

Perception des phonostyles et représentativité du phonogène : le cas du commentaire sportif en direct

Catherine Mathon

UMR 5263 CLLE-ERSS

Université Bordeaux Montaigne

<catherine.mathon@u-bordeaux-montaigne.fr>

Abstract

This paper reports a perceptual study conducted on 63 French subjects, in order to test the ability to discriminate different types of media comments with the prosodic information only. We used five categories of media comments: four categories represent comments for different types of sport (match, individual performance, artistic sport, race), the fifth category regroups non-sport media comments. The results show that subjects distinguish the artistic sport category on one hand, and the non-sport media comment on the other hand, from all the others sport categories, whereas rhythmic measures separate the five categories presented.

Mots-clés : prosodie, phonostyles, perception, commentaire sportif

1. Introduction

Nous proposons dans cette étude de tester la perception des phonostyles associés au commentaire sportif en direct de différents sports et plus largement au commentaire en direct d'événements médiatiques. Notre objectif est de déterminer si la seule information prosodique permet une catégorisation fine des phonostyles entre différentes catégories de commentaires en direct et notamment entre différentes catégories de sports commentés. Nous cherchons ainsi, au moyen d'un test de perception basé sur des stimuli filtrés, à définir quels sont les phonostyles les plus représentatifs du commentaire sportif en direct. Après avoir défini les notions de phonogène et de phonostyle au regard des études antérieures, et après avoir indiqué les critères retenus pour la classification des contraintes de genre et d'événement, nous décrivons le protocole expérimental qui a conduit au test de perception. Nous établissons ensuite les résultats statistiques de l'étude des réponses des sujets, que nous discutons avant de conclure.

2. Critères de classification des phonogènes

Notre réflexion part des travaux de Pršir, Goldman et Auchlin (2013), qui cherchent à établir une typologie de phonogènes, c'est-à-dire de genres oraux, en recoupant des situations de parole et des traits

prosodiques typiques. Dans leur étude, aucun phonogène n'est posé a priori, mais des critères de description pragmatique des différentes situations de parole présentes dans leur corpus, permettent de premiers regroupements. Ces traits situationnels sont organisés dans quatre dimensions :

- caractère médiatique
- type d'audience
- degré de préparation
- degré d'interactivité

Le commentaire sportif en direct, dans cette perspective, regroupe les mêmes caractéristiques quant à la situation de parole que le commentaire en direct d'autres événements médiatiques, tels qu'un mariage princier, ou la situation électorale devant un QG de campagne, un défilé militaire. Pour autant, notre intuition est que la réalisation prosodique, c'est-à-dire le phonostyle, associé au commentaire de chacun de ces événements sera sans doute très différent.

Par ailleurs, en ce qui concerne le phonogène du commentaire sportif, Pršir et al. (2013) remarquent que le commentaire sportif, au sein de leur corpus composé de commentaires de trois sports différents (foot, rugby et basket), est le phonogène le moins compact, en termes de spécificités prosodiques. Autrement dit, au sein d'un même phonogène, on rencontre une variation prosodique importante. Cette dispersion est attribuée à « des dynamiques cinétiques différentes qui sont reflétées dans la parole » (Pršir et al., 2013).

3. Le commentaire sportif en direct : un phonogène mais des phonostyles

Sur la base des études antérieures et de nos propres recherches (Augendre et al., 2014), nous considérons le phonogène comme un ensemble structuré de contraintes, et le phonostyle, la réalisation prosodique engendrée par ces contraintes. On suppose donc que, selon la combinaison des contraintes pertinentes au moment de la production du discours, le phonostyle fera l'objet de variations. Cela nous amène à considérer qu'il existe une variété de phonostyles pour un même phonogène.

3.1. Le commentaire sportif : un discours sous contraintes

Le commentaire sportif en direct, en tant que phonogène, est un discours sous contraintes. Celles-ci sont liées d'une part aux conditions de production de la parole, et d'autre part à la structure de l'événement commenté.

3.1.1. Contraintes liées au phonogenre

Nous reprenons à Pršir et al. (2013) l'idée que l'on peut établir une typologie des phonogenres en fonction de critères situationnels, que nous considérons comme des contraintes appliquées au discours. Nous retenons comme contraintes de production du discours :

- Le caractère médiatique du commentaire sportif. Cette condition amène le commentateur à réaliser son discours en portant des « émotions », qui sont à destination de son audience.
- Le caractère synchrone du commentaire sportif. Une partie du commentaire, notamment les phases de description, est réalisée de manière quasi-simultanée, avec l'action sportive.
- Le caractère « en direct » du commentaire sportif. Le commentaire est réalisé de manière spontanée, sans que les images commentées aient pu être visualisées à l'avance par le commentateur. Celui-ci prend connaissance des images en même temps qu'il les commente.

3.1.2. Contraintes liées à la structure de l'événement commenté

Certaines contraintes sont liées non plus seulement au phonogenre, mais également à la structure de l'événement commenté. Nous retenons comme contraintes de structure :

- La durée des actions commentées, qui va influencer directement le type de discours produit (purement descriptif ou apport d'informations complémentaires en marge de l'action). L'objectif pour le journaliste est d'apporter, quand il le peut, un maximum d'informations complémentaires aux images diffusées à l'attention du téléspectateur.
- Le caractère prééminent (qui doit être mentionné) ou non de l'action en cours.
- Le caractère marquant (qui a un impact sur la réalisation dans son ensemble) de l'action.

3.1.3. Contraintes liées au type de sport

Pour ce qui est plus particulièrement du commentaire sportif, nous considérons comme Audrit et al. (2012), que le rythme et la dynamique des sports ont une influence sur le discours et vont amener des contraintes variées, auxquelles le style du journaliste va nécessairement devoir s'adapter. Pour prendre en compte ces variations, nous proposons de classer les sports selon différentes catégories :

- Performance séparée et individuelle [Perf. Indi.] : Les compétiteurs effectuent leur performance, seuls, indépendamment des autres (exemples : slalom, saut en hauteur).
- Performance à dimension artistique, ou expression corporelle [Expr. Corp.] : L'émotion artistique dégagée par le compétiteur est prise en compte tout autant que les qualités sportives et techniques (exemples : patinage, nage synchronisée).
- Performance parallèle [Course] : Les compétiteurs effectuent leur performance en parallèle et simultanément aux autres (exemples : épreuve du cent mètres, aviron).
- Combat/Match [Match] : La compétition se fait entre deux adversaires (exemples : judo et rugby). Pour les matches, le jeu est le résultat de l'interaction non seulement entre les deux équipes adverses mais également à l'intérieur de chaque équipe.

3.2. *Le phonostyle : une réalisation des contraintes*

Comme le synthétise le schéma ci-dessous (cf. Figure 1) nous considérons que le phonostyle n'est pas seulement le résultat des contraintes liées au phonogène, mais plutôt le résultat du croisement des contraintes du phonogène avec celles du type d'évènement décrit, et pour le commentaire sportif, du type de sport. Par ailleurs, nous considérons que la structure interne à l'évènement commenté influence la production du discours et amène une variété de phonostyles correspondants.

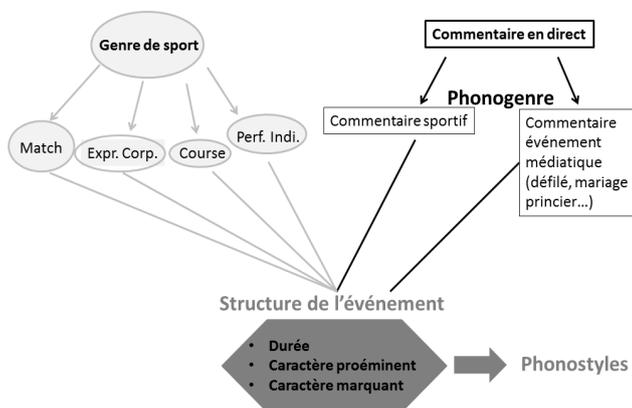


Figure 1 : Représentation des contraintes à l'œuvre dans la production des phonostyles de commentaire sportif en direct

4. Protocole expérimental

Nous partons du principe que le téléspectateur, sans avoir une pleine conscience des contraintes de production du discours, possède une ou des représentations du commentaire sportif en tant que phonogène et attend de la part du journaliste sportif certaines réalisations prosodiques prototypiques. L'objectif de notre étude est de déterminer si, en perception, certains phonostyles sont plus particulièrement représentatifs du commentaire sportif en direct.

4.1. Hypothèses

Nous souhaitons vérifier les hypothèses suivantes :

- La seule information prosodique permet de différencier le commentaire sportif du commentaire en direct d'un événement médiatique (type mariage princier, défilé militaire, ...), même si les deux commentaires appartiennent à un même méta-phonogène : le commentaire médiatique en direct.
- La seule information prosodique permet une distinction fine entre les différents types de sport et leur commentaire : match, course, performance individuelle et expression corporelle.

4.2. Corpus

Notre corpus, quant au commentaire sportif, est basé à l'origine sur un corpus d'enregistrements de matchs de rugby diffusés lors de la Coupe du Monde de Rugby 2007 (Lortal et Mathon, 2009). Nous utilisons le commentaire en français du match d'entrée dans la compétition : France-Argentine. Cet enregistrement a fait l'objet d'une transcription orthographique avec Transcriber et d'annotations discursives permettant notamment de différencier les périodes descriptives des périodes de commentaires (Augendre et al., 2012). Ce commentaire a été utilisé comme représentatif de la catégorie de sport « match ».

Pour compléter le corpus de commentaires sportifs avec des sports correspondants aux catégories établies en 3.1.3., nous avons sélectionné les commentaires en français de différentes épreuves qui ont pris place durant les Jeux Olympiques d'Hiver de Sotchi (2014) :

- Snowcross pour la course [Course]
- Slalom masculin et Snowpipe pour la performance sportive individuelle [Perf. Indi.]
- Patinage artistique pour la performance sportive et artistique [Expr. Corp.]

Pour ce qui est du commentaire des autres événements médiatiques, nous avons sélectionné les enregistrements de la soirée

électorale pour les élections législatives de 2012, et le commentaire du défilé militaire du 14 juillet 2013.

4.3. Sélection des stimuli

Les stimuli sont des extraits de chacun des commentaires de chaque type de sport et de chaque événement médiatique. Nous avons sélectionné ces extraits en fonction des critères définis dans la section 3.

4.3.1. Critères de sélection

Les stimuli ont été sélectionnés et extraits des commentaires de chaque catégorie d'événement médiatique, sportif ou non, avec l'idée d'obtenir un échantillon correspondant à chacun des critères identifiés plus haut. Nous avons ainsi différencié les événements commentés selon :

- Leur durée : action ponctuelle ou durative
- Leur importance (Proéminence)
- La valeur émotionnelle dégagée par le commentaire (Émotion)

Le tableau 1 marque pour chaque catégorie d'événement médiatique la répartition des stimuli selon les critères Durée, Proéminence et Émotion.

		Durée		Proéminence		Émotion	
		Ponctuel	Duratif	-	+	-	+
Événement sportif	Match	2	6	4	4	2	6
	Course	1	3	3	1	1	3
	Perf. Indi.	5	5	3	7	5	5
	Expr. Corp.	2	2	2	2	1	3
Autre événement médiatique	Défilé militaire	0	1	1	0	1	0
	Soirée électorale					4	7
Total		10	19	14	15	19	16

Tableau 1 : Répartition des stimuli selon les critères Durée, Proéminence et Émotion

Nous avons sélectionné au total 38 stimuli répartis comme suit dans chaque catégorie : Match (8), Course (4), Performance individuelle (10), Expression corporelle (4), Événement médiatique non sportif (12).

4.3.2. Délexicalisation

Les stimuli sélectionnés ont fait l'objet d'un filtrage passe-bas à l'aide du logiciel Praat (Boersma et Weenink, 2001). Le filtrage permet en supprimant les hautes fréquences de masquer l'information verbale tout en conservant l'information mélodique et rythmique. Plusieurs valeurs de fréquence pour le filtrage ont été testées, avant qu'un filtrage 500-1000 Hz soit finalement retenu. Le filtrage affecte l'intensité perçue du signal, de telle manière qu'on a l'impression que le son est produit à une grande distance. Pour contrebalancer cet effet, les stimuli après filtrage ont fait l'objet d'une augmentation de l'amplitude.

La délexicalisation des stimuli est un processus difficile quand il est fait a posteriori, sur des stimuli spontanés, pour lesquels on souhaite conserver au maximum l'information prosodique, tout en effaçant de manière efficace le contenu lexical. Ainsi, indépendamment du filtrage passe-bas, nous avons testé la resynthèse (toujours à l'aide de Praat). Avec cette méthode, l'information lexicale est parfaitement supprimée, mais le signal sonore a fait l'objet d'une telle modification qu'il est difficile de le reconnaître comme provenant de la parole humaine. Nous avons donc préféré, malgré ses imperfections, le filtrage passe-bas. Pour nous assurer de son efficacité quant à la suppression de l'information lexicale, nous avons demandé à nos sujets d'indiquer le degré d'accès à l'information lexicale.

4.4. Test de perception

Le test de perception a été conduit sur 63 sujets francophones (étudiants en première année et deuxième année de licence de Sciences du Langage, et en première année de licences LLCE allemand, russe, japonais). Il a été présenté par set de 12 stimuli sous la forme d'un Test de type CLOZE hébergé sur la plateforme Ecampus (Moodle) de l'Université Bordeaux Montaigne.

L'utilisation de l'interface Moodle pour la diffusion du test de perception nous a permis de toucher un panel large de sujets à distance, qui ont l'habitude de faire des exercices sous ce type de format et qui en maîtrisent donc l'interface. Les sujets ont reçu des consignes précises sur les conditions optimales pour effectuer le test : calme, utilisation d'un casque audio dans l'idéal, effectuer le test en une seule fois sans interruption. Les réponses au test sont fournies sous la forme d'une feuille Excel. Parmi les informations fournies en complément des réponses, est indiquée la durée d'exécution de l'exercice. En moyenne, les sujets ont mis entre 5 et 10 minutes pour effectuer le test. Les réponses des sujets qui avaient outrepassé

significativement la durée moyenne du test n'ont pas été prises en compte.

Pour chacun des douze stimuli du test, les sujets devaient répondre à deux questions présentées comme des questions à choix multiples. Dans un premier temps, ils devaient indiquer si d'après eux l'événement médiatique commenté était un événement sportif un autre événement médiatique (type mariage princier, défilé de mode). Dans un second temps, ils devaient indiquer quelle catégorie d'événement était couverte par le commentaire entendu : match, performance sportive individuelle, expression corporelle, course ou événement médiatique non sportif.

5. Résultats

Les réponses des sujets à la question 2 et la catégorie (Match, Perf. Indi., Expr. Corp., Course ou Autre événement médiatique) correspondant à chaque stimulus ont fait l'objet d'une analyse de correspondance à l'aide du logiciel SPSS. Le graphe ci-dessous (Fig. 3) représente les résultats de l'analyse de correspondance entre la catégorie appliquée à chaque stimulus et les réponses des sujets à la question 2 du test de perception (attribution d'une des catégories proposées au stimulus entendu).

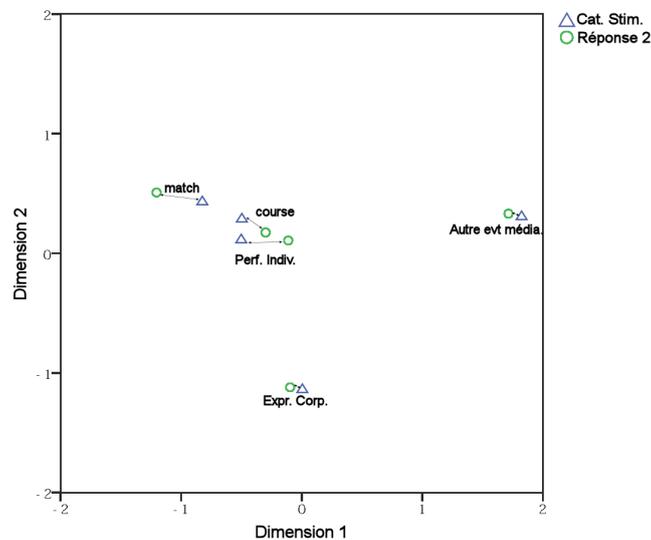


Figure 2 : Résultats de l'analyse de correspondance entre la catégorie des stimuli et les réponses des sujets à la question 2

Le graphe permet dans un premier temps de représenter la distance entre chaque catégorie et les réponses des sujets. On peut ainsi noter que cette distance est moindre quant aux catégories Expression Corporelle et Autre événement médiatique, ce qui laisse entendre une meilleure reconnaissance de ces deux catégories par les sujets. Le graphe permet également de mettre à jour trois pools d'événements. Les catégories Match, Course et Performance sportive individuelle semblent regroupées, tandis que la catégorie Autre événement médiatique d'une part et Expression corporelle d'autre part sont isolées chacune dans un coin du graphe.

Catégorie Stimuli	Réponses Q2				
	Match	Expr. Corp.	Perf. Indi.	Course	Autre évt. Média.
Match	66.7	7.1	11.9	11.9	2.4
Expr. Corp.	4.75	76.2	4.75	1.6	12.7
Perf. Indi.	32.8	10	20.6	24.4	12.2
Course	43.7	6.3	14.3	26.2	9.5
Autre évt. Média.	4.4	7.5	11.5	9.1	67.5

Tableau 2 : Matrice de correspondance des catégories des stimuli avec les réponses des locuteurs à la question 2. Les réponses sont transcrites en pourcentages

La matrice de correspondance des catégories des stimuli avec les réponses des locuteurs (cf. Tableau 2) montre que pour les catégories de stimuli, match, expression corporelle et autre événement médiatique, plus de 65% des sujets donnent la catégorie correspondante. Il semblerait donc que ces trois catégories ne fassent pas l'objet d'autant de confusion. La catégorie Expression corporelle semble particulièrement bien perçue avec 76.2% des réponses correspondantes.

6. Analyse rythmique des stimuli contrastés

Le test de perception nous a permis d'identifier trois catégories d'événements. Ainsi, l'expression corporelle semble perçue différemment des autres catégories de sport, et la catégorie Autre événement médiatique est également isolée des deux autres. Afin de déterminer quels pourraient être les différences perceptibles, nous avons procédé sur l'ensemble des stimuli à des mesures de débit (débit de parole et d'élocution) en syllabes par seconde et à des mesures des pauses en seconde.

Un test non-paramétrique de Kruskal-Wallis a été effectué pour déterminer si les valeurs des pauses, du débit de parole et du débit d'élocution étaient significativement différentes en fonction de chaque

catégorie (Match, Perf. Indi., Expr. Corp., Course ou Autre évt. Média.). Pour ce qui est des pauses, le test de Kruskal-Wallis montre une différence significative entre chacune des catégories avec une valeur de $p=0.012$. Il en est de même pour le débit de parole qui présente une différence significative entre chaque catégorie. En revanche, la différence entre chaque catégorie pour le débit d'élocution n'est pas significative ($p=0.103$). Le débit d'élocution ne prenant pas en compte les pauses des locuteurs, on peut faire l'hypothèse que ce sont les pauses, et leur durée qui permet essentiellement de faire la différence entre chaque catégorie.

Les valeurs mesurées, que ce soit de pause ou de débit, sont des valeurs globales pour l'ensemble du stimulus, et ensuite pour l'ensemble de la catégorie à laquelle appartient le stimulus. Elles ne permettent pas par exemple de rendre compte des différences intra-catégorie entre des événements duratifs et des événements ponctuels, ou à un niveau encore plus micro des différences de rythme associées aux différentes phases d'un événement duratif.

7. Discussion et conclusion

Le test de perception que nous avons mené nous a permis de vérifier que la seule information prosodique pouvait permettre de catégoriser de manière relativement fine des commentaires médiatiques en direct et notamment de distinguer le commentaire sportif du commentaire d'autres événements médiatiques. Nous avons vu également que les sujets isolaient le commentaire d'une activité sportive à dimension artistique des autres commentaires sportifs. En revanche, à l'intérieur de la catégorie du commentaire sportif, la distinction entre les différents sports n'est pas aussi facilement faite. Le match est la catégorie qui est la mieux reconnue, peut-être parce qu'il s'agit du mode sportif le plus diffusé (match de football, de rugby, ...), et pour lequel nous avons donc construit le plus de représentations. Les catégories Course et Performance individuelle font au contraire l'objet d'une confusion, le plus souvent avec la catégorie Match.

Au vu de ces premiers résultats, on peut émettre l'hypothèse que les contraintes liées au phonogenre sont finalement plus importantes que les contraintes liées au type d'événement, et que le commentaire sportif serait bien un phonogenre, indépendamment des sports commentés. Les études globales de rythme ne nous donnent toutefois que peu d'éléments sur l'effet des contraintes d'événement sur la parole des locuteurs. La prochaine étape de ce travail consistera donc à fournir une description métrique du commentaire sportif, afin d'essayer de faire émerger des patrons rythmiques corrélés aux différents types d'événements et à leur rythme propre.

Bibliographie

- Audrit, S., Pršir, T., Auchlin, A. & Goldman, J.-P. (2012). Sport in the media: a contrasted study of three sport live media reports with semi-automatic Tools. *Proceedings of Speech Prosody 2012*.
- Augendre, S. & Mathon, C. (2012). Comment l'information à transmettre influence-t-elle les constructions référentielle et syntaxique du commentaire sportif en direct ? *Actes du CMLF 2012 - 3ème Congrès Mondial de Linguistique Française* (pp. 2027-2040). Paris : EDP Sciences.
- Augendre, S., Mathon, C., Boyé, G., & Kupsc, A. (2014, soumis). Influence des contraintes extra-linguistiques sur le discours : cas du commentaire sportif télévisé, *Actes du CMLF 2014 - 4ème Congrès Mondial de Linguistique Française*, Berlin.
- Boersma, P. & Weenink, D. (2001). Praat, a system for doing phonetics by computer. *Glott International* 5 (9/10): 341-345.
- Goldman, J.-P., Auchlin A., & Simon A.C. (2011). Discrimination de styles de parole par analyse prosodique semi-automatique. In Yoo, H.-Y. & E. Delais-Roussarie (éds) *Actes d'IDP 2009* (pp. 287-301). Paris, Septembre 2009.
- Lortal, G. & Mathon, C. (2008). Motion and Emotion or how to align emotional cues with game actions. *Proceedings of EMOT Workshop, LREC, Marrakech*.
- Pršir, T., Goldman, J.-P., & Auchlin, A. (2013). Variation prosodique situationnelle : étude sur corpus de huit phonogenres en français. In Mertens, P. & A.C. Simon (éds), *Proceedings of the Prosody-Discourse Interface Conference 2013 (IDP-2013)* (pp. 107-112). Leuven, September 11-13, 2013.