tonal employé dans cet extrait.