



ACOUSTICS 2012

Intonation des phrases interrogatives et affirmatives en langue Berbère

H. Teffahi and R. Halimouche

USTHB electronics en computer Faculty, BP 32 Bab Ezzouar El Alia, 16111 Algiers, Algeria
hteffahi@gmail.com

This work is at the frontier between multimedia information retrieval and automatic speech processing. During the last years, a new task emerged in speech processing: the rich transcription of an audio document. An important *meta-data* for rich transcription is the information on sentence type (i.e. sentence of interrogative or affirmative type). The study on the prosodic differences between these two types of sentences in Berber language, the detection and classification of sentence type in Berber language is the main subject of this preliminary work. We've realized a system for segmentation and automatic detection of sentence type based on both prosodic, in Berber language, a language where all studies until now on prosodic system are still preliminary. We've carried a study on the prosodic differences between interrogative and affirmative sentences. A classification motor has been built.

1 Introduction

La prosodie concerne l'ensemble des éléments dynamiques de la chaîne parlée, tels que les variations de hauteur, d'intensité (ou d'énergie) et de durée, qui déterminent la mélodie, les tons, les pauses, les accents, le rythme, le débit...etc. Il s'agit d'un phénomène complexe, relativement difficile à étudier dans la mesure où ses manifestations (i.e. ; patrons accentuels et contours intonatifs) sont sujettes à de nombreuses variations. Les schémas prosodiques diffèrent ainsi selon les langues, les dialectes, les registres linguistiques, la structure du discours, la syntaxe des énoncés, les mots constituant ces énoncés et la structure phonétique des unités lexicales, mais aussi selon le genre, l'âge, l'origine sociale voire l'état émotionnel du locuteur [1].

Chaque langue possède un ensemble de contours intonatifs qui lui est propre. Ces contours sont liés aux types de phrases et à l'expressivité. A l'intérieur d'un même type de phrase, l'utilisation d'une intonation différente engendre des interprétations différentes. En effet, l'intonation est une des particularités prosodiques qui, pour une même phrase, donnent des sens (valeurs) divers.

Pour les langues occidentales non tonales (dont le français ou l'anglais sont des exemples) il a été validé que la prosodie de la phrase véhicule des informations extralinguistiques, comme les émotions ou l'état du locuteur ou bien comme la nature de la phrase (affirmative, interrogative ou exclamative) [2, 3]. Cependant, dans le cas des langues tonales (comme le Berbère), le contour mélodique de l'intonation est complexe parce qu'il est composé de macro-variations correspondant à l'intonation de la phrase et de micro-variations correspondant aux tons lexicaux appliqués sur chacune des syllabes des mots mono (ou bi) syllabiques. C'est pourquoi, appliquer sur ces langues directement les méthodes d'analyse validées pour les langues non tonales semble ne pas permettre de détecter ou de classifier la nature des phrases avec suffisamment de fiabilité car les micro-variations tonales semblent brouiller l'information extra-lexicale de la phrase. Dans le cas de la langue berbère d'ailleurs, pour différencier les phrases interrogatives des autres, l'emploi de mots spécifiques dit « classifieurs interrogatifs » est pratiquement systématique.

Jusqu'à ce jour, très peu d'études ont analysé la phonologie de la langue kabyle en profondeur. Nous pouvons citer quelques travaux récents portant sur les tons lexicaux et sur la prosodie de la phrase [4, 5, 6]. La question alors posée peut être résumée ainsi : existe-t-il,

pour le Berbère, langue à tons dont la prosodie est complexe, des informations extralinguistiques caractérisant le type de phrases, véhiculées par la prosodie et utilisées pendant les actes de dialogue pour la classification de ces types de phrases ? La réponse, outre le fait qu'elle nous permettra d'approfondir nos connaissances de la langue, si elle est positive, nous permettra d'envisager la réalisation de classifieurs automatiques indépendants des moteurs de reconnaissance.

2 Analyse du contour intonatif

2.1 Méthodologie et préparation du corpus

Partant de ces constats, nous souhaitons préciser les différences prosodiques entre les phrases interrogatives et les phrases affirmatives en langue Berbère. Pour cela, nous avons construit un corpus spécifique constitué des paires de phrases qui se composent chacune d'une phrase question et d'une phrase non-question. Dans notre étude, nous avons porté une attention toute particulière au naturel des phrases, pour les phrases non-questions comme pour les phrases questions. Pour cela, nous les avons extraites de corpus reproduisant des situations de la vie courante. Quand il y a des particules dans la phrase interrogative, nous avons, dans la mesure du possible, gardé aussi les mêmes mots, ou bien nous avons utilisé des mots à la prononciation peu différente afin que ces deux phrases soient les plus ressemblantes que possible au niveau de la prononciation. De cette façon, nous éliminons ainsi tous les phénomènes de co-articulation qui pourraient interférer avec notre analyse prosodique. Le fait de choisir les mêmes tons nous permet d'éliminer l'influence des tons des syllabes sur l'intonation générale de la phrase. Toutes ces phrases ont été intégrées dans des dialogues significatifs, afin que leur prononciation soit la plus naturelle possible. Chaque dialogue est répété cinq fois par six locuteurs (4 hommes et 2 femmes) originaires de la Kabylie. Après sélection, notre corpus contient en tout, 336 phrases. Les paires de phrases utilisées sont présentées dans le tableau 1.

2.2 Résultats d'analyse

Pour chaque enregistrement nous analysons le contour de la fréquence fondamentale F0 avec le logiciel Praat (exemple présenté figure 1).

Tableau 1 : Les paires de phrases affirmatives (a) et interrogatives (i) du corpus.

1.i	da cu dsa ?	<i>quel jour sommes nous ?</i>
1.a	asa d lkmi	<i>aujourd'hui c'est le jeudi</i>
2.i	atct agrum ?	<i>tu manges du pain ?</i>
2.a	ad tcag agrum	<i>je mange du pain</i>
3.i	mnhu i d kci ?	<i>qui es tu ?</i>
3.a	nki d lhusin	<i>je suis Hocine</i>
4.i	achal id yusan ?	<i>combien sont venus ?</i>
4.a	u sand atas	<i>beaucoup sont venus</i>
5.i	milmi ardyugal ?	<i>il revient quand ?</i>
5.a	adyugal azka	<i>il revient demain</i>
6.i	iwacu it atlat ?	<i>pourquoi t'as tardé ?</i>
6.a	idul wbrid	<i>la route est longue</i>
7.i	da cu itsbit ?	<i>qu'as-tu ramené ?</i>
7.a	abigd irdan	<i>j'ai ramené du blé</i>
8.i	achal wigi ?	<i>combien ceux là ?</i>
8.a	wigi sidrimn	<i>avec de l'argent</i>
9.i	idul wbrid ?	<i>le chemin est long ?</i>
9.a	abrid idul	<i>le chemin est long</i>
10.i	mnhu gkhm ?	<i>qui est à la maison ?</i>
10.a	akhm ytcu	<i>la maison est pleine</i>
11.i	usand inabgawan ?	<i>les invités sont venus ?</i>
11.a	An am usand	<i>oui ils sont venus</i>
12.i	anwa wigi ?	<i>qui sont ceux là ?</i>
12.a	wigi di mazigan	<i>ceux là sont des amazigh</i>

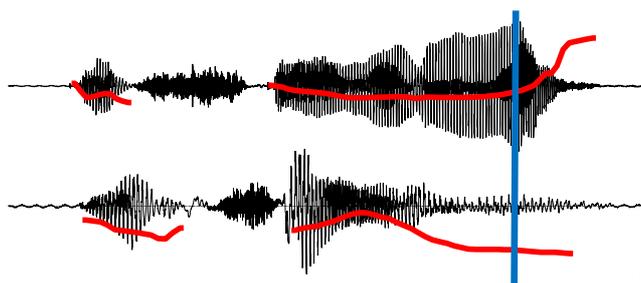


Figure 1 : Deux phrases à nombre de syllabes et tons identiques. En rouge le contour de F0.

Figure du dessus : phrase interrogative (*atct agrum ?*),
 Figure du dessous : phrase affirmative (*ad tcag agrum*).

En étudiant chaque paire de phrases présentées dans le tableau 1, nous remarquons que l'essentiel des différences d'intonation se situe à la fin de la phrase (zone située figure 1 après la barre verticale) : le contour de la dernière syllabe ou de la deuxième moitié de celle-ci semble être croissant pour les phrases interrogatives. Une étude statistique, présentée dans le tableau 2, confirme cette tendance : 70% des cas de phrases de type interrogative possèdent un contour de F0 croissant à la fin de phrase et 97% des phrases de type affirmative possèdent un contour décroissant. Nous retrouvons là une tendance bien connue pour les langues non tonales comme le français [7].

Tableau 2 : Nombre (et pourcentage) des contours de F0 de la dernière moitié de la dernière syllabe.

	Interrogative	Affirmative
Pente montante	255 (70%)	09 (3%)
Pente descendante	101 (30%)	327 (97%)

L'étude statistique, présentée dans la figure 2, donne le taux de détection correcte pour chaque locuteur suivant les contours de F0 de la dernière moitié de la dernière syllabe. Nous pouvons affirmer que les paramètres prosodiques de la phrase Berbère transportent des informations extralinguistiques qui peuvent permettre de discriminer le type de phrase. Comme pour les langues non tonales, ces informations sont essentiellement codées par le fait que l'intonation monte ou non en fin de phrase.

Au niveau production, cette première étape de l'étude a permis de caractériser la prosodie des phrases simples de la langue Berbère (dialogue), en éliminant l'influence des tons: les différences entre questions et affirmations sont essentiellement une différence de pente de F0 (croissante ou décroissante) en fin de la phrase (deuxième moitié de la dernière syllabe).

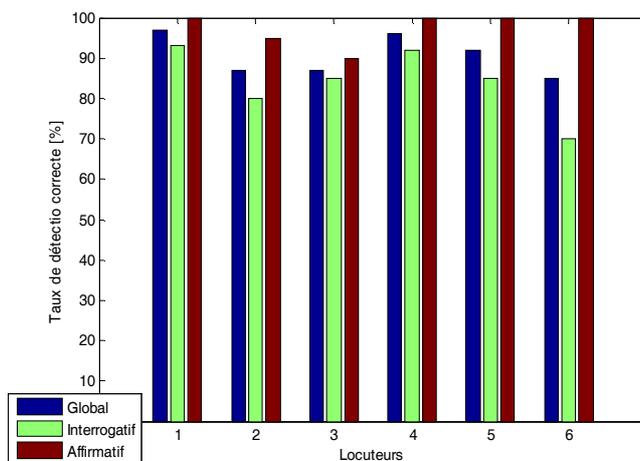


Figure 2 : Taux de détection correcte pour chaque locuteur suivant les contours de F0 de la dernière moitié de la dernière syllabe.

3 Classification Automatique

3.1 Principe

Pour évaluer automatiquement le type des phrases à des fins de classification en interrogatives et affirmatives, nous présentons dans ce paragraphe une méthode utilisant le contour prosodique. Le paramètre mesuré et analysé prend en compte l'évolution de l'intonation pendant l'énoncé de la phrase : augmentation des valeurs de F0 en fin de phrase.

Pour chaque enregistrement nous avons extrait le contour de la fréquence fondamentale F0 avec une méthode de détection de F0 basée sur la méthode cepstrale. Pour classifier nos phrases, nous avons adopté le formalisme des méthodes à Vecteur de Support Machine (SVM) comme outil de décision discriminant. Le vecteur caractéristique, qui représente l'information pertinente, est la fréquence fondamentale.

Les Machines à Vecteurs de Support ou Séparateur à Vaste Marge (SVM) sont des techniques discriminantes dans la théorie de l'apprentissage statistique.

Elles ont été proposées en 1995 par V. Vapnik [8]. Elles permettent d'aborder plusieurs problèmes divers et variés comme la régression, la classification, la fusion etc...

Nous avons appliqué dans notre travail un SVM mono-classe sous Matlab. Sachant que les SVM nécessitent une base de données d'apprentissage (2/3 de corpus) et une autre pour le test (1/3 de corpus) avec des vecteurs caractéristiques de même taille. Dans notre cas le corpus utilisé est constitué avec des phrases de différentes tailles ce qui donne des vecteurs caractéristiques différents, donc il est quasiment impossible de les utiliser dans un moteur de classification à base de SVM.

Pour s'affranchir de ce problème, nous avons opté pour une méthode de limitation des vecteurs caractéristiques (même taille) avec un taux maximum d'identification. Puisque l'information caractérisant le type se trouve à la fin de la phrase, le balayage commence de la fin du signal arrivant jusqu'au début du signal le plus court ; cela en calculant à chaque longueur du vecteur caractéristique le taux d'identification.

La taille du vecteur caractéristique est définie par le nombre de valeur de F0 qui donne le taux d'identification maximum. Pour notre corpus, la taille du signal le plus court contient 75 valeurs de F0 donc nous obtenons 75 valeurs de taux d'identification.

La procédure est illustrée par le schéma de la figure 3.

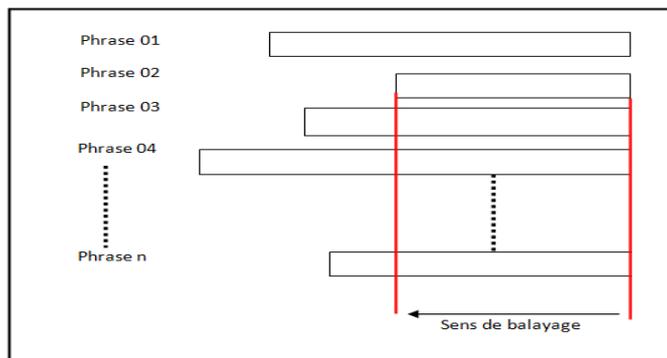


Figure 3 : Procédure de détermination de la taille du vecteur caractéristique.

Nos expériences, sur la base de données précédente, ont montrées que le taux d'identification atteint la valeur maximum pour un vecteur caractéristique de 32 valeurs de (F_0) (maximum de l'information significative). Au-delà de cette valeur le taux de classification se dégrade d'une manière significative.

Nous avons remarqué pendant nos analyses que le corpus contient du bruit au début et à la fin de chaque phrase due à l'enregistrement (environnement bruité). Pour diminuer l'effet de ce bruit nous avons utilisé une méthode qui consiste à détecter les frontières des phrases et des silences. Cette méthode est basée essentiellement sur l'estimation de l'énergie. Un silence ou bruit est détecté si l'énergie des 30 échantillons du début et de la fin du signal est inférieure à 10% de l'énergie maximum de la phrase. Cette opération permet de mieux estimer le contour intonatif.

3.2 Résultats obtenus

La figure 4, donne le taux de classification correcte pour chaque locuteur suivant les contours de F0. Nous remarquons que les phrases de type interrogatif sont les mieux identifiées pour les locuteurs (1, 4, 5 et 6) par rapport aux phrases de type affirmatif. Par contre le locuteur (3) représente un taux plus élevé pour les phrases affirmatives. La figure 3 montre que le locuteur 3 présente le taux le plus élevé de bonne identification global. Ceci est dû essentiellement à un bon positionnement de ces articulateurs (notons, que les locuteurs ne sont pas issus de la même région de Kabylie). Le taux d'identification globale est de 86%. Dans l'ensemble, les phrases interrogatives sont identifiées avec un taux de 90%, pour un taux de 81% pour les phrases affirmatives.

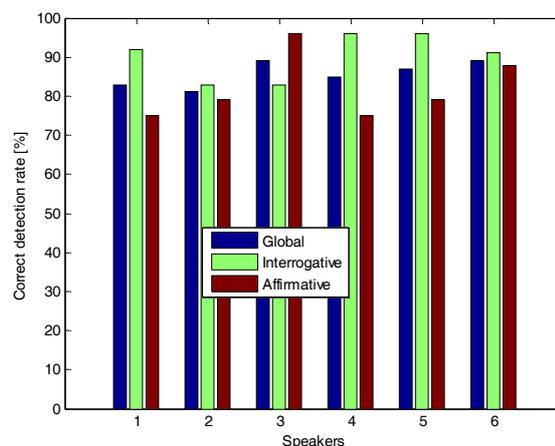


Figure 4 : Taux de détection correcte pour chaque locuteur par un classifieur SVM.

4 Conclusion

Ces résultats préliminaires, pour la langue Berbère, montrent que les différences entre questions et affirmations, sont essentiellement une différence de variation de F0 (croissante ou décroissante) en fin de la phrase (deuxième moitié de la dernière syllabe).

Ainsi, nous pouvons confirmer que l'allure du contour intonatif basé sur la fréquence fondamentale F0 est largement suffisante pour l'identification de phrases de type interrogatif. La nouveauté apportée dans ce travail réside dans l'introduction d'une méthode automatique de détection qui utilise les SVM mono-classe avec un balayage spécifique respectant les propriétés de la prosodie. Nous avons aussi montré qu'un prétraitement adéquat permet d'améliorer les résultats d'identification.

References

- [1] M. Vu, E. Castelli, A. Boucher, L. Besacier “Classification de parole en Question et Non-Question par arbre de décision” *SFC 05, 12èmes Rencontres de la Société Francophone de Classification* - Montréal, (2005).
- [2] D.Hirst, A. Di Cristo “*Intonation Systems. A Survey of 20 Languages*” Cambridge University, Press. Verlag,
- [3] M. Rossi “L’intonation, le système du français :description et modélisation” Ed Ophrys, (1999).
- [4] S. Chaker “ données exploratoires en prosodie berbère : l’accent kabyle” , *parues dans les Comptes rendus du GLECS*, 31, 27-54, (1995).
- [5] S . Chaker “ Eléments de prosodie berbère : quelques données exploratoires” *Etudes et documents berbère*, 8, 5-25 , (1991).
- [6] S . Chalah “Le rôle de l’intonation en syntaxe et en sémantique : étude de cas portant sur l’opposition d’état du nom kabyle” , *Cahiers de l’ISL*, 22, 47-62, (2007).
- [7] E. Shriberg, R. Bates, P. Taylor, A. Stolcke, D. Jurafsky, K. Ries, N. Cocarro, R. Martin, M. Meteer, C. Van Ess-Dykema “Can Prosody Aid the Automatic Classification of Dialog Acts in Conversational Speech?” *Language and Speech* 41, 439-487, (1998).
- [8] J. Platt, “Fast Training of Support Vector Machines Using Sequential Minimal Optimization”, in *Advances in Kernel Methods - Support Vector Learning*. eds. B. Schoelkopf, C. Burges, and A. Smola, MIT Press, (1998).