

Vaissière, J, 2010, «Le français, langue à frontières par excellence», dans *FRONTIÈRES, DU LINGUISTIQUE AU SÉMIOTIQUE*, ouvrage collectif initié par Nelly Andrieux-Reix, dirigé et édité par Dominique Delomier en collaboration avec Mary-Annick Morel, éditeurs Lambert Lucas.

## Preversion

**Jacqueline Vaissière**

**Laboratoire de Phonétique et de Phonologie, UMR/CNRS/PARIS3/7018**

**Université Paris Sorbonne Nouvelle**

### Plan

1. Les indices acoustiques de la segmentation en phonèmes et en syllabes.....	2
1a. Segmentation en phonèmes.....	2
1b. Segmentation en syllabes.....	4
2. Les divers indices de la segmentation en mots et les groupes de sens.....	5
2a) Phénomènes liés à la tension des articulateurs au début des mots lexicaux.....	6
2b) Phénomènes d’allongement.....	10
2c) Phénomènes au niveau du mot, harmonie vocalique.....	10
2d) Les contraintes phonotactiques, lexicales, syntaxiques, sémantiques et pragmatiques.....	11
3. Les indices de frontières supérieures au mot.....	11
3a) Les groupes de sens et les deux parties de la phrase.....	11
3b) Résumé des indices prosodiques.....	15
4. Discussions et conclusion.....	16
Remerciements.....	18
Références.....	18

### 1. Introduction

Tenir une conversation, comprendre un message verbal, acquérir une langue implique une *segmentation* du continuum sonore. Les mouvements des organes de la parole génèrent un flux continu mais le signal est décodé par l’auditeur comme une suite discontinue de *syntagmes*, de *mots* ou de *syllabes*. Il existe des indices acoustiques de frontières de ces unités (que nous détaillerons), mais leur présence dans la parole de tous les jours n’a pas un caractère obligatoire, loin s’en faut.

Depuis longtemps déjà, les linguistes, phonéticiens et psycholinguistiques se sont penchés sur le mystère de la *segmentation* du continuum vocal. Les résultats issus des expérimentations contrôlées ont mis à jour la multiplicité des indices susceptibles de jouer un rôle dans le processus de segmentation auditif. Le poids respectif des indices acoustiques contenus dans le signal, d’une part, et des indices issus d’autres niveaux, linguistiques, contextuels ou probabilistes, d’autre part, dans le processus de segmentation et de compréhension du message reste une inconnue dans les situations de communication. Les problèmes principaux sont que la quantité et l’évidence des indices acoustiques varient largement dans les situations réelles de communication, que les différents indices interagissent au niveau perceptif, et que les facteurs pragmatiques sont difficiles à estimer.

Les indices acoustiques porteurs d'information prosodique sont contenus par les *écarts* et *déviations* dans la réalisation des voyelles et des consonnes qui sont non prévisibles à partir de la seule nature des phonèmes en séquence. Les variations impliquent tous les niveaux de la production de la parole, du niveau *sous-glottique* au niveau *supra-glottique* en passant par le niveau *glottique*. Le niveau *sous-glottique* concerne la gestion de l'expiration de l'air nécessaire à la production de la parole, les pauses respiratoires et les différences d'*effort articulatoire* (mesurables par les variations de pression sous-glottique) accordé par le locuteur pour prononcer telle ou telle partie de son discours. Cet effort peut augmenter, par exemple, sur un élément du message à focaliser. Le niveau *glottique* implique la génération de la source vocale, à la fois les modulations contrôlées de la fréquence du fondamental –Fo- (Fo est le nombre de vibrations des plis vocaux par seconde), la régularité des cycles (ou mélodicité, selon l'expression de Fónagy, 2003) et les écarts de qualité de voix par rapport à la voix usuelle du locuteur. Le *niveau supra-glottique* concerne les organes articulatoires, c'est-à-dire la langue, les lèvres, le voile du palais et la mâchoire. Les déviations se traduisent pour les voyelles par des écarts dans les valeurs de formants par rapport aux valeurs attendues, et donc un timbre différent. En ce qui concerne les consonnes, la déviation peut être un degré de fermeture plus fort ou plus faible conduisant respectivement à une fortition ou à une lénition. D'autres indices, temporels, dérivent du timing de la réalisation des gestes respiratoires, phonatoires et articulatoires. Toutes ces variations résultent en une multitude de détails différents dans la réalisation d'un phonème donné, généralement visibles sur un spectrogramme (un spectrogramme représente la répartition de l'énergie dans l'échelle des fréquences en fonction du temps). Ces détails phonétiques (qui ont attiré du reste l'attention des chercheurs ces dernières années) sont disponibles pour participer au processus de segmentation. Notre article ici se restreint aux seules variations allophoniques qui renseignent sur les présences des frontières, et qui ont donc un rôle de démarcation. Pour des raisons de place, le nombre de références est réduit.

Notre communication résume les indices acoustiques de frontières en français, en insistant sur les unités inférieures, c'est-à-dire les phonèmes, les syllabes (partie1), et les mots (partie2). Nous rappelons également (mais succinctement) les indices acoustiques, en général moins subtils, qui participent à des regroupements supérieurs au mot (partie3). Cet article s'adresse à un public non phonéticien. Notre connaissance est en partie basée sur nos propres expériences de segmentation de la parole en mots en reconnaissance automatique de la parole et en lecture de spectrogrammes.

## **1. Les indices acoustiques de la segmentation en phonèmes et en syllabes**

### **1a. Segmentation en phonèmes**

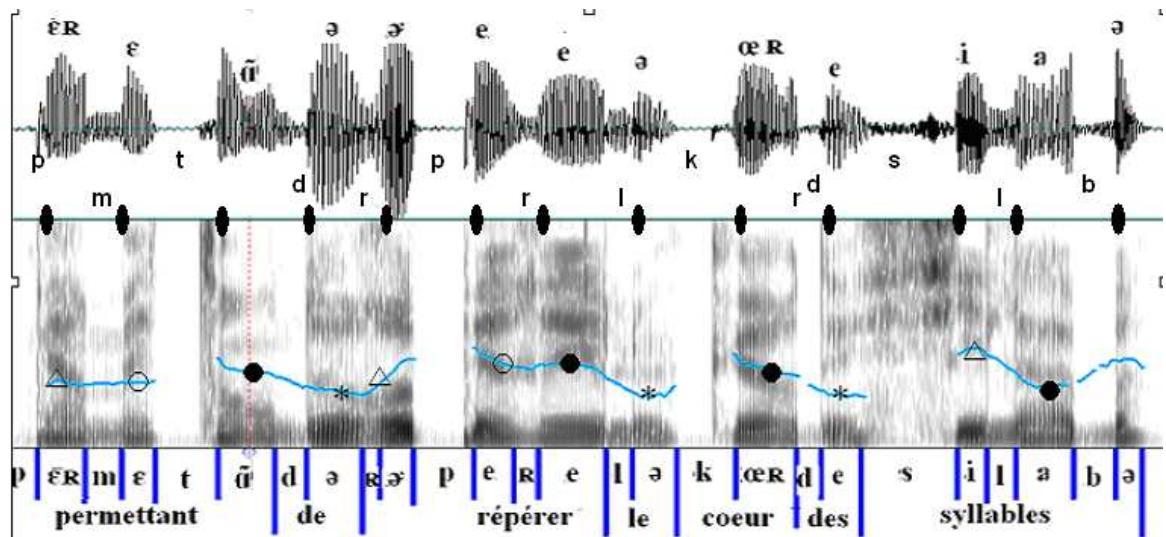


Figure 1: Extrait de phrase: «... permettant de repérer le cœur des syllables ». En haut, le signal, au milieu le spectrogramme (différents niveaux de gris) avec superposition de la courbe de  $F_0$  (en bleu); en bas: la segmentation proposée par l'auteur. Les triangles, les ronds noirs et les ronds creux sur la courbe de  $F_0$  correspondent respectivement aux premières syllabes, aux syllabes finales et aux syllabes intermédiaires des mots lexicaux. « \* » correspond aux mots outils et les ovaies pleins marquent les débuts des segments vocaliques.

Il est relativement facile de repérer la présence d'une voyelle ou d'une consonne sous-jacente sur le signal ou le spectrogramme qui lui correspond. En effet, chaque maximum local d'énergie sur le signal de parole correspond (en général) à l'ouverture du conduit vocal pour la production d'une *voyelle* (voir les treize maxima locaux d'énergie sur la figure 1). Une voyelle correspond à un seul segment vocalique, mais un segment vocalique épenthétique peut apparaître entre deux consonnes, comme dans « ours blanc » prononcé [urs\*blã], ou « avril » prononcé [av\*ril]. Il est bien évident que la segmentation cognitive en phonèmes dépend en partie de la langue. Les images acoustiques du mot « baille » français ou du mot « buy » anglais évoquent une séquence de trois phonèmes chez les auditeurs français et de deux seulement chez l'auditeur anglais. En français, l'absence de diphtongues ou d'opposition phonologique entre voyelles tendues et relâchées simplifie grandement le processus de segmentation : les mouvements des formants durant les voyelles sont exclusivement dus dans notre langue maternelle au lieu d'articulation des phonèmes environnants. Les creux entre les maxima correspondent à la présence d'un ou de plusieurs segments consonantiques. Une consonne correspond à un seul segment ou plus : les occlusives sourdes correspondent en général à deux segments, une zone de silence et un zone de bruits (voir par exemple, les différentes réalisations des consonnes sourdes /p/, /t/ et /k/ à la figure 1).

Si la présence d'une voyelle ou d'une consonne est relativement facile à repérer sur un spectrogramme, déterminer leurs frontières exactes peut être plus ou moins difficile selon la nature des phonèmes en séquence. La réalisation de chaque trait distinctif correspondant à un phonème s'étend généralement sur plusieurs segments acoustiques, affaiblissant l'évidence des frontières. On remarquera sur la figure 1 la variation du second formant durant la consonne /R/ dans les contextes symétriques [ərə] et [ere]: la valeur plus ou moins élevée de F2 dans ce cas précis traduit respectivement une articulation plus ou moins antérieure et l'image acoustique de [r] tend à se rapprocher de celles des voyelles environnantes, affaiblissant l'évidence des frontières. Malgré les phénomènes de coarticulation, des discontinuités acoustiques correspondent généralement à des changements de mode dans l'articulation entre deux sons successifs (Fant, 1967: 105 et suivantes), dans le passage d'une consonne à une voyelle et vice versa. Ces discontinuités permettent de localiser la très grande majorité des frontières des segments acoustiques. Certaines frontières entre segments apparaissent avec plus d'évidence que d'autres. Par exemple, les silences correspondant à la tenue des occlusives sourdes (telles que /p/, /t/ et /k/ à la figure 1) sont particulièrement faciles à repérer. De façon générale, la frontière est plus nette entre l'attaque et le noyau qu'entre le noyau et la coda, surtout quand la coda est une sonante. Par exemple, la consonne /R/ en coda (voir « permettant » et « cœur » à la figure 1) est impossible à séparer du noyau de la syllabe: la voyelle est uvularisée à cause de son contact avec /R/.

Les discontinuités du signal permettent de repérer automatiquement environ 80% des frontières, du moins dans des textes lus et oralisés.

### 1b. Segmentation en syllabes

La connaissance des tendances générales de la langue, des contraintes phonotactiques, des contraintes rythmiques et des détails fins peut aider le processus de décision sur la localisation des frontières de syllabes.

Les contraintes *phonotactiques* apportent leur contribution au processus de segmentation en syllabes. Le nombre de structures syllabiques possibles est limité dans une langue. Certaines séquences de consonnes  $C^1C^2$ , comme « rt » ne peuvent constituer l'attaque d'une syllabe, en français. Soit l'auditeur insérera inconsciemment une frontière de syllabes entre les deux consonnes ( $C^1C^2 = C^1-C^2$ ) (« partir », « opté » = « -par-tir- », « -op-te- »), ou soit un e muet ( $C^1C^2 = C^1əC^2$ ) (« retour », « petit » = « -rtur- », « -pti- »). Les phénomènes d'enchaînement et de resyllabification (par exemple : « Madame est mon amie » prononcé [-ma-da-mɛ-mə-na-mi-]) et l'absence du phénomène d'ambisyllabité augmentent l'impression du français comme une langue à syllabes ouvertes. Une approche développée pour le français, la SOSH (*Syllable Onset Segmentation*

*Heuristic*), propose de considérer les débuts de syllabes comme points d'alignement efficaces dans le processus de segmentation (Content & al, 2001).

L'application de contraintes *temporelles* permet de raffiner les résultats. On remarquera sur la figure 1 la régularité des intervalles entre les débuts des voyelles (marqués par des ovales). Les syllabes se succèdent à environ toutes les 160 ms ou plus, dans la parole à débit modéré. Deux maximums très proches ne peuvent donc pas (en général) appartenir à deux syllabes différentes. Cette contrainte permet de regrouper deux maximums locaux en un seul segment vocalique (comme les voyelles nasales et les voyelles suivies de /r/ en coda à la figure 1). Ce type de contraintes temporelles est inapplicable dans le cas de débits très rapides.

Des *détails fins* sur le signal peuvent également aider à la décision. La comparaison montre que les consonnes initiales, comparées aux consonnes en coda, sont, toutes choses égales par ailleurs, plus « contoïdales » (c'est à dire moins « sonantes »), plus longues, moins nasales (si elles sont nasales), moins co-articulées avec le phonème précédent, avec une courbe d'intensité plus montante durant la consonne (si visible) que dans les autres positions. Chaque indice se révèle efficace en parole de laboratoire pour distinguer sur un spectrogramme entre les mêmes suites de phonèmes ne différant que par la position des frontières de syllabes. Lehiste (1960) a étudié en profondeur les indices fins qui permettent de distinguer entre « a nice man - an ice man »; « grade A - gray day »; « two lips – tulips »; « a name- an aim » et bon nombre de ses observations sont également valables pour le français. Ces indices disparaissent souvent en parole rapide.

## 2. Les divers indices de la segmentation en mots et les groupes de sens

Les différences allophoniques dues aux frontières de mots (et de syllabes) ont pleinement montré leur rôle décisif dans les changements phonétiques. Au XIX<sup>ième</sup> siècle, les spécialistes de phonétique historique distinguaient déjà entre positions forte ou faible des syllabes dans le mot, et positions forte ou faible des phonèmes dans la syllabe. Les positions fortes dans le mot sont l'initiale et sous l'accent lexical. Les phonèmes en position forte dans la syllabe sont à l'attaque et la coda est en position faible. Par exemple, dans le passage du latin au français, seules les syllabes en position forte dans le mot, c'est-à-dire à l'initiale de mot ou sous l'accent ont résisté au temps dans les mots les plus courants, les autres ont en général disparu (**MUS**culum > moule; **CLAR**itate > clarté); les consonnes en coda de syllabe et en finale de mot ont eu tendance à disparaître (**HOS**pite > hôte). Les variations synchroniques reflètent aujourd'hui les mêmes tendances (pour une revue, voir Vaissière, 1996). Nous avons proposé d'ajouter à la liste des traits distinctifs de chaque phonème un indice quasi continu de «force » de chaque phonème, selon sa position dans la syllabe et dans le mot, et la fréquence du mot. Cet indice recouvre l'ensemble des différences, glottiques et

supraglottiques et leurs interactions, entre les réalisations des phonèmes en positions forte et faible (Vaissière, 1986 : 537).

Les instrumentations modernes ont permis d'identifier les conditions articulatoires, acoustiques, ou aérodynamiques qui distinguent les allophones positionnels et qui permettent d'expliquer, en partie, leur comportement de fortition ou de lénition au cours du temps. Elles ont permis de nuancer fortement les affirmations anciennes (et parfois récentes !) sur le manque de frontières dans la parole continue, en français, affirmations qui étaient basées sur la seule impression perceptive.

Le tableau suivant résume des indices susceptibles de contribuer au marquage des frontières de mots. Il n'existe pas, à notre connaissance, d'études testant l'ensemble des indices et leurs impacts respectifs sur le plan de la perception.

Nous nous proposons d'illustrer un nombre réduit de points: 1) les conséquences de la « tension » initiale (trait [+strong]) sur les plans supraglottique et glottique qui marque le début des mots, 2) les phénomènes d'allongement qui en marquent la fin, et 3) l'harmonie vocalique qui renforce le mot en tant qu'unité acoustique, en français.

Tendances prosodiques	Niveau supra-glottique	Renforcement initial du mot lexical Allongement de l'attaque du mot Contact accru des articulateurs Allongement de la rime finale du mot lexical Harmonie vocalique dans le mot
	Niveau glottique	Saut initial de Fo en début de mot Glottalisation de la voyelle initiale de mot Cible basse en fin des mots grammaticaux
Montée de continuation finale ou pic		
Dévoisement initial		
Autres contraintes		phonotactiques
		Syntaxiques, sémantiques, pragmatiques

*Tableau 1: Quelques indices de frontières gauches et droites de mot en français*

## 2a) Phénomènes liés à la tension des articulateurs au début des mots lexicaux

La potentialité du début de mot à être porteur en français d'un accent d'insistance (ou didactique, Lucci, 1980) a été attestée depuis fort longtemps. Cet « accent » correspond à un renforcement de la « tension » initiale qui existe dans la parole neutre.

L'utilisation du terme « tension » n'est pas facile à justifier, car il n'existe pas d'études électromyographiques qui comparent l'activité de tous les articulateurs de la parole dans les différentes positions du mot. La « tension » initiale se traduit, entre autres, par un contact accru des

articulateurs supra-glottiques (a1), par le dévoisement des consonnes voisées initiales (a2), par un saut du fondamental en début de mot (a3).

### 2a1) Variations dans le degré de contact des articulateurs supra-glottiques

La figure suivante (extraite d'un article de Marguerite Durand, 1929) illustre six palatogrammes. Un palatogramme est une photographie de l'empreinte de la langue sur le palais dur durant la réalisation du son. Il renseigne sur le degré de contact entre la langue et le palais dur durant l'articulation des sons antérieurs et leur mode articuloire (degré de fermeture plus ou moins grand). Les palatogrammes A et B représentent l'empreinte de la consonne /g/ à l'initiale absolue (après une pause) des mots « gamme » et « garçon »: la fermeture est complète, comme attendu (le segment est occlusif). F présente une occurrence finale de /g/: dans ce cas, la consonne n'est pas lénifiée. Les trois autres palatogrammes illustrent des cas de lénition (la fermeture n'est pas complète). C correspond à l'empreinte pour la consonne /g/ à l'intérieur du mot « agape », E et F à la consonne initiale lénifiée de /g/ dans le mot « garçon » inséré dans un groupe de sens (Le lecteur intéressé pourra consulter les travaux de Straka, 1979, et l'ensemble des travaux, plus récents, de Fougeron, 1998, sur le français). Un style plus relâché et un débit plus rapide induisent plus de phénomènes de lénition qu'un style soigné et un débit lent.

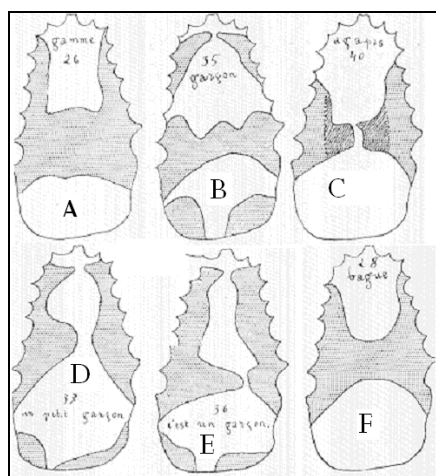


Figure 2: Palatogrammes de M. Durand (1929). Voir le texte.

Delattre (1940) a noté qu'il était possible d'observer dans certains cas que la consonne d'enchaînement (« petite amie » = [pti-ta-mi]) a un comportement de coda plutôt que d'attaque (« petit amis » = [pti-ta-mi]). Des différences subtiles apparaissent sur les spectrogrammes en lecture soignée, surtout par comparaison de paires.

### 2b2) Variations dans le degré de voisement/assourdissement des consonnes

La tension initiale et la « relaxation » non initiale concernent aussi le niveau glottique. Les plis vocaux vibrent moins facilement en début de mot en français et les consonnes voisées ont

tendance à se dévoiser à l'initiale. A l'intervocalique, les obstruents sourdes se voisent (et s'ouvrent, comme vu précédemment). Par exemple, « d'accord » a tendance à être prononcé avec un dévoisement du /d/ initial et un voisement de /k/ en français spontané (Basset, 1999). Les variations synchroniques rappellent naturellement les règles qui ont présidé bon nombre de changements dans le passage du latin au français (lat. « ripa » devenu « rive » en français, où /p/ a été spirantisé et voisé).

### **2b3) Variation dans la réalisation du saut initial de fréquence du fondamental**

La tension de début de mot se traduit également en français par un saut éventuel du fondamental. Ce saut que nous avons dénommé « montée initiale », (*initial rise*, noté « [H » sur nos figures) ou mieux « saut initial » se situe entre le mot grammatical précédent (noté « b ») et le mot lexical suivant. Il peut être réalisé même en cas d'absence d'accent d'insistance et en lecture (Vaissière, 1975). Le lecteur remarquera sur la figure 1 le saut positif du fondamental entre le cœur de la dernière syllabe du mot grammatical et la première voyelle du mot lexical – dans les suites « de REpérer », « le COEUr », « des SYllabes ». L'amplitude de ce saut initial par rapport à la plage de variations de  $F_0$  varie beaucoup selon les locuteurs (Vaissière 1974) et elle tend à diminuer du début vers la fin de la phrase. Le saut initial n'a pas un caractère obligatoire. Contrairement à ce qui pourrait être attendu pour un mot lexical de plus de deux syllabes (les mots longs ont tendance à avoir un plus grand nombre de sauts initiaux que les mots courts), la première syllabe de la proposition « permettant » n'est marquée par aucun signe d'accentuation initiale (voir Vaissière et Michaud, 2006). Il peut s'estomper sur un mot, parfois pour mettre ce mot en retrait, ou parfois pour mettre en valeur un mot environnant (l'absence de saut initial au début du mot « permettant » conduit à mettre en relief le mot suivant, ici « repérer »), ou encore par anticipation d'une montée de continuation (Vaissière, 1975), lorsque le mot commence par une voyelle, ou lorsque le débit est rapide. Sa fréquence d'utilisation dépend aussi du style (Lucci, 1980, Fónagy, 1980).

La nature des phonèmes sous-jacents influence la position du maximum local de  $F_0$ , conséquence de la tension initiale. Le maximum local atteint, conséquence de la tension initiale, n'est pas une cible précise sur l'axe du temps. Sa position exacte varie avec le caractère voisé ou non voisé de la consonne initiale (figure 3). Si le mot lexical commence par une *consonne sourde*, la « montée initiale tend à prendre la forme d'un saut de  $F_0$  entre la voyelle initiale du mot et la voyelle précédente ; le maximum local de  $F_0$  apparaît de façon précoce dans la voyelle, souvent à son début (observez le saut du fondamental entre le mot grammatical et le début des mots « cœur » et « syllabes » à la figure 1 et dans « compatibilité » à la figure 3). La tension des plis vocaux nécessaire pour la réalisation du trait sourd de la consonne favorise la réalisation du saut du fondamental d'origine prosodique. Si la consonne initiale du mot est *voisée*, la cible est plus

tardive : des plis vocaux trop tendus ne peuvent pas vibrer. La courbe de  $F_0$  durant la première syllabe prend l'allure d'un glissando montant (voir « repérer » à la figure 1 et « l'incompatibilité » à la figure 2), qui peut être confondu en reconnaissance automatique de la parole avec une montée de continuation (Vaissière, 1982). Si le mot lexical commence par une *voyelle*, sans liaison et sans enchaînement, un coup de glotte (non obligatoire) au début de la voyelle permet d'éviter le report du saut initial sur la seconde syllabe, ou encore sa suppression. La position de la valeur haute varie aussi avec la longueur du mot: si le mot est long, le maximum tend à se déplacer vers la droite.

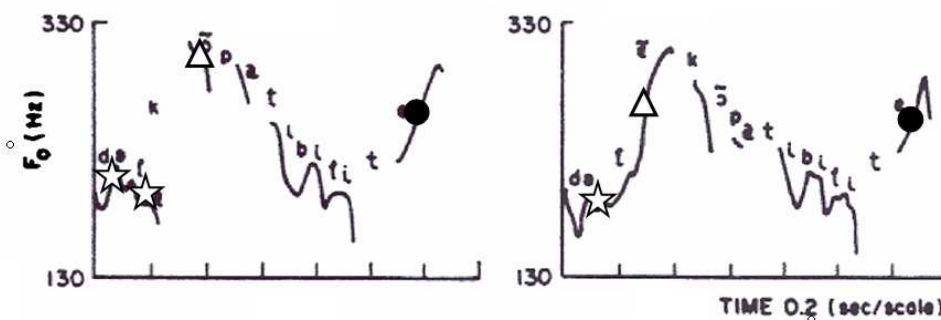


Figure 3: Courbe de  $F_0$  de la séquence « de la compatibilité » (à gauche) et de « de l'incompatibilité » (à droite), illustrant la basse de  $F_0$  durant les mots grammaticaux et la réalisation différente du saut initial après le(s) mot(s) outil(s) en fonction des phonèmes sous-jacents (d'après Vaissière, 1975).

L'amplitude de la tension initiale (et l'amplitude du saut initial) tend à diminuer du début à la fin de la phrase ; elle augmente après une frontière importante entre deux parties de l'énoncé, et elle est sensible au poids sémantico-pragmatique du mot dans l'énoncé.

Les mots outils tendent à être prononcés avec une cible basse en leur fin. La cible basse (notée « b » dans notre notation) se situe souvent très exactement à la frontière entre le mot grammatical et le début du mot lexical (Vaissière, 1975). Selon le contexte,  $F_0$  est soit plat soit descendant durant les mots grammaticaux, ou montant en tout début de phrase (la montée est due à la superposition de la montée initiale de phrase qui s'étale sur plusieurs syllabes). Le lecteur est invité à remarquer la valeur basse du fondamental (sur la ligne de base) sur les mots grammaticaux « de », « le » et « des » à la figure 1 et « de » et « la » à la figure 3, et « de », « la », « à » à la figure 4, et les valeurs plus élevées (c'est-à-dire au dessus de la ligne de base) dans « à » à la figure 4 « à Nîmes », ou « en » (« en bête ») et « et » (« et m'embête »), à la figure 5, où ils sont insérés dans la descente finale de  $F_0$  en fin de phrase.

## 2b) Phénomènes d'allongement

### 2b1) Allongement de la consonne initiale

La première consonne des mots tend à être légèrement allongée. Une longue durée d'une consonne observée sur un spectrogramme ne permet pas cependant d'affirmer que la consonne est initiale de mot, car une consonne en coda peut également être allongée (Vaissière, 1982). La syllabe est souvent considérée comme l'unité prosodique minimale. Ce choix n'est pas entièrement justifié en français: pour une durée donnée d'une syllabe CV, un allongement relatif plus important de C ou de V suggère respectivement une syllabe initiale de mot ou une syllabe finale (Vaissière, 1977).

L'allongement initial n'est pas un phénomène propre au français (voir une revue dans Vaissière, 1995).

### 2b2) Allongement de la rime finale

L'allongement évoque souvent la notion de fin chez l'auditeur (sur des stimuli non parole, Fraisse, 1974) et l'allongement final en parole a été attesté dans de nombreuses langues. L'allongement de la rime finale des groupes de sens est une caractéristique essentielle du français (Delattre, 1962 ; Delattre, 1939). L'allongement relatif de la dernière syllabe des mots a été vérifié statistiquement sur des grands corpus de parole journaliste en français (Adda-Decker & al, 2008). Un allongement final de mot ou de syntagme suffit à la distinction de 'bords durs ' versus 'bordures ', 'Jean, Pierre et Jacques' versus ' Jean-Pierre et Jacques ', « l'eau tarit » versus 'l'otarie ' (voir, entre autres, Banel et al, 1994, Rietveld, 1980). Les mots outils sont relativement plus courts, du moins en général (voir cependant la longueur relativement longue de la première proposition « de » à la figure 1). L'allongement en fin de groupes de sens n'est pas obligatoire, les frontières pouvant être marquées par le seul contour de Fo ; inversement, des allongements suffisent à marquer les frontières droites des mots ou des groupes de sens dans des parties de parole avec un contour plat.

### 2c) Phénomènes au niveau du mot, harmonie vocalique

Des phénomènes d'harmonie entre les phonèmes en séquence renforcent la cohésion du mot. Le locuteur français a tendance à ouvrir toutes les voyelles *moyennes* dans « phonologue » (prononcé [fɔnɔlɔg], dans « aimait » (prononcé [ɛmɛ] et dans « Elle était belle » ([ɛlɛtɛbɛl) et à fermer toutes les voyelles moyennes dans « phonologie » (prononcé [fonolozi]), dans « aimer » (prononcé [eme]), et dans « c'était l'été » [setelete] ou à nasaliser tous les phonèmes de « amener » (d'où la confusion avec « emmener »). Les enfants français prononcent « petit » [piti] et « surtout » [surtut]. L'harmonie de voyelle à voyelle pour les voyelles *non moyennes* est conditionnée par la résistance de la consonne intermédiaire à laisser la langue et/ou les lèvres se coarticuler avec les

voyelles environnantes. Par exemple, les consonnes coronales sont plus résistantes que les consonnes labiales et vélares.

Il existe aussi une tendance générale observée dans plusieurs langues à la réduction des durées moyennes des phonèmes en fonction de la longueur du mot.

## **2d) Les contraintes phonotactiques, lexicales, syntaxiques, sémantiques et pragmatiques**

Les *contraintes phonotactiques* limitent les combinaisons de phonèmes qui peuvent constituer un début légal de mots. La présence de groupes de consonnes illégitimes suggère la présence d'une frontière sous-jacente de mot. L'utilisation de contraintes phonotactiques sur les débuts de mots est affaiblie en français du fait de la possible disparition d'un e muet: [rt], [mn] est un début possible de mots comme « retour » et « mener », un phénomène que nous avons mentionné plus haut.

Il existe des contraintes de plus haut niveau. Au delà des évidences phonétiques, l'auditeur cherche à regrouper des phonèmes ou des syllabes en un mot légal dans la langue (*contraintes lexicales*), à reconstituer une suite de mots qui soit syntaxiquement correcte (*contraintes syntaxiques*), qui fasse sens (*contraintes sémantiques*) et dont le sens soit compatible avec les mots précédents reconnus et la situation de communication (*contraintes pragmatiques*). Il existe un très grand nombre de recherches très intéressantes dans ces domaines, qui dépassent le cadre de cet article.

## **3. Les indices de frontières supérieures au mot**

Il existe des indices moins subtils, toujours audibles, pour instancier des structures supérieures au mot.

### **3a) Les groupes de sens et les deux parties de la phrase**

Les anciens, qui ont essentiellement étudié la prosodie dans la parole spontanée, à partir de l'écoute (Grammont, 1933) ont toujours insisté sur l'importance du sens (et non de la syntaxe) dans le découpage de l'énoncé français oral. Les locuteurs français tendent à segmenter le continuum en petits paquets de mots qui font sens (en anglais: « sense group »). Il existe certes un grand flou sur le nom de cette unité juste supérieure au mot. Les unités de la taille du groupe de sens ont été parfois définies sur une base *perceptive*, comme un ensemble de syllabes se terminant par une syllabe « accentuée », d'où la dénomination de groupe ou syntagme accentuel, groupe accentué, groupe rythmique, groupe intensif, unité accentuelle (en anglais: « stress group »); sur une base *acoustique*, comme un groupe rythmique se terminant par une syllabe allongée (Delattre, 1939), ou encore par un mouvement mélodique (Hirst et Di Cristo, 1998); sur une base *syntaxique*, comme un syntagme grammatical, etc. Au-delà de cette foison d'appellations, du reste compatibles, des consensus se sont faits autour de la composition du groupe de sens ou unité équivalente: 1) il doit

contenir au moins un mot plein (lexical), précédé du ou des mots outils qui en dépendent; 2) il peut contenir plusieurs mots pleins liés par le sens, comme un nom et un adjectif antéposé ou postposé; 3) les locuteurs ont tendance à équilibrer la longueur des groupes de sens successifs: ils diront plus volontiers « Pierre part / en vacances » (plutôt que « Pierre/part en vacances »; 4) enfin et surtout, une même phrase peut être divisée à l'oral en un plus ou moins grand nombre de groupes de sens, en fonction du contexte, de la vitesse d'élocution, du libre arbitre du locuteur, etc. (Grammont, 1933).

Lorsque deux mots pleins sont unis dans un même groupe de sens, Delattre (1939) avait déjà noté que la « désaccentuation » (ce sont ses termes) des mots « graphiques » à l'intérieur du groupe n'était pas complète. En fait, il ne s'agit pas uniquement de fait de « désaccentuation » de la dernière syllabe des mots internes au groupe de sens, mais d'un ensemble de changements qui conduisent à conférer à l'ensemble du groupe de sens un profil qui se rapproche plus ou moins de celui d'un seul mot.

Selon la tradition, les groupes de sens se regroupent au sein de l'énoncé pour former deux parties dans la phrase déclarative: la partie dite montante (correspondant en général mais non obligatoirement au thème), et la partie dite descendante (correspondant au rhème : Coustenoble et Armstrong, 1934; Karcevski, 1931). Cette division était essentiellement définie, comme le groupe de sens pour certains, sur des bases *pragmatico-sémantiques*: « La partie montante est celle qui annonce quelque chose et suscite une attente; la partie descendante est celle qui satisfait l'attente et conclut la phrase (Grammont, 1933: 152-3). Les courbes mélodiques dessinées de façon impressionniste par les auditeurs représentaient la première partie avec une mélodie de plus en plus haute (avec des décrochements locaux montants à la fin de chaque groupe de sens), un sommet mélodique à la fin de la première partie et une descente progressive le long de la seconde partie.

Les indices acoustiques qui permettent le repérage de la frontière entre partie montante et descendante de l'énoncé ou *frontière majeure* de l'énoncé furent clairement explicités par des études expérimentales subséquentes, car les progrès techniques ont permis de visualiser enfin les courbes de Fo. La tendance générale et régulière de la Fo à décliner régulièrement (ou ligne de déclinaison) du début jusqu'à la fin de la phrase (représentée sur certaines de nos figures par des lignes en pointillés) avaient échappé perceptivement aux anciens : on n'entend pas ce à quoi on s'attend physiologiquement (Vaissière, 1995).

On a pu constater aussi que la frontière entre les deux parties de la phrase correspond rarement à une Fo la plus élevée de la phrase, du moins dans les phrases longues. Une pause plus longue peut faire l'affaire (voir exemple plus bas). Pour Delattre (1966), c'est une ampleur plus grande du mouvement de Fo du groupe de sens qui marque la différence entre la continuation

majeure et les continuations mineures dans l'énoncé. Nous verrons par la suite que ces différences de longueur de pauses ou d'amplitude du mouvement montant de Fo ne sont que des options parmi d'autres pour marquer différents degrés de frontière à l'intérieur de la phrase. Delattre a aussi illustré le fait qu'une phrase pouvait contenir, si on s'en tient au plan acoustique, plusieurs continuations « majeures » de même amplitude, s'écartant de la définition de la « continuation majeure » comme jointure entre thème et rhème, mais rejoignant en cela la définition de « groupe intonatif » (ou GI) dans les descriptions plus modernes.

Donnons deux exemples, le premier concernant une phrase longue, et le second des phrases courtes.

La figure 4 illustre une schématisation d'une courbe de Fo possible du célèbre holorime de Marc Monnier (1829-1885), divisé ici en 6 unités, les 5 premières étant terminées par une (forte) montée de continuation et la dernière par une descente marquant la finalité, et se terminant par la valeur de Fo la plus basse de la phrase, un fait bien connu. La ligne de base de Fo et le plateau (les termes de Maeda, 1976) ont été indiqués sur la figure par des lignes (parallèles ici, mais souvent asymptotiques, Maeda 1976) en pointillés. La frontière perçue comme la plus importante se situe après la troisième unité : « Gal, amant de la reine, alla tour magnanime /// galamment de l'arène, à la tour Magne, à Nîmes ». Une chute rapide du fondamental (signalée par des flèches sur la figure 4) marque la fin des mots lexicaux « amant », et de « galamment ». Dans la suite « alla tour magnanime », prononcée sans pause, le mot « alla » a le comportement prosodique d'un mot grammatical, et le mot « tour » est marqué par un pic initial de Fo. Les durées relatives confirment, dans ce cas, les informations délivrées par la mélodie: toutes choses égales par ailleurs, les rimes finales aux articulations majeures sont les plus longues, suivies des allongements en fin de mots (« amant », « tour », « galamment »), et les mots outils sont raccourcis par rapport aux premières syllabes des mots pleins. La frontière majeure n'est marquée, ici, ni par un glissando de plus grande ampleur qu'à la fin des unités 1 et 2, comme l'aurait prédit Delattre (1966), ni par une cible de Fo la plus élevée de la phrase, selon la description des anciens, mais par la présence d'une plus grande pause et par un rehaussement de la ligne de base du fondamental à la reprise de parole après la pause. Un simple rehaussement de la ligne de base et/ou une durée plus longue de la rime avant la frontière majeure auraient également suffi (sur des références sur la complexité de la perception de l'intonation, Vaissière, 2005).

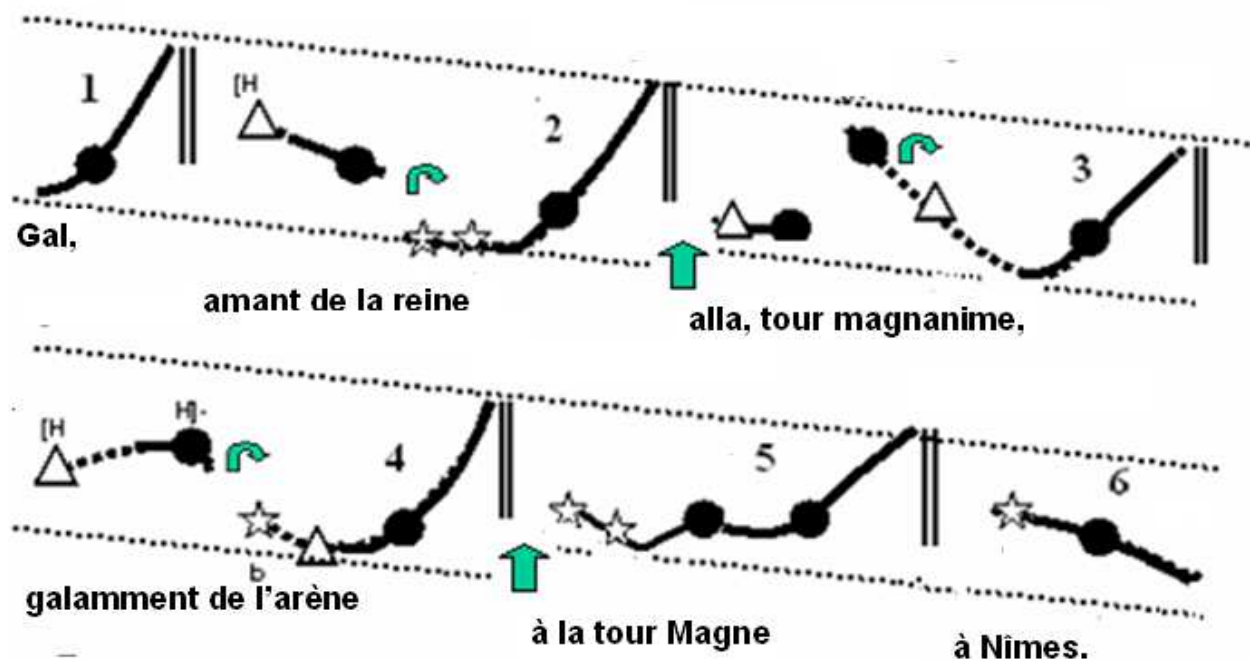


Figure 4: Schématisation de la courbe de Fo de l'holorime de Monnier : « Gal, amant de la reine, alla, tour magnanime, galamment de l'Arène à la tour Magne à Nîmes ».

La figure 5 illustre la désambiguïation de trois courtes phrases par les pauses, des contrastes de pente sur la dernière syllabe des mots lexicaux et des degrés d'allongement différents. Comme attendu, la pente ascendante de Fo sur la deuxième ou cinquième syllabe dans chaque phrase marque l'articulation majeure. Pour ce locuteur, et pour ces phrases courtes, l'articulation majeure correspond au Fo le plus haut de chaque phrase. Comme attendu, les mots outils tendent à se situer sur la ligne de base (sauf s'ils sont inclus dans la descente finale de la phrase, comme mentionné auparavant). Le saut de Fo entre des mots outils et le mot plein suivant, à la frontière gauche de mot est modéré chez ce locuteur. Les signes «::» «:» indiquent deux degrés différents d'allongement, établis par comparaison à l'intérieur d'une même phrase: les fins de groupe de sens tendent à être plus allongées («::») que les fins de mot et l'allongement en fin de mots est facultatif.

Il existe des indices prosodiques de plus haut niveau qui intègrent plusieurs énoncés en un seul paragraphe, et ces indices ne sont pas propres au français (Lehiste, 1979 sur l'anglais ; pour une revue, Vaissière, 1983). Alors que les rapports partiellement congruents entre prosodie et syntaxe ont déjà fait l'objet de très nombreuses dans plusieurs langues depuis les années 70 et que l'essentiel a sans doute été déjà découvert, les études sur les rapports entre prosodie et discours ne font que commencer et elles ouvrent des perspectives passionnantes.

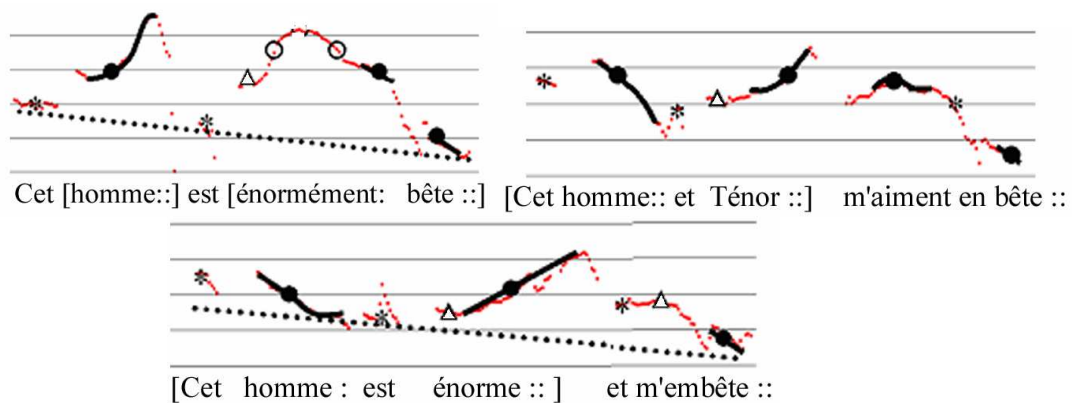


Figure 5: *Fo* de trois phrases désambiguïsées par la prosodie (extrait de Vaissière, 2007).

### 3b) Résumé des indices prosodiques

#### 3b1) Pauses

La pause *respiratoire* répond à une nécessité physiologique et le locuteur en profite pour respirer entre les articulations majeures de son discours, entre les phrases et au milieu des phrases très longues, quand il lui est nécessaire de reprendre son souffle. Une pause non respiratoire peut être insérée à d'autres endroits, par exemple, pour marquer d'autres frontières, pour respecter les signes de ponctuation du texte en lecture, et il existe des pauses d'insistance, placées avant les mots à mettre en valeur et des pauses d'hésitation, qui n'ont pas de fonction démarcative (Duez, 1982). Toutes choses égales par ailleurs, la longueur de la pause tend à être modulée selon la profondeur de la frontière à marquer. Le nombre de pauses dépend également d'autres facteurs, comme le style de discours, le débit et l'état émotif du locuteur (voir également, Grosjean & al, 1979).

#### 3b2) Autres procédés mélodiques

Il existe d'autres procédés prosodiques, désormais bien connus : les contrastes de pics de *Fo* ou d'amplitude de montée finale de *Fo*, un rehaussement de la ligne de base et des allongements de durée de la rime finale de constituants. Tous ces procédés semblent physiologiquement motivés et se retrouvent donc dans plusieurs langues (Vaissière, 1995). Une plus grande ampleur d'une des montées de continuation (b) et/ou un rehaussement de la ligne de base (c) permettent d'augmenter le degré de frontière perçue et de contraster les différentes structures prosodiques (pour le français, Vaissière, 1975).

Bien que le français donne subjectivement l'impression d'être une langue où toutes les syllabes tendraient à être perçues comme étant d'une durée égale, c'est justement le contraste entre les durées qui est porteur d'informations (Wenk et Wioland, 1982). De façon générale, plus le mot est indépendant du mot suivant, et plus l'allongement final est important (et plus le contour de *Fo*

est montant, et/ou haut). La durée seule des rimes et la durée de la consonne initiale suffisent à contraster des phrases du type « Ton tonton Thon tond tonton », ou « tonton Thon tond ton thon Thon », en synthèse, même si le contour de Fo est plat.

#### 4. Discussions et conclusion

Le marquage fort des frontières d'un niveau quelconque se fait toujours au détriment du marquage des autres niveaux. Par exemple, le locuteur peut « syllaber », en détachant bien chaque syllabe, ou parler « mot à mot », en marquant bien chaque mot. Il peut également délivrer son message avec de grandes envolées lyriques. L'allongement final, le renforcement initial, et la tendance à utiliser un ton haut ou montant pour marquer la continuation ont été attestés dans un grand nombre de langues (mais pas dans toutes), mais ils dominent en français: le bébé français babille avec des mélodies plus fréquemment montantes, avec un allongement final plus net que les bébés anglais ou japonais (de Boysson-Bardies, 2004) et devenu adulte, il tendra à réaliser spontanément des montées de continuation et un allongement finals en fin de groupe de sens dans toutes les nouvelles langues qu'il apprend.

Selon les points de comparaison avec les autres langues, le français a été classé différemment. Il a été considéré comme une *langue sans accent* (Rossi, 1980) du point de vue fonctionnel, comme une *langue à accent fixe final* du point de vue étymologique, à *accent primaire final* et à *accent secondaire initial*, comme une *langue à double accentuation* du point de vue phonétique ou comme une *langue à accent « probabilitaire »* (Fónagy, 1980). D'autres ont préféré le classer comme une *langue à accent de groupe* (Delattre, 1939) du point de vue *sémantique* ou à *accent final de syntagme* du point de vue *syntactique*. Enfin, dans certaines typologies, le français est catégorisé comme une *langue à intonation* (comme l'anglais et l'allemand) par opposition aux langues à tons, car la courbe de Fo dépend essentiellement de l'énoncé, et non de caractéristiques mélodiques différences de chaque mot. Du point de vue rythme, il est généralement classé comme une *langue à rythme iambique*, ou encore à *double rythme*, c'est-à-dire avec des accentuations en début et en fin des constituants (pour une revue Vaissière, 1991). Parmi les langues à intonation, nous avons proposé de considérer le français comme « *une langue à frontières* », où domine l'aspect démarcatif de la prosodie, par opposition aux *langues dites à accent*, comme l'anglais (Vaissière, 1991), où domine son aspect culminatif. En français, la congruence entre accent lexical (qui se situe sur la dernière syllabe pleine des mots lexicaux, c'est-à-dire leur frontière droite) et les marques intonatives de continuation (sur les frontières droites des constituants en anglais) rend le marquage des frontières droites plus régulières et évidentes que dans les autres langues. Il est alors possible de manifester une grande variété de degrés de frontières droites, en jouant sur les pauses, la durée, Fo (et l'intensité) (Vaissière 1986 : 535). Par ailleurs, la possibilité de jouer sur des degrés

différents de renforcements initiaux, à la frontière gauche des constituants, engendre une double structuration prosodique, l'une fortement liée à la syntaxe, et l'autre à la structure informationnelle de l'énoncé. Elle permet, toutes choses égales par ailleurs, de renforcer/diminuer le degré de frontières entre les constituants marqué par les frontières droites, d'une part, et d'assigner à la phrase une structure pragmatico-sémantique qui ne soit pas strictement équivalente à sa structuration prosodico-syntaxique, d'autre part (les mots importants du point de vue de la pragmatique voient leur début renforcé, surtout s'ils sont longs).

Les déviations par rapport aux attentes sont porteuses d'informations (Fónagy). Le poids relatif porté sur la frontière droite ou gauche des unités dépend d'un grand nombre de facteurs, comme du style et aussi de la région: les politiciens et les enseignants (Lucci, 1980), les gens de l'Est et du Nord appuient plus fortement sur la frontière gauche, alors que les gens du Sud rendent plus évidente la dernière syllabe qu'ils font « chanter », ce qui modifie naturellement les attentes. Cette double accentuation (et double structuration, selon nous) semble être, pour l'instant, une autre caractéristique du français. Nous recommandons la lecture de Di , à paraître, 1999 et de Rossi, 1999 pour des vues complémentaires.

La fonction démarcative, qui nous concerne ici, n'est pas, loin s'en faut, la seule fonction de la prosodie. La prosodie permet également d'exprimer des nuances de sens, d'attitudes et d'émotions (voir par exemple, l'article de Fónagy, 2003, sur les diverses fonctions de la prosodie), qui sont également marqués par des écarts par rapport aux attentes et où la phonation joue un grand rôle.

Les premiers essais d'utilisation de certains indices prosodiques en reconnaissance automatique de la parole datent des années 70-80 (voir Vaissière, 1988, pour une revue). Wayne Lea (1980) a été le premier à proposer pour l'anglais une estimation automatique de la distance perceptive de l'intervalle inter-accentuel, qui est basée sur la profondeur de l'échancrure de la courbe de  $F_0$  entre deux syllabes lexicalement accentuées (le fameux « *fall-rise pattern* »): plus la vallée était importante, et plus la séparation syntactico-sémantique entre les deux mots était perçue comme importante. Le programme conçu pour l'anglo-américain ne cherchait aucunement à repérer les frontières elles-mêmes. Le but du programme *Proséidon* (Vaissière, 1984), que nous avons nous-même conçu pour le français, était de repérer dans un premier temps toute évidence de début et fin de mots lexicaux à partir des seuls indices mélodiques et temporels contenus dans le signal et dans un second temps de dériver une structure en hiérarchisant la force des différentes frontières repérées, le nombre de niveaux marqués étant grosso modo proportionnel à la longueur des énoncés. Les essais d'utilisation de la prosodie en reconnaissance ont été moyennement couronnés de succès, entre autres à cause du caractère non obligatoire de la présence des indices prosodiques, de leur

caractère subtil, de la variabilité interlocuteurs et de la difficulté de pondérer les évidences prosodiques face aux autres évidences (syntaxiques, par exemple).

Delattre a insisté sur le groupe de sens comme unité prosodique prégnante en français et il a affirmé que le français « manque beaucoup de moyens phonétiques pour délimiter les mots ». La comparaison avec les autres langues montre que le français est une langue où la fonction démarcative domine. Le locuteur marque en général les frontières droite des groupes de sens (par un allongement final, par un glissando montant, pour un pic de Fo, par une chute de Fo, etc.) et éventuellement gauche (par un saut initial, un renforcement de la consonne initiale, par une glottalisation de la voyelle, etc.). Si tel est son désir, ou si le contexte l'exige, il a la possibilité de marquer ou d'estomper les frontières des mots internes aux groupes de sens, et d'exprimer ainsi le degré de dépendance entre les mots en séquence avec une grande finesse. Les paramètres acoustiques, tels que la Fo, durée, et intensité qui sont communs au marquage des accents lexicaux et des frontières dans les langues à accent (comme l'anglais) ne sont pas monopolisés en français par le marquage d'accents lexicaux et le français peut être considéré comme une langue à frontières, par excellence.

### **Remerciements**

Nous remercions très sincèrement les membres de LPP, Patricia Basset, Florentina Fredet, Michaela Pernon, Martine Toda, Lucile Wallet pour une relecture. Nous remercions également Cécile Fougeron pour ses remarques pertinentes.

### **Références**

Adda-Decker, M., Gendrot, C. & Nguyen, N., 2008, Contributions du traitement automatique de la parole à l'étude des voyelles orales du français. *Revue Traitement Automatique des Langues*, vol 49, n°3, p. 1-34.

Banel, M., H., Bacri, N., 1994, On metrical patterns and lexical parsing in French, *Speech Communication*, n° 15, 115-126.

Basset, P., 1999, *Disparition des phonèmes en français spontané*, Diplôme d'Etudes Approfondi de Phonétique, Université de la Sorbonne Nouvelle, France.

Content, A., Kearns, R.K. & Frauenfelder, U. H., 2001, Boundaries versus onset in syllabic segmentation, *Journal of Memory and Language*, 45-2, 177-199.

Coustenoble, H. N., Armstrong L. E., 1934, *Studies in French intonation*, Cambridge, W. Heffner, Sons.

Boysson-Bardies, B., 2004, *Comment la parole vient aux enfants : de la naissance jusqu'à deux ans*, Odile Jacob.

Delattre, P., 1939, Accent de mot et accent de groupe, *The French Review*, n° 13, 2, p. 141-146.

Delattre, P., 1940, Le mot est-il une entité phonétique en français ? , *Français Moderne*, n° 8, 47-57.

Delattre, P., 1962, Comparing the prosodic features of English, French, German and Spanish, *International Review of Applied Linguistics*, n° 1, 3, p. 193-210.

Delattre, P., 1966, Les dix intonations de base du français, *The French Review*, n° 40, 1, p. 1-14.

Di Cristo, A., 1999, Le cadre accentuel du français : essai de modélisation : première partie, *Langues*, vol. 2, n°3 : 184-205. seconde partie : *Langues*, vol. 2, n°4 : 258-269.

Di Cristo, A., à paraître, Regards sur la prosodie du français, Première Partie : Accentuation et Phrasé prosodique.

Duez, D., 1982, Salient pauses and non salient pauses in three speech style, *Language and Speech*, n° 25, 7, 11-28.

Durand, M., 1929, Études sur les phonèmes postérieurs dans une articulation parisienne, *Revue de Phonétique*, VI, 3-4.

Fant, G., 1967, Sound, features and perception, *STL-QPSR*, n° 2-3/1967, 1-14.

Fónagy, I., 1980, L'accent français: accent probabilitaire. Dynamique d'un changement phonétique, in L'accent en français contemporain, *Studia Phonetica*, n° 15, 4, p. 123-233.

Fónagy, I., 2003, Des fonctions de l'intonation : essai de synthèse, *Flambeau*, 29, Tokyo, Université des langues étrangères, p. 1-20.

Fougeron, C., 1998, *Variations articulatoires en début de constituants prosodiques de différents niveaux en français*. Thèse de Phonétique, Paris III Sorbonne Nouvelle, France.

Fraisse, P., 1974, *Psychologie du rythme*, Presses Universitaires de France, Paris.

Grammont, G., 1933, *Traité de phonétique*, Delagrave, Paris.

Grosjean, F., Grosjean, L. & Lane, H., 1979, The patterns of silence: Performance structures in sentence production, *Cognitive Psychology*, n° 11, 58-81.

Hirst, D. & Di Cristo, A., 1998, A survey of intonation systems, dans Hirst & Di Cristo (eds). *Intonation Systems: A Survey of Twenty Languages*, Cambridge University Press, p. 1-44.

Karcevski, S., 1931, Sur la phonologie de la phrase, *Travaux du Cercle de Linguistique de Paris*. n° 4, p. 188-227.

Lea W., 1980, Prosodic Aids to Speech Recognition, dans W. Lea (ed.), *Trends in Speech Recognition*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, p. 166-205.

Lehiste, I., 1960, An acoustic-phonetic study of internal open juncture. *Phonetica*, n° 5, p. 5-54.

Lehiste, I., 1979. Perception of sentence and paragraph boundaries, dans *Frontiers of speech communication*. B. Lindblom and S. Ohman (eds.), New York: Academic Press, 191-201.

Lucci, V., 1980, L'accent didactique. L'accent en français contemporain, *Studia Phonetica*, n° 15, p. 107-121.

- Maeda, S., 1976, *A characterization of American English intonation*, Ph. D., Massachusetts Institute of Technology. [<http://dspace.mit.edu/handle/1721.1/29189>]
- Rietveld, A. C. M., 1980, Word boundaries in the French language, *Language and Speech*, n° 23, 3, p. 289-96.
- Rossi, M., 1980, Le français, langue sans accent ?, *Studia Phonetica*, n° 15, p. 85-101.
- Rossi, M., 1999, *L'intonation, le système du français. Description et modélisation*, Ophrys.
- Straka, G., 1979, *Les sons et les mots*, Klincksieck, Paris.
- Vaissière, J., 1974, On French prosody, *Quarterly Progress Report*, Research Laboratory of Electronics, Massachusetts Institute of Technology, n° 114, p. 212-223 [halshs-00363971]
- Vaissière, J., 1975, Further note on French prosody, *Quarterly Progress Report*, Research Laboratory of Electronics, Massachusetts Institute of Technology, n° 115, p. 251-261. [halshs-00363974]
- Vaissière, J., 1977, Premiers essais d'utilisation de la durée pour la reconnaissance en mots dans un système de reconnaissance, *VIIIes Journées d'Etudes sur la Parole*, p. 345-352.
- Vaissière, J., 1982, Utilisation des paramètres suprasegmentaux en reconnaissance automatique de la parole comme aide à la segmentation. In : A. D. Cristo, J.-P. Haton, M. Rossi, & J. Vaissière (Eds.), *Prosodie et reconnaissance automatique de la parole* (p. 123-139). Groupement des Acousticiens de Langue Française.
- Vaissière, J., 1984, PROSEIDON, Automatic Detection of Prosodic Cues in Continuous Speech, *XIIIes JEP*, Bruxelles, pp. 53-66
- Vaissière, J., 1983, Language-independent prosodic features, dans A. Cutler & D. R. Ladd (Eds.), *Prosody: Models and measurements* (p. 53 -66). Heidelberg, Springer-Verlag.
- Vaissière, J. 1986, Variance and Invariance at the Word Level. dans J. S. Perkell & D. H. Klatt (Ed.), *Invariance and Variability in Speech Process*, Invited Contribution, (p. 534-539). Lawrence Erlbaum Associates, Publishers. [halshs-00368898]
- Vaissière, J., 1988, The use of prosodic parameters in automatic speech recognition, dans *Recent advances in speech understanding and dialog system*, (p. 71-99), NATO ASI Series, Eds H. Niemann & al., Springer Verlag. [halshs-00363982]
- Vaissière, J., 1991. Rhythm, accentuation and final lengthening in French, dans *Language, Music, Speech and Brain*. (p. 108-120), J. Sundberg, L. Nord and R. Carlson, Macmillan Press. [halshs-00363980]
- Vaissière, J., 1995, Phonetic explanations for cross-linguistic similarities, *Phonetica*, n° 52, 123-130.
- Vaissière, J., 1996, From Latin to Modern French : on diachronic changes and synchronic variations, *Arbeitsbericht, Insitut für Phonetik and digitale Sprachverarbeitung*, Universität Kiel, n° 31, p. 61-74. [halshs-00365006]
- Vaissière, J., 2005, Perception of intonation, dans D. B. Pisoni & R. E. Remez (Eds.), *The Handbook of Speech perception*, p. 236-263. Padstow, Cornwall: Blackwell Publishing.
- Vaissière, J., 2007, *La phonétique*, Que-Sais-Je?, PUF (seconde édition)

Vaissière, J., & A. Michaud, 2006, Prosodic constituents in French: A data-driven approach dans: *Prosody and syntax: Cross-linguistic perspectives*, I. Fónagy, Y. Kawaguchi and T. Moriguchi (Eds.), Amsterdam: J. Benjamins, p.47-63. [hal-00130794]

Wenk, B.J. & Wioland, F., 1982, Is French really syllable timed ?, *Journal of Phonetics*, n° 10, p. 193-215.