

# ANALYSE PHONOSTYLISTIQUE DU DISCOURS RADIOPHONIQUE

## Expériences sur la mise en fonction professionnelle du phonostyle et sur le lien entre mélodicité et proximité du discours radiophonique

Baptiste Hupin, Anne-Catherine Simon<sup>1</sup>

### La radio, un mode de communication à vive voix

La radio est un média intime. Un phénomène de masse qui pourtant touche individuellement chacun de ses auditeurs. En radio, le présentateur ne dispose que de sa prose et de sa prosodie pour nouer une relation avec chacun de ces auditeurs. Pierre Bourdieu affirme que « *la compétence suffisante pour produire des phrases susceptibles d'être comprises peut-être tout à fait insuffisante pour être écoutée*<sup>2</sup> ». En d'autres termes, il ne suffit pas de parler dans un micro pour être écouté.

---

1 Baptiste Hupin est licencié en information et communication de l'UCL et journaliste à la RTBF. Anne-Catherine Simon est professeur à l'UCL (Département d'Etudes romanes, Institut Langage et Communication). Jean-Philippe Goldman (Université catholique de Louvain, Université de Genève) a développé les logiciels d'analyse sans lesquels cette étude n'aurait pas été possible. Les auteurs l'en remercient.

2 P. Bourdieu, *Ce que parler veut dire, l'économie des échanges linguistiques*, Fayard, Paris, 1982.

*Recherches en communication*, n° 28 (2007).

## Le style vocal, ou phonostyle

Au-delà des mots et de leur sens, c'est au style vocal du présentateur qu'incombe la tâche d'informer et de fidéliser l'auditeur. Cette manière de parler, constitue la dimension non verbale du langage radiophonique. Le fonctionnement du style vocal, encore appelé le phonostyle par Léon<sup>1</sup>, a été largement décrit par Fónagy dans son ouvrage *La vive voix*<sup>2</sup>. Ivan Fónagy se représente la production de la parole, dont le langage radiophonique est un exemple, comme deux actes d'encodage successif du message.

Le premier encodage est linguistique. Il transforme une idée en une séquence de phonèmes et est régi par la grammaire. Le deuxième encodage est para-linguistique. Le codage para-linguistique correspond à l'actualisation, la mise en son des phonèmes. Il greffe le message secondaire, sonore, sur le message primaire. C'est à ce deuxième encodage que nous nous intéresserons particulièrement.

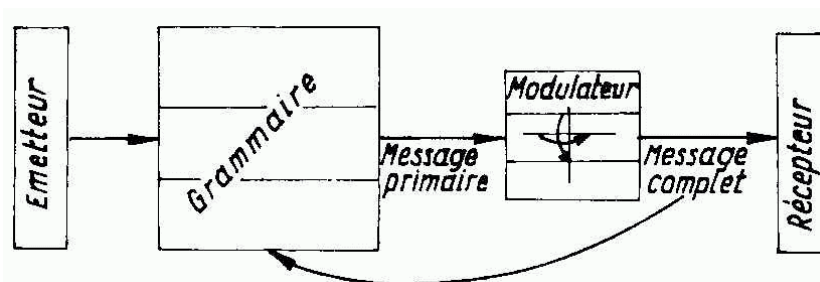


Figure 1. Modèle de l'encodage double. "La phrase engendrée par la Grammaire passe nécessairement par le Modulateur (ou Détracteur) qui greffe des messages secondaires sur la phrase" (Fónagy 1983: 229)

L'encodage para-linguistique est le fait d'un *modulateur* qui déforme le message linguistique. Pour Fónagy, cet encodage fonctionne par métaphore. Ce deuxième encodage est dit motivé et continu. L'allongement du mot «énorme», par exemple, permet de montrer à quel point une chose est énorme. Il y a un rapport intrinsèque de causa-

1 P. Léon, *Précis de phonostylistique, Parole et expressivité*, Nathan Université, Paris, 1993.

2 I. Fónagy, *La vive voix. Parole et expressivité*, Payot, Paris, 1983.

lité entre l'allongement des voyelles ou des consonnes et la taille de l'objet désigné. Le mimétisme occupe une place prépondérante dans la production et la compréhension des signes vocaux. D'après Fónagy, les organes de la voix ont la capacité d'imiter les gestes de notre corps.

Les signes vocaux sont donc des signes gestuels, « *des espèces d'interfaces mimétiques dont l'effet pragmatique premier est d'assurer une connexion identificatoire entre locuteurs et allocutaires* »<sup>1</sup>. En d'autres termes, il faut considérer le vocal comme une communication non verbale qui fonctionne de manière analogique. Or les modes de communication analogiques, comme l'ont affirmé les psychosociologues de Palo-Alto, concernent avant tout la relation.

Le style vocal dispose d'une force pragmatique. « *Le style vocal est d'une importance capitale dans la pragmatique des rapports sociaux. Sous les paroles qui s'échangent, quel que soit le sujet de nos conversations, le style vocal établit, maintient et module continuellement la relation sociale (...)* »<sup>2</sup>. Il confère à chacun des interlocuteurs son rôle dans la situation de communication. Le phonostyle du présentateur radio, sa manière de s'exprimer, modulerait par conséquent la relation qu'il entretient avec l'auditeur. On peut donc supposer que le journaliste parle différemment lorsqu'il est en studio, dans son rôle de journaliste, que lorsqu'il discute de manière informelle avec des collègues. Plusieurs journalistes confirment d'ailleurs cette impression. « *C'est sûr, ne serait-ce que parce que tu fais une lecture et que quand tu lis tu ne vas jamais parler de la même façon* »<sup>3</sup>. Pierre Léon évoque cette possibilité lorsqu'il parle de *mise en fonction professionnelle* du phonostyle.

### Une radio, un style ?

La question se pose donc de savoir si les journalistes changent effectivement de ton, et de quelle manière, lors de leur passage à l'antenne, conformément à la théorie de la mise en fonction profession-

1 J.-P. Meunier, D. Peraya, *Introduction aux théories de la communication, Analyse sémio-pragmatique de la communication médiatique*, De Boeck, Bruxelles, 2004, p. 288.

2 *Idem*.

3 Déclaration d'un journaliste de La Première au cours d'une entrevue informelle (RTB-S-CO).

nelle. Par ailleurs, les chaînes de radio se différencient, entre autres, par leur auditoire. Et la relation qu'elles entretiennent avec leurs auditeurs peuvent être des critères de ralliement à une chaîne. Une chaîne d'information s'adressera par exemple de manière différente à son public qu'une chaîne musicale destinée à des jeunes. Existerait-il dès lors des différences phonostylistiques entre les présentateurs de l'une ou l'autre radio ? Chacun a déjà pu faire l'expérience d'identifier intuitivement une chaîne de radio sans pour autant faire attention à la fréquence choisie ou aux propos tenus. Par ailleurs, s'il existe une différence phonostylistique entre des chaînes radio, il importe d'en déceler la nature et l'origine.

Nous nous sommes posé cette question à l'égard de deux chaînes concurrentes du paysage radiophonique belge : Bel-RTL et La Première, dans le contexte plus particulier de la présentation de journaux parlés. Puisque nous ne disposons pas d'outils de mesure de la relation qu'une chaîne entretient avec ses auditeurs, nous avons présupposé la nature de cette relation en fonction de la manière dont les chaînes se présentent. D'emblée, Bel-RTL se présente comme une radio de proximité alors que La Première se présente comme une radio de contenu dont on peut supposer qu'elle favorise une plus grande distance avec l'auditeur. Il suffit d'ailleurs d'observer les slogans de ces deux chaînes : « Vivre Ensemble » pour Bel-RTL, « Restons Curieux » pour La Première.

De nombreux auteurs se sont déjà penchés sur la question du phonostyle journalistique. Pierre Léon le caractérise comme un « *phonostyle haché* ». <sup>1</sup> Piet Mertens caractérise le style journalistique, entre autres, par les accents d'insistance, le changement de registre mélodique et les liaisons sans enchaînement <sup>2</sup>. Brigitte Besse estime que le style radiophonique relève de *l'écrit-parlé*. <sup>3</sup> Selon Lucie Ménard, qui analyse un corpus de lecture de bulletins de nouvelles télévisés, le style des nouvelles lues est « *un style construit, appris de la part des locuteurs, mais qui représente fidèlement l'image ou la norme qu'une*

---

1 P.Léon, *Précis de phonostylistique, Parole et expressivité*, Nathan Université, Coll. Fac. Linguistique, Paris, 1993, p. 164.

2 P.Mertens, *Phonétique française*, Leuven, Acco, 2005.

3 Brigitte Besse, Docteur en linguistique et animatrice de séminaires de méthode de journalisme audiovisuel à la RTBF, citée lors d'un cours de François Ryckmans donné à l'UCL sur l'écriture radiophonique.

*communauté linguistique projetée d'elle-même*»<sup>1</sup>. Elle met en évidence que ce phonostyle représente une référence dans les pratiques d'une communauté linguistique.

Pour Jean-Pierre Meunier et Daniel Peraya, le phonostyle journalistique se caractérise par un haut degré de mélodicité. Ce style connaît des variations selon le type d'émission, mais son effet serait globalement le même : « *envelopper l'audience dans une même attitude spectatorielle faite de participation à distance* ».<sup>2</sup> Ils décrivent la lecture d'un bulletin de nouvelles comme une courbe au haut degré de mélodicité qui emporte destinataire et destinataires dans un même mouvement.

Conformément à cette description et à cette interprétation du style vocal du journaliste, on peut émettre l'hypothèse que plus la chaîne développe une relation de proximité avec le public, plus la mélodicité du discours est forte. Meunier et Peraya considèrent effectivement que la mélodicité sollicite au maximum la fusion mimétique. « *Adressée à l'auditeur, elle [la mélodicité] appelle cette vibration à l'unisson qui donne un sentiment d'unité.* »<sup>3</sup> La relation qu'une chaîne entretient avec ses auditeurs est, à la lumière de cette théorie, un facteur influençant la mélodicité du phonostyle. Plus la chaîne est proche de son public, plus la mélodicité de l'énoncé radiophonique sera élevée. Et inversement.

### Aspects prosodiques du phonostyle

La prosodie concerne les modulations de la voix qui affectent des mots ou des suites de mots, comme les variations du débit ou la ligne mélodique qui caractérise un énoncé. Ces phénomènes ont une fonction dans la communication. Cette fonction n'est pas de même nature celle des éléments dits segmentaux, comme les phonèmes, grâce auxquels on distingue par exemple la suite "je reviens" de "je retiens", où /v/ s'oppose à /t/. La prosodie est dite suprasegmentale, puisqu'elle court sur des portions de parole supérieures aux simples segments (les

1 L. Ménard, «Le registre en tant que marqueur prosodique de dialectalité», dans M. Demers, *Registre et voix sociale*, Nota Bene, Québec, 2003, p. 39.

2 J.-P. Meunier, D. Peraya, *Introduction aux théories de la communication, Analyse sémio-pragmatique de la communication médiatique*, De Boeck, Bruxelles, 2004, p. 288.

3 *Ibid.*, p. 286.

phonèmes). Elle comprend «*les variations de hauteur, d'intensité et de durée qui constituent le matériau de base de la mélodie, de l'accentuation, du rythme et des pauses*<sup>1</sup>».

### Éléments essentiels de la prosodie: hauteur, durée, intensité

Hauteur, durée et intensité sont les éléments de base de la prosodie. La combinaison et les variations de ces trois paramètres confèrent ses particularités prosodiques à un énoncé en déterminant son rythme, son accentuation et son intonation.

La hauteur d'un son correspond en fait à la fréquence de vibration des cordes vocales. Plus la fréquence de vibration des cordes vocales augmente, plus le son émis sera aigu. Cette fréquence de vibration est appelée la *fréquence fondamentale* (F0). Mesurée en Hertz (nombre de cycles par seconde), la fréquence fondamentale est convertie en demitons (comme en musique) afin de correspondre à une échelle (logarithmique) perceptive<sup>2</sup>. Cette conversion permet la comparaison de mouvements mélodiques réalisés par des voix qui n'ont pas la même hauteur moyenne.

Dans l'énoncé, la **hauteur** permet de distinguer des syllabes de hauteur différente. L'intervalle de hauteur, encore appelé intervalle mélodique, représente la distance mélodique entre deux syllabes. Les hauteurs de plusieurs syllabes successives déterminent la ligne mélodique (ou **contour** mélodique) d'un énoncé.

La hauteur mélodique est déterminée par la **tessiture** d'un locuteur. La tessiture est définie par la hauteur et l'étendue moyennes de la voix d'un locuteur<sup>3</sup>. Généralement, lorsqu'un locuteur parle, il ne se sert pas de l'ensemble de sa tessiture. C'est-à-dire qu'il parle rarement dans les fréquences les plus basses ou les plus hautes qu'il peut atteindre sauf, par exemple, dans des cas de conversation à forte valeur

---

1 Piet Mertens, *Phonétique française*, Leuven, Acco, 2005, p. 42.

2 Pour nos mesures, nous réalisons une conversion des Hz en ST (demi-tons) relativement à 1 Hz.

3 Généralement, la tessiture d'une voix de femme est plus haute que celle d'un homme. Les différences de tessiture sont provoquées entre autres par les différences physiologiques de longueur, d'épaisseur et de tension des cordes vocales.

émotive. La partie de la tessiture qu'un locuteur utilise habituellement est appelée son **registre** usuel. Les différences de registre entre deux locuteurs n'empêchent pas des similitudes entre les contours mélodiques utilisés par ces locuteurs. De même, deux locuteurs à la tessiture identique peuvent exploiter des étendues différentes de registre, créant par là un style vocal qui leur est propre.

La **longueur**, ou la **durée**, permet d'opposer des syllabes (ou des pauses) brèves et longues. Le débit en fournit une mesure globale. Le débit de parole correspond au nombre de syllabes réalisées par unité de temps, en tenant compte des pauses silencieuses. Le débit d'articulation<sup>1</sup> correspond au nombre de syllabes par seconde sur le temps d'articulation, c'est-à-dire en soustrayant les silences et les pauses silencieuses. Plus le débit est élevé, plus la durée des syllabes réalisées est courte.

L'**intensité**, quant à elle, permet de distinguer les sons forts et les sons faibles. Mesurée en dB, l'intensité correspond à l'amplitude des vibrations sonores, à l'énergie avec laquelle les cordes vocales sont mises en vibration. L'intensité élevée d'un son contribue à l'impression du volume élevé de ce son. Toutefois, en radio, les sons ont tendance à être égalisés. On ne perçoit donc que très peu de modifications d'intensité.

### Représentation de la hauteur, de la durée et de l'intensité

Chaque syllabe est caractérisée par une certaine hauteur mélodique, par une durée et par une intensité. L'outil d'analyse prosodique Prosogramme, développé par Piet Mertens<sup>2</sup>, permet de calculer automatiquement les valeurs de chaque syllabe (f<sub>0</sub>, durée, intensité) et d'en donner une représentation graphique.

Un prosogramme se lit comme une partition musicale.

- 
- 1 Un débit moyen correspond à 5 ou 6 syllabes par seconde. Un débit rapide peut aller jusqu'à 9 syllabes articulées à la seconde, tandis qu'un débit lent peut descendre jusqu'à 3 ou 4 syllabes par seconde.
  - 2 Piet Mertens, «Le prosogramme : une transcription semi-automatique de la prosodie», *Cahiers de l'Institut de Linguistique de Louvain* 30/1-3, 2004, pp. 7-25. Le script Prosogramme fonctionne dans le programme d'analyse acoustique Praat (Paul Boersma et David Weenink, 2008, *Praat: doing phonetics by computer (Version 5.0.20) [Computer program]*). Page consultée le 10 avril 2008, <http://www.praat.org/>.

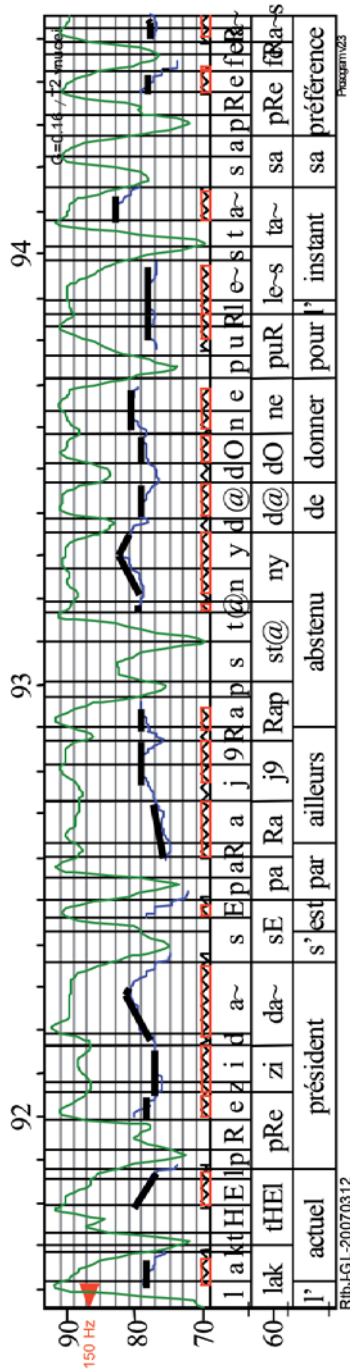


Figure 2. Prosogramme de l'extract l'actuel président s'est par ailleurs abstenu de donner pour l'instant sa préférence (RTB-E-JP2)

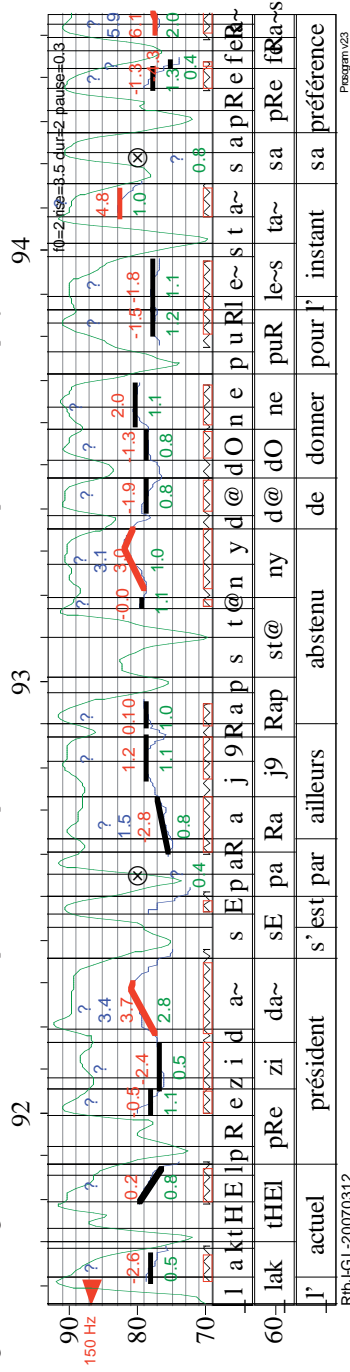


Figure 3. ProsoProm de l'extract l'actuel président s'est par ailleurs abstenu de donner pour l'instant sa préférence (RTB-E-JP2). Chaque syllabe est décrite par sa longueur relative (chiffre souscrit), par sa hauteur relative et par son mouvement mélodique dynamique. Par exemple, la dernière syllabe de *président* [da~] est 2.8 fois plus longue que les trois syllabes qui l'entourent, 3.7 demi-tons plus haute et contient un mouvement mélodique de +3.4 demi-tons. Elle est proéminente.



- Chaque syllabe voisée est stylisée<sup>1</sup> par un trait noir qui représente sa hauteur mélodique sur une échelle en demi-tons (les valeurs en demi-tons, calculés relativement à 1 Hz, sont représentées sur le bord gauche du schéma et la valeur absolue de 150 Hz est donnée à titre indicatif). La fréquence fondamentale non stylisée est représentée par une fine ligne discontinue qui apparaît derrière les syllabes stylisées.
- La longueur de la syllabe correspond à la longueur du trait qui la représente. Cette longueur est mesurable grâce à une graduation en dixièmes de secondes sur le bord supérieur du schéma.
- La variation de l'intensité est représentée par une fine courbe continue, en forme de montagnes et de vallées, dans la partie supérieure du prosogramme. Les pics d'intensité correspondent en partie aux noyaux syllabiques. Ils sont utilisés par le prosogramme pour la segmentation de la courbe mélodique.
- Sous le graphique, trois lignes d'annotation fournissent une segmentation en mots graphiques, en syllabes et en phones (en utilisant l'alphabet phonétique SAMPA).

À partir du prosogramme, des mesures de hauteur, de durée ou d'intensité peuvent être prises automatiquement, sur tout ou partie du signal sonore.

### Représentation de l'accentuation

Certaines syllabes sont allongées ou elles contiennent un mouvement mélodique qui leur donne un caractère particulièrement saillant. C'est le cas de la dernière syllabe du mot *président*, illustrée à la Figure 2, qui est allongée et montante. Ces syllabes sont perçues comme proéminentes: elles se détachent des syllabes environnantes et on considère généralement qu'elles sont accentuées.

L'outil ProsoProm<sup>2</sup> permet de détecter automatiquement les syllabes proéminentes sur la base d'une comparaison de chaque syllabe

---

1 Le prosogramme fournit une représentation segmentée et stylisée de la F0. Ce parti pris est justifié par un modèle de perception sur la base duquel ne sont représentées que les variations audibles de la courbe mélodique. La segmentation de la courbe s'effectue en identifiant le noyau (vocalique) de chaque syllabe.

2 A.C. Simon, J.-Ph. Goldman & M. Avanzi. (2008). La détection des proéminences syllabiques. Un aller-retour entre l'annotation manuelle et le traitement automatique.

avec les syllabes qui l'entourent. Les syllabes plus longues, plus hautes ou plus dynamiques sont considérées comme proéminentes et sont représentées par un trait gris (Figure 3).

On a souvent décrit le style journalistique comme usant (et abusant) d'accents initiaux sur la première syllabe des mots<sup>1</sup>. Dans notre étude, nous nous sommes contentés d'identifier les syllabes proéminentes sans les catégoriser selon leur localisation (début ou fin de mot). La proportion des syllabes proéminentes a tendance à augmenter dans les styles plus formels (discours politique, journal parlé) et à diminuer dans les styles informels (conversation spontanée) où le débit est souvent plus rapide.

## Méthodologie

### Présentation des données

Pour les besoins de la recherche, il fallait identifier des émissions comparables sur les deux chaînes étudiées. C'est pourquoi le corpus se compose d'enregistrements de journaux parlés (JP) sur La Première et sur Bel-RTL. Les JP sont des émissions qui, tant du point de vue de leur contenu que de leur forme, se ressemblent. De plus, les JP symbolisent le lieu où les intentions de la chaîne et la relation qu'elle entretient avec ses auditeurs se marquent le plus. Elles s'y dévoilent. Les choix des sujets traités et des angles développés sont faits en fonction des attentes du public ciblé par la chaîne. En outre, les JP ont la particularité de s'articuler autour d'une structure simple qui facilite le traitement du corpus.

Au total, 14 journaux parlés présentés par 8 journalistes (4 par chaîne) ont été recueillis. Pour 8 journaux, la partie où le journaliste énonce les titres a été transcrite et traitée séparément. En outre, deux journalistes de La Première ont accepté de se prêter à une rencontre informelle, destinée à recueillir un échantillon de leur parole hors de l'exercice de leur fonction professionnelle (il s'agit des enregistrements «conversationnels» RTB-S-CO et RTB-E-CO).

---

*Actes du Congrès mondial de Linguistique française, Paris.*

1 Voir Monique Callamand, «Aspects prosodiques de la communication», *Études de linguistique appliquée* 66, 1987, Paris, Didier.

Pour les journaux parlés, seules les interventions du présentateur du JP ont été retenues, c'est-à-dire les lancements de sujets et les brèves. Les questions du présentateur à un journaliste au cours des directs n'ont pas été retenues parce qu'il ne s'adresse plus uniquement au public et que l'hypothèse de recherche se centre sur la relation de la chaîne avec son public. Le corpus se compose de 156 minutes, enregistrées pendant la tranche d'informations de huit heures du matin<sup>1</sup>.

La partie traitée du corpus (durée utile) correspond à 78 minutes d'enregistrement réparties entre les deux chaînes (41 minutes pour La Première, 37 pour Bel-RTL). Les titres de huit JP sur les quatorze traités ont été retenus; ils correspondent à environ 7 minutes, tandis que les deux extraits de conversation totalisent 9 minutes (en ne gardant que les interventions du journaliste et en supprimant celles de son interlocuteur). Les JP eux-mêmes, en totalisant 62 minutes de parole, constituent la partie la plus importante des données. Le corpus complet compte 23.793 syllabes qui ont fait l'objet des mesures détaillées ci-dessous.

### Traitement des données

Notre étude vise principalement à comparer différents phonostyles, qu'il s'agisse des styles d'un même journaliste dans des situations différentes ou des styles de journalistes qui officient sur des chaînes radio différentes. Le traitement des données doit donc avant tout permettre la comparaison des différents enregistrements (journaux parlés, titres et conversations) qui constituent notre corpus.

La première étape du traitement consiste simplement à transcrire orthographiquement les enregistrements et à aligner cette transcription avec le signal sonore, au niveau des mots, des syllabes et des sons (ou segments)<sup>2</sup>. Le script Prosogramme (voir 2.2) est appliqué aux enre-

---

1 Un journal parlé date de 2006; huit journaux ont été enregistrés entre le neuf et le quatorze mars 2007 et cinq journaux ont été enregistrés en avril 2008.

2 À partir d'une transcription orthographique réalisée manuellement, l'alignement est réalisé de manière semi-automatique avec le script EasyAlign (Jean-Philippe Goldman), <http://latlcul.unige.ch/phonetique/>. Le résultat de l'alignement est donné sous la forme d'un fichier texte au format TextGrid édité sous le logiciel Praat et contenant des index temporels correspondant aux différentes unités transcrites (mots, syllabes, phones).

gistements alignés. Il effectue une détection et une stylisation de la  $f_0$  et il fournit les valeurs de hauteur, de durée et d'intensité de chaque noyau syllabique. Ces valeurs servent de point de départ aux calculs réalisés par le script ProsoProm (voir 2.3) et débouchent sur une identification automatique des syllabes proéminentes. Une dernière étape est réalisée via le script ProsoReport<sup>1</sup>. Ce dernier fournit, sous la forme d'un tableau, une "photographie générale" des propriétés prosodiques d'un enregistrement. Il calcule le débit moyen du locuteur, son registre et son étendue; il compare la proportion de syllabes proéminentes et non proéminentes, et précise quels types de mouvements mélodiques (montants, descendants, plats) sont les plus fréquents; etc.

Les 24 enregistrements de notre corpus ont été traités par ProsoReport de sorte à produire une série de mesures permettant de comparer des (groupes) d'enregistrements, selon nos hypothèses. Ces mesures sont détaillées et analysées au point suivant.

## Mesures et interprétation des mesures

Parmi les nombreuses mesures possibles de l'outil ProsoReport, nous n'avons retenu que celles qui permettent de tester nos hypothèses de recherche, principalement sur la mélodicité. Nous avons formulé deux hypothèses pour répondre à notre double question de recherche.

### Hypothèses et indicateurs

Comme l'a observé le sociolinguiste William Labov<sup>2</sup>, n'importe quel locuteur varie sa manière de parler en fonction de la situation dans laquelle il se trouve. C'est la variation stylistique, ou situationnelle. Dans ce cadre, nous faisons l'hypothèse (H1) que le phonostyle du journaliste répond à une mise en fonction professionnelle. Il parle différemment selon qu'il est en contexte de présentation de journaux parlés

---

1 Goldman, J.-Ph., A. Auchlin, A.C. Simon & M. Avanzi. 2007. "Phonostylographe : un outil de description prosodique. Comparaison du style radiophonique et lu". *Nouveaux cahiers de linguistique française* 28, 219-237. <http://clf.unige.ch/num.php?numero=28>

2 William Labov, *Language in the inner city: studies in the black English vernacular*, Oxford, Blackwell, 1976 (traduction en français par A. Kihm, *Le parler ordinaire: la langue dans les ghettos noirs des Etats-Unis*, Paris, Minuit, 1978).

ou en conversation informelle. Pour vérifier cette première hypothèse, nous avons inclus dans le corpus l'enregistrement de deux conversations de journalistes de La Première afin d'en comparer le phonostyle avec celui utilisé lors de leur passage à l'antenne.

- Nous avons comparé ces phonostyles selon le **débit d'articulation**, c'est-à-dire le nombre de syllabes prononcées par seconde.
- Nous avons également comparé **les tessitures** en contexte de conversation et de présentation.

Deuxièmement, nous faisons l'hypothèse (H2) que le phonostyle des journalistes de Bel-RTL est plus mélodique que celui des journalistes de La Première, conformément à la théorie développée par Jean-Pierre Meunier et Daniel Peraya<sup>1</sup> qui lie la proximité entretenue avec l'auditeur et degré de mélodicité du style vocal.

Pour vérifier la seconde hypothèse, il importe de dégager les indicateurs de la mélodicité du phonostyle. Trois mesures ont été retenues.

- Nous avons d'abord mesuré **l'étendue de la fréquence fondamentale** (f0) des journalistes étudiés. Il s'agit de « *la plage de fréquence exploitée* » dans la tessiture du locuteur<sup>2</sup>. Elle se mesure en demi-tons. Pour valider H2, l'étendue doit être plus importante pour les journalistes de Bel-RTL
- La deuxième mesure retenue est la **proportion de syllabes statiques** par rapport aux syllabes dynamiques. Plus cette proportion est réduite, plus la mélodicité est réduite. Pour valider H2, il faut donc que cette proportion de syllabes statiques soit plus petite pour le phonostyle des journalistes de Bel-RTL
- Nous avons ensuite mesuré **l'agitation mélodique**. Il s'agit du chemin moyen parcouru par la f0 en une seconde. Elle se mesure en demi-tons. Pour valider H2, cette agitation doit être plus importante pour le phonostyle des journalistes de Bel-RTL.
- La dernière mesure est le calcul du **pourcentage de syllabes proéminentes**. On considère qu'une syllabe est proéminente lorsqu'elle ressort par sa longueur ou par sa hauteur particulière.

---

1 J.-P. Meunier, D. Peraya, *Introduction aux théories de la communication, Analyse sémio-pragmatique de la communication médiatique*, Ed. De Boeck, Coll. Culture et communication, Bruxelles, 2004.

2 L. Ménard, «Le registre en tant que marqueur prosodique de dialectalité», dans M. Demers, *Registre et voix sociale*, Nota Bene, Québec, 2003, p. 36.

Plus les syllabes proéminentes sont nombreuses plus l'impression de mélodicité est grande. Il faut donc que cette mesure soit plus importante pour les journalistes de Bel-RTL pour valider H2.

### **Vérification de H1 : le journaliste marque sa prise de fonction à l'antenne par un changement phonostylistique**

Pour valider H1, il faut constater une différence entre le phonostyle du journaliste en conversation et à l'antenne.

H1 se vérifie du point de vue du **débit d'articulation**. On constate une variation du débit pour les deux journalistes. L'un d'eux ralentit en passant de la conversation à la présentation du journal. L'autre accélère. Mais tous deux convergent vers un débit très similaire (6.06 syll/sec et 5.91 syll/sec). Ce débit vers lequel ils tendent est comparable à la moyenne des débits de tous les journalistes dans notre corpus (5.80 syll/sec). Cette mesure connaît un écart-type faible. Du point de vue du débit on pourrait donc conclure que les journalistes modifient leur phonostyle de départ pour aller vers un standard du débit d'articulation partagé par l'ensemble des journalistes pour la présentation des journaux parlés.

L'hypothèse de la mise en fonction professionnelle se vérifie également du point de vue du **registre moyen (f0 moyenne)** des locuteurs observés. S'il n'est pas pertinent de comparer la hauteur moyenne de la voix d'un journaliste à un autre (chacun ayant sa propre tessiture), il est intéressant d'observer que chaque journaliste, compte tenu de sa tessiture, peut modifier son registre usuel (étendue de f0) dans une situation particulière. Les deux journalistes pour lesquels nous disposons de cette mesure ont une f0 moyenne plus élevée quand ils passent à l'antenne. Le registre utilisé à l'antenne est donc plus élevé que celui utilisé en conversation.

L'**étendue** du registre est aussi soumise à variation. On constate qu'un des journalistes augmente cette étendue à l'antenne en exploitant une plus large palette de fréquences mélodiques. Cela se traduit donc par une mélodicité accrue à l'antenne<sup>1</sup>. On peut interpréter l'augmentation de l'étendue de la f0 comme une exploitation du "code d'effort" étudié par Gussenhoven à la suite des travaux de John Ohala: "*Increases in*

---

1 Pour des raisons techniques, la mesure n'a pas pu être réalisée pour l'autre journaliste. Il n'y a donc pas ici de généralisation possible.

*the effort expended on speech production will lead to greater articulatory precision, but also a wider excursion of the pitch movements»<sup>1</sup>. En dépensant plus d'énergie, avec comme conséquence une augmentation de l'étendue mélodique, le locuteur signifie l'importance du contenu de son message.*

L'**agitation mélodique** est plus importante en conversation qu'en présentation. Il y a donc une différence entre les deux contextes de parole. Mais cet indicateur va dans un sens opposés à ceux commentés jusqu'à présent, puisqu'il indique que la mélodicité serait moins importante à l'antenne qu'en conversation.

En revanche, il n'y a pas de différence en ce qui concerne la proportion de syllabes **proéminentes**. Le taux de proéminences reste stable en conversation et à l'antenne pour les deux journalistes. Toutefois, il y a probablement une différence dans le placement des proéminences, et donc des accents, que la méthode utilisée ne permet pas de vérifier. Pour Callamand, citée par Léon, *l'une de caractéristiques les plus souvent notées [du phonostyle journalistique] est la prolifération de l'accent didactique, qui se marque non seulement par une accentuation forte mais aussi par le coup de glotte<sup>2</sup>.*

La première hypothèse d'un changement de phonostyle lors de la prise de fonction du journaliste est donc partiellement vérifiée. Des différences de phonostyle sont effectivement observables entre les conversations et le passage à l'antenne. Sur les trois observations mélodiques (moyenne et étendue de f<sub>0</sub>; agitation mélodique), deux vont dans le sens de plus de mélodicité à l'antenne. Ce qui confirme en partie l'hypothèse de Daniel Peraya et Jean-Pierre Meunier.

## **Vérification de H2 : le phonostyle des journalistes de Bel-RTL est plus mélodique que celui de journalistes de La Première**

La validation de H2 passe par une mesure plus détaillée de la mélodicité du phonostyle des journalistes de Bel-RTL.

- 
- 1 C. Gussenhoven, Intonation and interpretation: Phonetics and phonology. In Bel, B. & I. Marlien (eds.) Proceedings Speech Prosody 2002, Aix-en-Provence, Laboratoire Parole et Langage, Université de Provence, p. 50.
  - 2 M. Callamand (1987), citée par Léon, *Précis de phonostylistique*, 1993, p. 165.

En ce qui concerne l'**étendue de f<sub>0</sub>**, H2 ne se vérifie pas. Les moyennes des étendues pour chaque chaîne sont très proches (11.84 ST pour Bel-RTL, 11.69 ST pour La Première). Si on examine la distribution de cette mesure, on constate qu'elle est quasiment similaire pour 6 des 8 journalistes. Un journaliste de chaque chaîne se démarque toutefois par une étendue nettement plus importante (13.65 et 15.70 ST)<sup>1</sup>.

Pour ce qui est de la **proportion de syllabes statiques** (avec une f<sub>0</sub> plate), les moyennes sont à nouveau proches pour les deux chaînes. En revanche la variation entre les journalistes, quelle que soit la chaîne, est très grande (de 57% à 76% pour Bel-RTL et de 62% à 73% pour La Première). Cette mesure ne semble donc pas aller dans le sens d'une mélodicité plus importante pour le phonostyle des journalistes de Bel-RTL

Lorsqu'on observe l'**agitation mélodique**, les moyennes indiquent une agitation plus importante pour les phonostyles de journalistes de La Première. On remarque toutefois que cette mélodicité plus importante est le fait d'un journaliste qui rehausse la moyenne. Si on retire cette valeur extrême on constate une assez grande similarité dans l'agitation mélodique des différents journalistes. Cette mesure infirme donc notre hypothèse.

L'étude des **syllabes proéminentes** révèle une similarité entre les deux chaînes observées. En moyenne, sur les deux chaînes, 27% des syllabes sont proéminentes<sup>2</sup>.

Toutefois, la mesure ne donne pas la nature des proéminences et une syllabe peut être proéminente par sa longueur et / ou par sa hauteur.

---

1 On observe aussi que l'étendue de la f<sub>0</sub> est systématiquement accrue, pour un même journaliste, dans la présentation des titres (vs le développement du journal). Les titres présentent une mélodicité plus importante, probablement liée à leur fonction d'accroche et à leur haute teneur informative (voir ci-dessus l'interprétation du «code d'effort» proposée par Gussenhoven). Ceci confirme indirectement H1: les journalistes adoptent des conduites vocales différenciées en fonction des tâches communicatives.

2 La proportion des syllabes proéminentes dans un discours varie de 25% (discours à débit rapide) à 50% (discours très formel, à débit lent, ou chaque (groupe de) mot(s) se voit attribuer un accent final (l'exemple typique est le discours officiel d'un jour de fête nationale). À titre de comparaison, une étude sur les phonostyles très typés des chroniqueurs de France Info (Goldman et al. 2007) a observé un taux de proéminence de 37%, supérieur à ce que nous observons dans notre corpus de journaux du matin (27%). À nouveau, les titres se distinguent de ce point de vue: le taux de proéminences passe à 35% pour un des journalistes dans nos données.



Seule une proéminence due à une hauteur accrue ajoute à la mélodicité du discours.

La seconde hypothèse ne se vérifie donc pas. Les mesures indiquent plus de ressemblances entre les chaînes que de différences. Ceci doit être nuancé à plusieurs titres. Premièrement, des différences au niveau d'autres indicateurs phonostylistiques ne sont pas à exclure. En effet, la méthode adoptée donne une description prosodique globale de chaque enregistrement, qui peut ainsi être comparé aux autres. Mais elle ne fournit pas d'indication sur des variations locales dans un même journal parlé<sup>1</sup>. Il n'est donc pas possible de voir si des ressemblances existent à ce niveau-là entre journalistes d'une même chaîne.

Une deuxième observation permet de nuancer notre conclusion. Il est important de noter que sur les huit journalistes, six obtiennent des scores proches pour la plupart des indicateurs, alors que deux journalistes (un pour chaque chaîne) se distinguent plus nettement. Leur étendue de  $f_0$  et leur agitation mélodique est plus élevée, ainsi que le taux de syllabes proéminentes qu'ils réalisent.

Enfin, nous n'avons pas tenu compte dans cette étude des différences phonostylistiques induites par les différences d'écriture entre les journalistes. Ces différences concernent le niveau grammatical, selon le modèle du double encodage (voir Figure 1), ce qui pourrait remettre en question la stricte indépendance posée par Fónagy entre le niveau de la Grammaire, où sont formées les phrases, et le Détracteur, qui «module» le message en fonction du style recherché. L'importance de l'écriture pour la vocalisation semble pourtant déterminante aux yeux des journalistes: *«le ton est clairement lié à la manière d'écrire les choses. Pour moi c'est clair, l'un ne va pas sans l'autre»*.<sup>2</sup>

- 
- 1 Le script ProsoReport donne des écarts-types pour chaque mesure, et donc une distribution des valeurs, mais il ne donne aucune représentation de l'évolution temporelle des paramètres (par ex. les endroits où le débit s'accélère ou ralentit).
  - 2 Déclaration recueillie au cours de l'un des entretiens informels réalisé avec un journaliste de La Première (RTB-G-CO).

## Conclusions

Cette étude visait à cerner en quoi consiste le style vocal des journalistes lorsqu'ils passent à l'antenne, en vérifiant si des chaînes radio-phoniques concurrentes, aux positionnements supposés différents, se distinguaient par ce biais.

Nous avons choisi d'adopter une hypothèse forte de Meunier et Peraya. Cette hypothèse suppose un lien entre mélodicité et proximité. Une première réserve doit être faite, puisque ce lien n'a pas été démontré. Une seconde réserve provient de notre méthode d'analyse qui, produisant des "rapports prosodiques" sur chaque enregistrement, en donne uniquement une vision globale.

Ceci étant dit, nous avons pu vérifier qu'il y a une mise en fonction professionnelle et qu'elle tend vers une standardisation du ton adopté durant les journaux parlés. Ceci est particulièrement vrai pour le débit d'articulation, qui est quasiment identique pour tous les journalistes durant les journaux. L'absence de différence significative dans la mélodicité des voix de la chaîne publique par rapport à la chaîne privée laisse penser que, lorsqu'il arrive à l'antenne, le présentateur est journaliste avant d'appartenir à une chaîne en particulier. Cette observation est confirmée par une impression d'un journaliste qui « *pense qu'il y a un patron commun [...] pour tous les journalistes radio et audiovisuels, avec des petites différences déjà entre radio et télé* ». Le phonostyle, par son caractère relativement stable et standardisé, fonctionnerait comme un gage de crédibilité et comme une marque de reconnaissance de la production médiatique en cours<sup>1</sup>.

---

1 I. Fónagy développe l'idée du style vocal comme signe d'appartenance à une profession donnée. « *Les membres des corporations ont [...] senti le besoin de se signaler en tant que membres d'un groupe professionnel, d'exprimer à l'aide des insignes et de vêtements représentant la profession leur appartenance à une corporation et le rang qu'ils y occupaient. Depuis qu'on a renoncé aux insignes et au langage vestimentaire, c'est le style verbal qui est censé jouer ce rôle. La prosodie [...] est prédestinée à cette fonction évocatrice, identificatrice.* » (I. et J. Fónagy, 1976, Prosodie professionnelle et changements prosodiques, *Le français moderne* 44/3, p. 193, cité par P. Léon, 1993, p. 157).

Toutefois, ce style prosodique commun n'exclut pas des déclinaisons entre chaînes et entre journalistes. D'une part, notre étude s'est limitée à l'analyse des voix, sans prendre en compte des aspects aussi importants que les indicatifs ou l'habillage sonore de la chaîne. D'autre part, nous avons constaté que, sur les huit journalistes analysés, deux se distinguent nettement du lot par une mélodicité accrue de leur style. Si la mélodicité est un facteur de proximité, ces deux journalistes seraient ceux qui parviennent le mieux à créer cette proximité avec les auditeurs.<sup>1</sup> C'est ici qu'une étude de perception, visant à mesurer l'impact de ces paramètres sur l'auditoire, donnerait une suite utile à notre recherche.

Le phonostyle des journalistes présente un caractère partiellement standardisé en même temps qu'il autorise une forme de distinction, la création d'une image de marque ou d'un style personnel. Il illustre une des fonctions de cette "vive voix" par laquelle les paroles, bien au-delà de leur contenu informationnel, créent de la signification, de l'identification sociales.

---

1 A ce titre, Pierre Léon, évoque la possibilité de certains journalistes de se singulariser grâce à leur ton. « *Les situations de communication radiophoniques et télévisuelles entraînent des types de verbalisation plus ou moins stéréotypés. Le degré de stéréotypie varie avec la personnalité du présentateur. Ce dernier peut fort bien créer sa propre batterie de stéréotypes (...). En dehors des styles de ces fortes personnalités se crée un phonostyle propre à chaque genre.* » (Pierre Léon, Précis de Phonostylistique, Parole et expressivité, Coll. Fac. Linguistique, Paris, 1993, p. 164.)

