

CHAPITRE XIX

PROSODIE, SENS ET AFFECTIVITE

Depuis quelques années, l'étude des « émotions » est devenu un nouveau domaine d'investigation dans le contexte des recherches internationales. Les progrès énormes réalisés dans les domaines des neurosciences et de la cognition, les développements parallèles en psychologie, en parole et en prosodie, en musicologie, et la perspective des applications commerciales qui ne cessent de se diversifier, expliquent ce foisonnement extraordinaire des recherches (pour une revue des études voir Caelen-Haumont, 2005).

Pour notre part, l'étude de la subjectivité en prosodie entamée depuis l'origine de mes travaux me permet de concevoir le domaine de l'affectivité, non pas comme le champ exclusif des « émotions », mais comme un champ diversifié, celui des états affectifs. En effet, au côté du champ des émotions et des attitudes bien recensé par la littérature, un état nous semble important en prosodie, sinon prioritaire, à savoir celui de l'émotion que nous appelons *l'émotion ordinaire* ou *émotion-racine*.

Ce chapitre a pour objectif d'aborder le champ de l'affectivité dans le domaine de la parole, qu'il s'agisse d'émotion ordinaire, d'émotions ou d'attitudes, en replaçant parmi les études antérieures et actuelles, nos propres résultats et ceux qui leur sont associés.

1. LES DOMAINES DE LA SUBJECTIVITE DANS LA PAROLE

Dans le domaine de la parole affective, et en dehors de la pathologie, il semble bien en effet que l'on puisse distinguer trois états subjectifs différents, à savoir ceux propres à « *l'émotion-racine* », aux « *émotions typées* », et aux « *attitudes* ».

Selon Caelen-Haumont et Bel 2000, *l'émotion-racine* est « l'émotion ordinaire », c'est-à-dire l'émotion à la source de l'individu, qui en dehors d'émotions spécifiques telles la joie, colère etc., instancie dans l'instant, souvent à son insu, la personne avec ses croyances, ses motivations, ses impulsions, ses inclinations et ses répulsions. Cet état subjectif, toujours présent, se manifeste spontanément dans le discours ou la conversation pour peu que la personne s'investisse dans sa performance. Il est en relation avec le contenu des échanges, se mobilisant et se démobilisant tour à tour, pour se manifester à nouveau de manière tout aussi transitoire. C'est le sentiment intérieur qui s'exprime, et s'il a le champ libre, il s'exprime en relief, au premier plan. Cet état est référencé au singulier, ce qui montre son caractère général lié à la subjectivité perçue comme l'élément fondamental de la personne. Il semble que ce soit essentiellement dans la parole que se manifeste cette émotion-racine.

Les *émotions typées* caractérisent des états bien connus de tout individu, expérimentés soit par le truchement de la parole, ou en dehors d'elle. Une émotion violente ou causée par surprise peut en effet bloquer momentanément l'expression verbale, de même qu'un interdit social par exemple peut réprimer son expression affective. On recense les émotions *d'extraversion*, telles que par exemple l'excitation, l'admiration, la joie, etc., et les émotions *d'introversion*, telles que la tristesse, l'ennui, la timidité, etc..

Les attitudes constituent un autre état affectif. La première tentative de définition du terme a été proposée par Couper-Kuhlen 1986 qui oppose ainsi l'émotion comme état du locuteur, à l'attitude, comme type de comportement.

Mais ceci laisse de côté les attitudes cognitives répertoriées par les psychologues sociaux qui sont fonction d'opinions, de croyances ou de connaissances. Ces dernières

sont appelées des *attitudes propositionnelles*, dans la mesure où elles s'expriment dans et relativement à des propositions : par exemple dans la phrase « je suis impressionnée par tes propos », « impressionnée » rend compte d'une émotion particulière telle que l'admiration exprimée vis-à-vis d'un énoncé.

Dans le domaine de l'intonation, Wichmann 2000 propose de réserver le terme d'*intonation expressive* aux énoncés porteurs d'indices prosodiques de l'émotion et aux énoncés exprimant une attitude propositionnelle, issue de croyances, connaissances ou opinions, et le terme d'*intonation attitudinale* aux énoncés qui dénotent un comportement de la part du locuteur, intentionnel et/ou inféré par son auditeur dans une situation donnée.

Or ces trois instances de la subjectivité dans l'énoncé possèdent en commun des indices prosodiques. Ainsi l'émotion-racine et les émotions d'extraversion qui manifestent toutes un investissement de la personne et son implication, se caractérisent généralement par des valeurs de F0 élevées, une variabilité des inflexions mélodiques, une forte amplitude tonale, et associés à des pauses non linguistiques, un débit plus rapide et une forte intensité. En fait le registre élevé est un indice discriminant des émotions et attitudes, attestant d'un état d'excitation et d'implication de la personne.

Inversement, les émotions d'introversion qui se caractérisent par le retrait de la personne dans l'énoncé manifestent un contour peu contrasté ou plat, un registre de F0 moyen à grave, un débit (beaucoup) plus lent, une intensité plus faible, et la présence plus marquée des pauses linguistiques.

Ce domaine connaissant à l'heure actuelle un développement considérable dans les sciences cognitives en général, il est intéressant d'en préciser les contours historiques au moins ici en ce qui concerne les approches linguistique et pragmatique. On trouvera dans Caelen-Haumont (2005), une revue détaillée de la recherche actuelle dans les sciences cognitives sur toutes ces questions.

2. L'APPROCHE LINGUISTIQUE / PRAGMATIQUE DE LA PROSODIE EMOTIONNELLE

A la fois linguiste des langues et de la parole, auteur de près de 350 écrits, Bolinger 1907-1992 s'est intéressé dès 1945 aux systèmes prosodiques de différentes langues. Dès 1972 il est l'un des premiers à lier de manière essentielle l'intonation aux aspects sémantiques et au contexte du discours, montrant que la prééminence prosodique signalant l'information nouvelle ou importante est liée à l'accentuation. Pour lui les fonctions grammaticales de la prosodie sont indirectes, l'intonation ne définissant pas les catégories grammaticales, et réciproquement. Malgré la vague conceptuelle de l'explication syntaxique qui est alors dominante, et avant le nouvel essor au niveau international des considérations sémantiques puis pragmatiques pour la prosodie, une perspective proche à cette époque, est tenue en France par Faure (1970, 1973) puis par Caelen-Haumont (1978/1981).

Concernant les émotions, Bolinger 1986 montre que la prosodie renforce les intentions sous-jacentes du locuteur, à savoir les croyances et le désir d'attention. Il distingue pour l'anglais, les accents de pouvoir qui donnent de la force à l'assertion et qui sont positionnés à la frontière droite du mot, par le positionnement d'un point culminant, et les accents d'intérêt. Une manière de réduire l'accent de pouvoir en anglais est de faire glisser l'accent vers la frontière de gauche du mot. Bolinger souligne aussi, en dehors de la question de l'accent, que la mélodie fournit pour l'expression des sentiments, les indices de l'excitation mélodie haute vs. du relâchement mélodie basse, avec la fin de la tension.

Parmi les tout premiers essais d'interprétation des indices prosodiques liées à

l'expression des émotions et attitudes, Léon 1970 associe pour le français perceptivement un registre élevé à « la gaieté, timidité, légèreté », un registre bas à « la tristesse, assurance, gravité », et l'augmentation de l'amplitude tonale à la joie, la colère, ainsi que sa réduction à la tristesse. Le contour inversement n'est pas considéré à lui tout seul comme discriminant, tandis que l'intensité renforce le sentiment exprimé.

Il semble qu'une caractéristique des nombreux travaux dans le domaine est d'avoir associé les émotions à la conjonction paramètres de F0 / types ou modalités de phrase. Ainsi pour O'Connor et Arnold 1973, les « wh-questions » avec « chute haute » sont qualifiées de « vives, affairées, attentionnées, non inamicales, vivantes, ou intéressées ». Pour Halliday 1994, les « wh-questions » avec un ton montant signalent une continuité, tandis que les « questions oui-non » avec un ton descendant sont péremptoires. Une déclarative finissant par une montée dénote « un défi, de l'agressivité, une défensive ou de l'indignation ».

Comme nous le verrons au paragraphe 2.1.3., Scherer et al. 1984b, dans le modèle « configuration » reprend cette idée de lier l'émotion à la relation contour / type de phrase. Ainsi dans l'étude des questions, le fait d'inverser les patrons standards de l'anglais chute finale pour les wh-questions, montée finale pour les questions oui-non produit une perception plus négative. Comme le souligne Wichmann 2002, ces résultats montrent que les choix marqués génèrent des attitudes, en particulier négatives. On peut aussi avancer l'idée qu'une autre source de sens affectif peut être générée dans les associations conventionnelles contour / acte de parole qui sont rompues intentionnellement ou pas. Dans ces conditions, il serait bien d'établir le patron des choix non marqués pour tous les types de phrases pour comparaison. Mais la tâche n'est pas simple car elle possède des déterminants qui ne sont pas uniquement linguistiques.

Pour le français, Fónagy et Bérard 1973, Fónagy et al. 1982c, ont étudié en détail les différents clichés mélodiques du français parisien, clichés qui sont dépendants du style de parole. Fónagy 1982a et b, spécialise ses études dans l'analyse des styles de parole par exemple lecture de contes de fée, informations journalistiques ..., des attitudes, et des émotions (Fónagy 1983), détaillant avec précision contours et cibles mélodiques.

Dans une étude pragmatique de 88 contours prosodiques de phrases anglaises se terminant par une requête « please », se réalisant la plupart du temps par une interrogation avec une finale soit montante, soit descendante, Wichmann 2002 montre que les requêtes avec finale haute sont utilisées dans une relation privée amis, collègues, alors que celles avec finale basse le sont en situation publique statut social inégal. Ces résultats sont en congruence avec le Code de Fréquence proposé par Ohala 1983 (voir plus en détail au paragraphe 2.1.2.), attribuant pour l'anglais un effet de soumission au registre élevé, et un effet de pouvoir au registre grave. De fait Wichman conclut qu'un registre élevé dans une situation informelle laisse à l'interlocuteur la possibilité d'achever l'échange, tandis que dans une situation formelle et plus impersonnelle, un contour à finale descendante est la propriété de celui qui possède l'autorité, qui la manifeste sous la forme d'une invitation qu'il n'est pas bon de refuser, ou encore d'une directive.

Concernant le français, dans la continuité du Code de Fréquence défini comme on l'a vu ci-dessus par Ohala, Piot 2001 se propose de tester sur les questions à contour montant l'hypothèse que la montée de F0 serait « plus grande lorsque le locuteur est plus désarmé, c'est-à-dire dans ce cas plus ignorant ». Utilisant les ressources de la synthèse de la parole pour manipuler plus facilement les items du test et leurs combinaisons acoustiques croisées F0, intensité, débit, deux expériences perceptives ont été menées, la première sur la question totale, pour l'évaluation 1° de « l'ignorance exprimée », 2° de « l'envie de savoir exprimée », la deuxième sur l'assertion à finale

haute pour évaluer 1° « l'ignorance d'autrui exprimée », 2° « l'envie d'informer exprimée ».

De ces deux expérimentations, il ressort tout d'abord que la hauteur du pic tonal est déterminante pour exprimer l'ignorance du locuteur vs. de l'allocutaire, et l'envie de savoir vs. d'informer. Viennent ensuite l'importance de l'amplitude et de la finale des contours pour la perception des attitudes, ainsi que l'existence d'une gradience pour la relation hauteur tonale / nouveauté. Un plus grand débit renforce la perception d'une plus grande envie de savoir, mais pas l'envie d'informer, et exerce une influence négative sur l'expression de l'ignorance de l'allocutaire. Enfin l'intensité finale a un effet sur les assertions mais pas sur les questions.

Dans la continuité des travaux de Scherer sur la covariance et la configuration (voir plus loin le paragraphe 2.1.3.), et leurs implications, Ladd et al. 1985 entreprennent trois expérimentations du même ordre. Les sujets devaient juger de l'émotion véhiculée par les énoncés où l'amplitude de F0, le type de contour, et la qualité de la voix étaient systématiquement modifiés. Une tâche de classement était proposée pour les états liés à une excitation, et une autre pour les attitudes cognitivement liées. L'objectif de la dernière expérimentation était de tester si les variations dans l'amplitude de F0 avaient des effets continus ou catégoriels sur la perception des émotions.

Les résultats ont montré d'une part que l'amplitude de F0 et la qualité de la voix avaient une forte incidence à la fois sur l'inférence que fait le sujet sur l'état d'excitation du locuteur, et sur l'inférence des attitudes cognitivement liées. Il n'a pas pu être montré que le type de contour affectait massivement le classement des attitudes, mais au contraire plutôt le classement des émotions. Comme on s'y attendait, les variations de l'amplitude de F0 ont causé des effets plus continus que catégoriels dans les jugements relatifs à l'émotion du locuteur. Une conclusion importante a été tirée, à savoir que les trois indices prosodiques, amplitude de F0, type de contour et qualité de la voix fonctionnent indépendamment les uns des autres.

2.1. L'APPROCHE PAR LES MODELES

2.1.1. LES INDICES PROSODIQUES DES DIMENSIONS DE LA PROSODIE : VALENCE, ACTIVATION, POUVOIR

Ayant constitué dans le cadre de l'allemand, une base de données constituée de dialogues spontanés recueillis dans une situation de construction de Lego, Kehrein 2002 a sélectionné sur la base de consensus perceptuels, les meilleurs candidats parmi les énoncés porteurs d'émotion. Après transcription orthographique, sous Praat, les paramètres du F0 moyen, début / fin de F0, amplitude de F0, F0 maximum / minimum, nombre de syllabes par seconde ont été extraits. L'objectif était de tester la relation entre ces divers paramètres et les dimensions de la prosodie : valence, activation, pouvoir, auxquels a été ajouté le caractère inattendu.

Il apparaît très clairement que l'information inattendue est liée à l'élévation du maximum de F0 : plus F0 est haut et plus la force du caractère inattendu est importante. Ceci confirme les résultats obtenus par Caelen-Haumont 1991, et dans l'ouvrage présent, pour le français avec un modèle des informations inattendues. Dans l'étude de Kehrein, l'émotion correspondante est celle d'un locuteur « alarmé, surpris, horrifié ».

Concernant la dimension de l'activation, c'est le tempo qui est le facteur dominant : plus le rythme est rapide, plus le locuteur est perçu comme excité, et inversement. C'est le cas lorsque les locuteurs sont perçus comme « excités, agités, incertains, anxieux ou en colère ». Pour le rythme lent, les locuteurs sont perçus comme « calmes, relaxés, satisfaits ou encore démotivés ».

Pour la dimension du pouvoir ou dominance, la force est liée à une augmentation de

la proéminence, ce qui correspond à une augmentation de l'amplitude de F0 et de l'intensité, et inversement pour la qualité opposée. Les locuteurs sont alors perçus comme « énergiques, sûrs, en colère et ennuyés », mais aussi « ravis ». Dans le cas contraire, ils sont perçus comme « peu sûrs », « perplexes, s'excusant, résignés, frustrés, déçus ».

Concernant la valence positive, les paramètres sont un étirement temporel de la syllabe, un contour avec un F0 maximum, et un contour alterné faible / fort pour l'intensité. Les locuteurs sont alors qualifiés de « ravis », et d'autres études mentionnent aussi pour une telle configuration par exemple le plaisir gastronomique !

2.1.2. LES CODES BIOLOGIQUES ET AUTRES CODES

Comme on l'a précisé ci-dessus, Ohala 1983 a proposé d'instancier un *Code biologique* de la fréquence fondamentale *Frequency Code* fondé sur le fait que les larynx de petite taille ont tendance à produire des notes plus élevées que ceux de plus grande taille. Il en découle que les fréquences élevées sonnent comme vulnérables et soumises, les fréquences basses comme protectrices et dominantes. Des expérimentations ont montré que pour le code fréquentiel, le registre fréquentiel F0 moyen et l'étendue de F0 étaient pertinents. Quand l'étendue de F0 augmente, le degré du caractère amical du message augmente, mais le degré de confiance diminue. Ceci varie cependant en fonction de la langue, et cette perception, à étendue de F0 égale, est reçue comme plus amicale et générant plus de confiance dans les stimuli de langue britannique que leurs correspondants hollandais. L'explication en est que les différences linguistiques des étendues de F0 induisent des effets non uniformes sur la perception des significations signalées par de telles variations.

Chen et al. 2002 proposent à leur tour deux autres codes, le *Code de Production* et le *Code de l'Effort*. Ce dernier est fondé sur le fait qu'un effort articulatoire plus intense a tendance à créer des réalisations phonétiques plus élaborées et plus explicites. Les codes biologiques se prêtent à deux types d'interprétations, information et affectivité, et dans le cas du Code de l'Effort, l'emphase répond à une information sur l'importance subjective que le locuteur accorde à son message, tandis qu'une émotion comme la surprise répond à celle de l'affectivité. Dans les deux cas, le locuteur dépense de l'énergie pour signaler ces fonctions. Trois indices de F0 manifestent cet effort et se révèlent pertinents pour l'interprétation des significations : la hauteur des pics de F0, le retard dans l'alignement des pics, et le registre de F0.

Concernant l'emphase, les résultats interlangues montrent une corrélation positive avec la hauteur du pic de F0, mais à hauteurs identiques, les stimuli hollandais sont perçus comme plus emphatiques que les stimuli de l'anglais. Ceci trouve sans doute son explication dans le fait que l'étendue de F0 en hollandais est plus faible qu'en anglais. Concernant le registre du pitch, une corrélation négative est cette fois trouvée entre les deux langues, le degré d'emphase croissant avec un registre du pitch plus aigu pour le hollandais, avec un registre du pitch plus grave pour l'anglais. Une explication possible est que la hauteur du pitch pourrait être utilisée pour un autre type de signification.

Concernant la surprise, dans les deux langues, on trouve une corrélation positive d'une part entre la hauteur du pic et le degré de surprise, d'autre part entre ce dernier et la production tardive du pic de F0. Dans les deux langues en effet retarder l'occurrence d'un pic de F0 est perçu avec la même fonction qu'un pic plus haut non retardé. Dans les deux cas, à hauteurs de pic identiques et à délais identiques de production du pic, les stimuli hollandais sont perçus comme générant plus de surprise. Ceci tient au fait déjà constaté de l'usage d'une étendue de F0 plus restreinte dans le cas du hollandais.

Les deux significations sont dérivées du Code de l'Effort qui stipule que les

excursions les plus larges et les hauteurs les plus aiguës correspondent aux significations associées à la dépense de l'effort, telle que la signifiante du message emphase, insistance ou l'état d'agitation du locuteur surprise. Un usage différent des indices de F0 / signification apparaît dans les deux langues, les anglais, relativement au Code de Fréquence, réservant le registre haut comme un signal de soumission et de disposition amicale vs. protection et autorité, les hollandais, relativement au Code de l'Effort, l'utilisant pour l'insistance et l'emphase dans la mesure où la hauteur de F0 reste un composant de l'étendue du pitch.

2.1.3. LES MODELES DE SCHERER

Dès le début des années 1970, Scherer monte un groupe de recherches spécialisé dans le domaine des paramètres vocaux comme indicateurs potentiels des états du locuteur tels que motivation, émotion et stress. Menés en collaboration avec un groupe dirigé par Ekman, les résultats d'expérimentations 1976 montrent que sous le stress, le niveau moyen de F0 augmente. Toutefois, les travaux ultérieurs de Scherer ont montré que les organismes répondaient de manière non uniforme aux différents types de stress.

Il se trouvait que les modèles transactionnels défendaient cette thèse en argumentant que le type d'évaluation et le potentiel de confrontation de l'individu co-déterminent le type de réponse au stress. Le cadre du travail de Scherer (1986, 1987) était ainsi dessiné : focaliser les recherches « sur les mécanismes d'évaluation de l'individu en essayant de comprendre les réponses physiologiques et motrices aux événements stressants ».

Un temps de latence existant manifestement entre l'évaluation d'un événement et l'exécution d'une réponse comportementale, ce temps pouvait être l'indice d'une activation de sous-systèmes de l'organisme. Scherer 1984a propose alors un modèle liant les composants de l'émotion aux différents sous-systèmes exerçant chacun une fonction spécifique, suggérant ainsi que les états émotionnels devaient être étudiés dans les termes d'un « component process model ». L'originalité des travaux se situent plus particulièrement dans la définition d'une série de contrôles dans l'évaluation du stimulus. Cette série de contrôles concerne, dans l'ordre, la dimension de la nouveauté, de l'agrément situation agréable ou pas, de la proximité dans la satisfaction des buts ou des besoins, du potentiel de maîtrise de soi, de la compatibilité aux normes personnelles et sociales. Scherer et al. 2003 fournissent une description différenciée des états émotionnels en fonction de ces divers contrôles, sous la forme d'une matrice de traits distinctifs, permettant de prédire les indices vocaux correspondants. Du fait que ces contrôles varient grandement d'un individu à un autre, même pour un stimulus identique, le modèle prédit que les réponses des individus au stress seront très différentes, qu'il s'agisse des sentiments, de la réponse physiologique ou de l'expression motrice. Scherer 1986 propose différents tableaux prédisant quels types de réponses sont attendus de l'organisme en fonction de ces domaines.

Par ailleurs, Scherer et al. 1984b proposent de distinguer pour l'intonation deux approches, l'approche *covariance*, et l'approche *configuration*. La première stipule que l'information sur l'émotion et l'information sur le contenu linguistique au sens strict fonctionnent de manière indépendante, ce qui entraîne la conséquence que les traitements linguistiques et paralinguistiques puissent s'effectuer en parallèle. Les indices de F0 associés à cette approche sont le F0 global et l'amplitude de F0. L'approche configuration pose que le type de contour de F0 est un élément linguistique. De ce fait on distingue entre les variations linguistiques de F0 et les variations paralinguistiques. Cela sous-entend que l'information liée au contour et à l'émotion n'est valide que si elle est traitée en interaction avec les traits linguistiques de l'énoncé.

L'indice de F0 correspondant est le type de contour. Au terme d'expérimentations perceptuelles, il apparaît que les deux approches sont pertinentes, la covariance semblant plus adéquate pour décrire la parole affectée par les facteurs biologiques, la configuration, pour la parole affectée par les conventions linguistiques et socioculturelles.

Scherer 1989 par ailleurs propose de distinguer deux types d'effets au niveau des expressions de la parole émotionnelle, les *effets push* et les *effets pull*, que nous envisageons ci-dessous au paragraphe 2.1.

Les recherches de l'équipe Scherer semblent maintenant s'orienter d'une part vers l'analyse des processus sous-jacents à l'encodage et décodage des émotions, et d'autre part grâce à l'analyse des cartes topographiques de potentiels évoqués, vers le traitement du décours temporel dans les processus cognitifs des composantes émotionnelles.

2.1.4. LE MODELE UNIFIE DE ZEI-POLLERMANN

Une autre validation concernant cette fois « l'émotion-racine » est apportée également par Zei-Pollermann 2002. En effet s'appuyant d'une part sur les conceptions de Piaget 1970 et de Prieto 1975, et d'autre part sur les dimensions reconnues généralement à la prosodie, valence, pouvoir, activation, l'auteur propose un modèle unifié cognition / émotion, argumentant que les processus cognitifs incluent nécessairement ces dimensions, puisque les connaissances sur le monde sont construites en fonction de la signification que l'individu accorde à ses intérêts personnels et ses comportements. Dans ces conditions le processus d'identification du stimulus et son évaluation inclut l'attribution de trois étiquettes en tant que trois dimensions de l'espace cognitif-affectif. Ce modèle suggère donc que les processus cognitifs impliqués dans l'interaction avec l'environnement incluent nécessairement ces dimensions, dimensions jusque-là réservées à la caractérisation des états émotionnels.

Cette conception est intéressante à plusieurs égards. Elle permet d'abord de classer les réponses de l'organisme émotionnelles vs. non-émotionnelles en fonction d'un seuil, un état non émotionnel se caractérisant par une variation habituelle de valeurs liées aux normes individuelles, l'état émotionnel étant atteint avec le dépassement du seuil pour au moins l'une des trois dimensions. Ensuite elle instaure une continuité entre les états et processus émotionnels vs. non émotionnels. Enfin elle rend compte du fait que tout état d'un organisme n'est jamais affectivement neutre.

3. INTONATION ET MELISME

3.1. ESPACE PROSODIQUE SOCIAL ET ESPACE INDIVIDUEL

Dans notre perspective (Caelen-Haumont et Bel 2000), la prosodie possède une structure composite, associant deux strates dont l'une est liée à l'expression de la structure linguistique, c'est-à-dire à l'expression de la langue en tant que convention sociale. Cette strate qui découpe l'énoncé en phrases et groupes, selon l'usage de la langue, est l'*intonation*. La seconde strate est liée à l'expression affective, et correspond au phénomène prosodique que nous dénommons « *mélisme* » (voir ci-dessous au paragraphe 3 pour une analyse plus précise). Ce dernier est d'expression locale, souvent mono-lexicale, caractérisé d'une part par une amplitude de F0 très importante, ce qui implique des valeurs maximales de F0, et d'autre part souvent lié à une rupture mélodique ou prosodique de la trame linguistique, c'est-à-dire en dehors des frontières de syntagmes (Caelen-Haumont 1978/1981).

Plus précisément, dans cette acception, l'intonation est la contrepartie prosodique de la chaîne linguistique : non seulement en effet elle sert de support acoustique à l'énoncé, mais au sein de cadre socio-culturel propre à la langue, elle met en oeuvre des règles

d'enchaînement linguistique des groupes prosodiques (pseudo-) syntaxiques, facilitant ainsi le décodage de l'énoncé.

Si l'intonation est le domaine de l'expression sociale de la langue, le mélisme est celui de l'expression individuelle. Le mélisme est ainsi l'expression prosodique d'implication personnelle, subjective, affective, *hic et nunc*, et de ce fait en complémentarité avec l'intonation. Ce phénomène possède, au moins en français, une extension souvent limitée au mot lexical (voire parfois grammatical) et éventuellement à ses déterminants, plus rarement une suite de mots lexicaux (qui est plutôt l'expression prosodique propre aux émotions typées). Ses caractéristiques prosodiques sont de ce fait beaucoup plus ouvertes.

Ce double traitement prosodique intonation / mélisme, établi à partir de l'observation des données, pourrait être la contrepartie prosodique des processus cognitifs que les études en psychologie et neurophysiologie établissent.

Comme on l'a précisé ci-dessus, au niveau des expressions de la parole émotionnelle, Scherer 1989 distingue deux types d'effets, les *effets push* produits par des « *déterminants internes au sujet* » causant des modifications physiologiques en réponse à une activation émotionnelle, et les *effets pull* produits par des « *déterminants externes au sujet* » en relation avec le contexte environnemental social où se trouve le locuteur. Toutefois, selon Scherer, même s'il est parfois possible d'évaluer la charge respective de ces effets pour un locuteur dans une situation donnée, les effets se combinent dans les indices acoustiques de la voix sans qu'on puisse les dissocier.

Plus encore et antérieurement, Scherer et al. 1984b définissant l'approche *configuration*, et l'approche *covariance* (cf ci-dessus), distinguent d'une part un traitement linguistique et socioculturel, et d'autre part, un traitement affectif.

Ceci est par ailleurs confirmé par les travaux récents en neurophysiologie, au niveau des aires cérébrales tout d'abord. Selon Pell 2002 en effet, il semble attesté que l'hémisphère gauche soit plutôt spécialisé dans le traitement de la linguistique et de la prosodie linguistique niveau représentationnel, alors que l'hémisphère droit aurait plus d'aptitude à traiter la prosodie émotionnelle, avec les indices de F0, véhicules attestés des intentions liées aux émotions.

Ainsi d'une part les modèles *configuration* et *covariance* de Scherer qui interviennent dans le traitement de la prosodie, d'autre part les travaux de Pell, valident totalement nos conceptions (relation intonation / mélisme), le modèle de covariance sur le plan théorique, et l'hémisphère droit sur le plan anatomique, rendant compte ou traitant des effets à la fois émotionnels et linguistiques, et symétriquement, le modèle de configuration, et l'hémisphère gauche, se spécialisant dans le traitement linguistique.

3.2. ASYNTAXE ET DISYNTAXE PROSODIQUES

Comme nous l'avons vu, les mélismes s'opposent à l'intonation en tant que deux expressions prosodiques différentes, de type psychologique vs. social. Plus précisément l'usage social de la langue a créé un espace de liberté contrôlé pour l'expression subjective individuelle. De la même façon que l'harmonie d'un groupe peut être remise en cause par l'action d'un seul individu, sur le plan de la parole, le mélisme, en tant qu'expression individuelle et subjective, peut remettre en cause la relation prosodique à l'organisation syntaxique de l'énoncé. En effet les mélismes peuvent en effet être accompagnés de deux processus, à savoir les *disyntaxe* et *asyntaxe* prosodiques (Caelen-Haumont 1978/1981). La disyntaxe prosodique correspond au non-respect des relations syntaxiques. Ce phénomène s'observe par définition en position initiale et interne de groupe et se manifeste par une rupture de niveaux mélodiques entre le mot courant et le mot suivant, un ralentissement du débit, souvent accompagnés d'une insertion de pause.

Son corollaire, qui peut exister parallèlement ou non, est l'asyntaxe prosodique qui se caractérise par un regroupement non syntaxique d'unités, c'est-à-dire d'unités qui ne sont pas en relation de dépendance syntaxique. Moins courants en lecture (sauf si le locuteur s'approprie le texte et s'implique personnellement), ces processus appartiennent surtout au discours spontané.

4. LES MELISMES DANS LA PAROLE

4.1. DEFINITION DU MELISME

Ce terme est emprunté au domaine musical où il désigne dans le chant, un mot dans lequel le nombre de notes est supérieur au nombre de syllabes. Dans notre acception, il désigne une *forme* acoustique et mélodique, structurée phonologiquement. Plus précisément il correspond à une structure phonologique de surface (non dérivable) enchaînant une suite de séquences monotonaux (« monotonal ») ou paire de cibles de niveau tonal identique, par exemple /hh/) ou bitonales (« bitonal ») ou paire de cibles de niveau tonal différent, par exemple /se/), ce qui peut donner par exemple pour un mot une suite telle que /hh he es se/. Par analogie et pour plus de commodité, ces séquences mono- ou bitonales sont appelées des « syllabes tonales ».

Sur le plan prosodique, le mélisme implique les niveaux les plus aigus, intégrés dans des pentes ou des plateaux. Les pentes manifestent une large excursion de F0, interne au mot lexical (voire grammatical), ou externe lorsque le déterminant ou le pronom personnel est par exemple intégré dans cette pente. Cette large excursion mélodique est souvent accompagnée d'un net ralentissement du débit, et parfois d'une augmentation de l'énergie.

Ton	<i>MELISMES</i>			élevé <i>e</i>	moyen <i>m</i>	centré <i>c</i>	bas <i>b</i>	infra <i>i</i>	grave <i>g</i>
	<i>aigu a</i>	<i>supra s</i>	<i>haut h</i>						
<i>a</i>	<i>aa</i>	<i>as</i>	<i>ah</i>	<i>ae</i>	<i>am</i>	<i>ac</i>	<i>ab</i>	<i>ai</i>	<i>ag</i>
<i>s</i>	<i>sa</i>	<i>ss</i>	<i>sh</i>	<i>se</i>	<i>sm</i>	<i>sc</i>	<i>sb</i>	<i>si</i>	<i>sg</i>
<i>h</i>	<i>ha</i>	<i>hs</i>	hh	he	hm	<i>hc</i>	<i>hb</i>	<i>hi</i>	<i>hg</i>
<i>e</i>	<i>ea</i>	<i>es</i>	eh	ee	em	ec	eb	ei	eg
<i>m</i>	<i>ma</i>	<i>ms</i>	mh	me	mm	mc	mb	mi	mg
<i>c</i>	<i>ca</i>	<i>cs</i>	<i>ch</i>	ce	cm	cc	cb	ci	cg
<i>b</i>	<i>ba</i>	<i>bs</i>	<i>bh</i>	be	bm	bc	bb	bi	bg
<i>i</i>	<i>ia</i>	<i>is</i>	<i>ih</i>	ie	im	ic	ib	ii	ig
<i>g</i>	<i>ga</i>	<i>gs</i>	<i>gh</i>	ge	gm	gc	gb	gi	gg

Tableau 1 : Matrice des 81 séquences tonales pour la description des configurations mélodiques des mots, en particulier les proéminences mélodiques subjectives (ou mélismes). Les séquences tonales en italiques gras sur fond grisé correspondent dans notre conception aux mélismes.

Ces proéminences mélodiques, souvent de forme complexe, sont produites dans le cadre du mot ou d'une suite de mots, en relation avec une intention subjective et/ou une affectivité sous-jacentes. La procédure MELISM décrite ci-après permet à partir de la courbe de fréquence fondamentale stylisée par MOMEL, de détecter les mélismes, de les découper en « syllabes tonales », et de les coder en fonction de leur hauteur au sein d'une échelle à 9 niveaux (niveaux absolus exprimés en demi-tons), le tout automatiquement. Ces syllabes mono- ou bitonales sont obtenues sous certaines conditions aux endroits de changement d'orientation de pente mélodique. Ce codage phonologique de surface permet donc de coder la suite des modulations mélodiques significatives.

Pour ce faire, relativement au modèle originel de Delattre, chacun des 4 niveaux a été

divisé en 3, le niveau central de ces 3 parties représentant 50% du niveau originel, chacune des marges inférieure et supérieure, 25%. Ceci a pour effet de créer une succession de 7 plages de 50% (échelle logarithmique), par l'addition 2 à 2 de ces marges de 25%, déterminant ainsi les 9 niveaux (tableau 2 ci-dessous) : aigu (a), supérieur (s), haut (h), élevé (e), moyen (m), centré (c), bas (b), inférieur (i), grave (g). Les cases en grisé correspondent aux séquences tonales qui dans cette conception identifient un mélisme (Caelen-Haumont, 2004).

Comme on le voit sur le tableau 1 ci-dessus représentant la matrice des 81 séquences tonales, une partie seulement des séquences (soit 40) correspond à la définition des mélismes, à savoir celles qui intègrent les niveaux les plus aigus, *a*, *s*, et sous condition que l'amplitude tonale soit suffisante, le niveau *h*. La limite a été fixée empiriquement à une amplitude supérieure à 2 niveaux (soit pour les 7 niveaux jusqu'à *h*, $2/7$ de l'amplitude *g/h*), ce qui exclut les séquences *eh* et *mh* pour le mélisme.

4.2. UN OUTIL DE CALCUL AUTOMATIQUE DES SEQUENCES TONALES : MELISM

Dans le domaine du codage de la prosodie émotionnelle, un outil d'analyse automatique a donc été récemment défini (Caelen-Haumont et Auran, 2004) permettant de décrire finement les modulations produites sous l'effet de l'affectivité (émotions, « émotion-racine », attitudes ...), ces modulations ayant la propriété d'être remarquables par leurs registres, leurs contours, leur amplitude et/ou leur complexité. Script et Manuel d'utilisation sont téléchargeables (<http://www.lpl.univ-aix.fr/~caelen-haumont/>).

MELISM est une procédure de codage discret de la mélodie, fonctionnant en quasi temps réel, et intégrée sous PRAAT (Boersma 2001) sous la forme d'un script. Généralement appliquée à des séquences sonores courtes (mot ou suite de mots), la procédure peut aussi s'appliquer à tout extrait de parole, quelle que soit sa longueur.

Pour fonctionner, MELISM suppose :

- une segmentation préalable du fichier sonore (format lisible par Praat) en unités linguistiques jugées pertinentes (mots, mots prosodiques, etc.) sur lesquelles l'étude des excursions mélodiques repose,
- une stylisation de la courbe de fréquence fondamentale par détermination de points cibles effectuée par l'algorithme MOMEL [HIR 93], également intégré sous Praat, et transparent à l'utilisateur.

La figure 1 ci-dessous montre un « TextGrid » de Praat avec diverses tires propres à la procédure MELISM, avec respectivement de haut en bas, la fenêtre « manipulation » de Praat, comportant le signal de parole, la courbe mélodique stylisée, puis le codage de la procédure MELISM, présentant de bas en haut :

- les valeurs de F0
- leur conversion en demi-tons
- le codage en cibles tonales selon la procédure MELISM (premier passage avant segmentation)
- le recodage au sein de la séquence segmentée, quelle que soit sa nature : mot, groupe, énoncé ... (deuxième passage), et détermination de la cible la plus haute au sein de l'unité segmentée,
- l'annotation en séquences mono- ou bitonales (les « syllabes tonales »).
- la segmentation en mots.

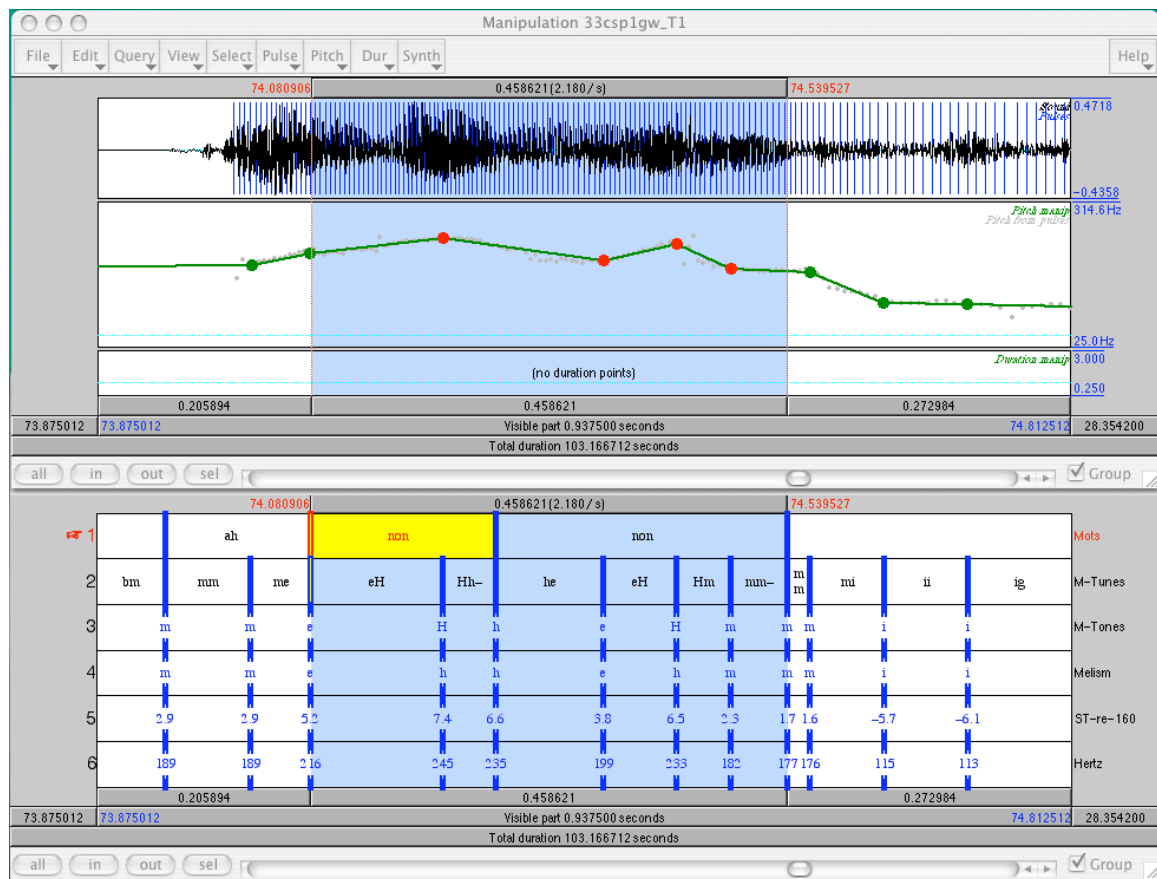


Figure 1 : Codage MELISM et TextGrid de Praat-MOMEL (Boersma 2001 ; Hirst et al. 1993) de la séquence : « ah non non » prononcée par une locutrice du sud-ouest de la France.

Les séquences bitonales lient nécessairement deux plages de hauteurs fréquentielles différentes. Par exemple, dans la figure 1 ci-dessus, deux séquences bitonales sont réalisées au sein du mélisme du mot *non*, à savoir *eH* et *Hh-*. La majuscule signale simplement le niveau le plus élevé dans le mot mélismé.

Inversement, les séquences monotonaux recouvrent des amplitudes faibles ou nulles. Si l'amplitude est nulle, il s'agit d'un plateau. Dans ce cas, la valeur de F0 est stable et concrètement dans la limite du seuil, elle s'écarte au plus de 3%. Les séquences sont de la forme : *ss*, *mm*, *gg*, etc. Dans l'exemple de la figure 1 ci-dessus, la séquence *Hh-* dans le mot « *non* » est un plateau. Ce dernier mot n'est pas mélismé.

Si l'amplitude est un peu plus importante (amplitude faible), il n'est plus question de plateau, mais de « variation intra-plage ». La variation est alors comprise entre 3% et l'amplitude maximale du niveau où elle se réalise, sinon plus importante, la variation devient bitonale. Ces variations intra-plage sont identifiées par les signes « + » pour une pente montante dans le cadre du niveau, et « - » pour une pente descendante. dans l'exemple ci-dessus, nous avons *Hh-* pour le premier *non*, et *mm-* pour le deuxième. On peut considérer que ces variations intra-plage constituent des allotones mélodiques.

En guise d'illustration, deux autres extraits de parole spontanée sont proposés ci-dessus, montrant le premier, l'émotion primaire de la joie, le second, deux attitudes. Le premier ci-dessus cf. figure 2 : « [*cela m'a marquée petite (rires)*] *j'me rappelle le grenier (rires) ...* », est énoncé par une locutrice relatant avec une émotion de *joie* un souvenir d'enfance, en l'occurrence l'expérience d'un grenier à gouttières. On peut ainsi observer sur cet extrait :

- la portion d'énoncé sélectionnée dans la fenêtre du logiciel Praat (Boersma 2001/2005) « *j' me rappelle le grenier* » qui est prononcée sous l'emprise de la joie, portion ponctuée par les rires de la locutrice,
- dans cet extrait, l'intensité de cette séquence est plus forte que la portion précédente. Comparer à ce sujet, les intensités respectives du signal de parole de la séquence précédente « *cela m'a marquée petite* » et de la séquence émotionnelle (encadré bleu : « *j' me rappelle le grenier* »),
- la hauteur de F0 en observant la courbe de la fréquence fondamentale qui monte vers les régions aiguës du registre de la locutrice, bien au-delà de sa valeur de F0 moyenne de 170 Hz. En effet sur les 9 niveaux retenus par ordre décroissant, *Aigu, Supra, Haut, Elevé, Moyen, Centré, Bas, Infra, Grave* pour étudier avec précision les motifs mélodiques de la voix affective [CAE 04], on note que les valeurs de F0 appartiennent aux niveaux les plus élevés soit *A, S et H* (en majuscules ou minuscules).

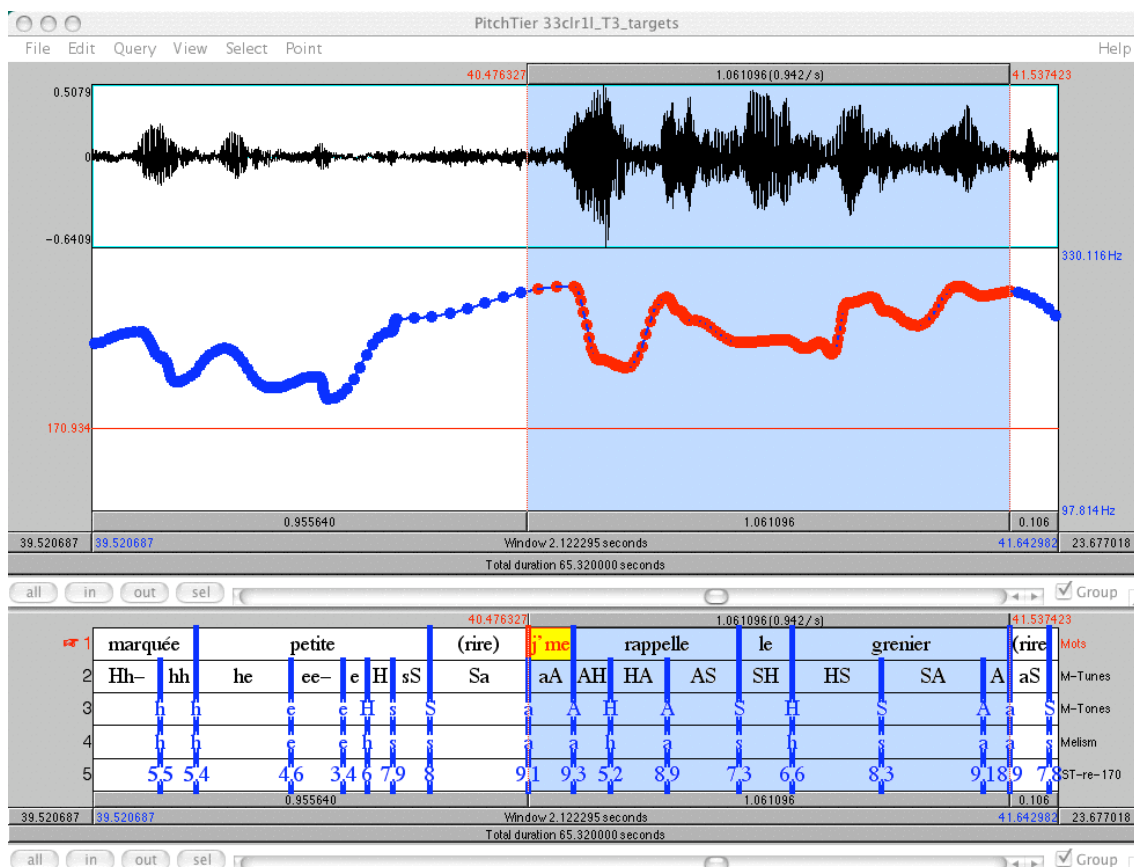


Figure 2. Extrait d'un corpus de parole spontané sous Praat-MOMEL (Boersma 2001 ; Hirst et al. 1993) rappelant un souvenir d'enfance exprimé avec joie: "[cela m'a marquée petite(rires)] je me rappelle le grenier (rires) ... De bas en haut, le signal de parole, la courbe de F0 représentée à l'aide du logiciel MOMEL, avec la moyenne de la F0 de la locutrice 170 Hz représentée sous la forme d'une ligne rouge horizontale, le texte de l'énoncé, le codage des séquences tonales relatives aux unités linguistiques, les cibles tonales codant les valeurs de F0 exprimées en demitons sur la ligne du dessous.

Le deuxième exemple d'une autre locutrice propose une succession de deux attitudes. On peut ainsi voir sur la figure 3 ci-dessous la portion d'énoncé sélectionnée dans la fenêtre du logiciel Praat « *non, c'est des matières sur l'étude de l'agronomie pause comme c'est intéressant ...* ». Dans cet extrait, l'énoncé est exprimé par la jeune-fille à l'encontre de sa famille qui l'a inscrite à la rentrée suivante dans un lycée agricole, alors qu'elle souhaite faire des études commerciales. La partie avant la pause est prononcée avec l'attitude d'un « *agacement contrôlé* », la deuxième partie sous le signe de l'*ironie*,

- on constate que l'ensemble de cet énoncé ne se situe plus dans le registre élevé, mais de part et d'autre de la valeur de F0 moyenne pour cette locutrice 234 Hz,

- la partie avant la pause, exprimant un agacement à l'encontre de sa mère qui tente de lui présenter sa future scolarité sous un jour agréable, se compose d'une succession de pentes alternées assez rapides qui martèlent chacun des mots lexicaux. Les niveaux sont compris entre *b* et *h*, la grande majorité entre *b* et *c*,

- la partie après la pause, s'oppose par sa forme à la précédente. Elle est dite avec *ironie* dans un registre proche mais avec une amplitude de F0 plus restreinte : entre *b* et *m*, avec un plateau initial niveau *b*, ce qui lui confère un contour assez plat et monotone, contrastant vivement avec le contenu de l'énoncé.

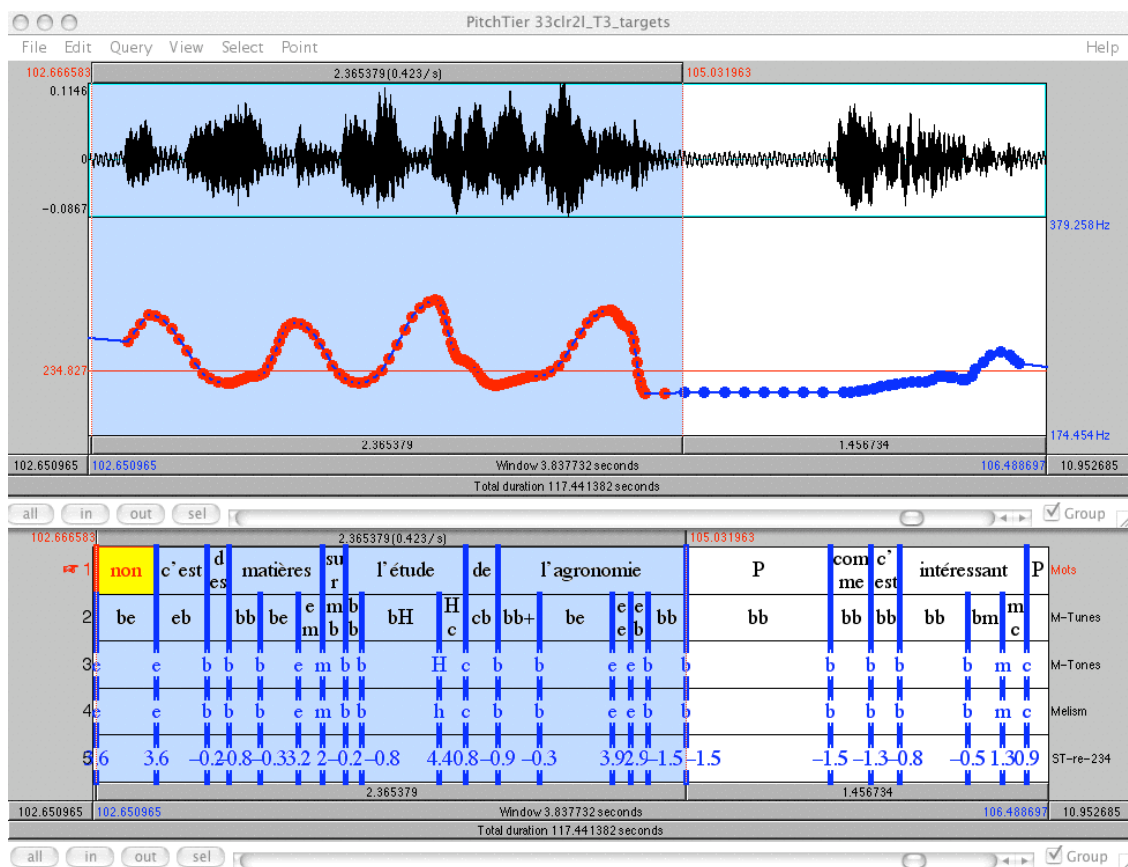


Figure 3. Extrait d'un corpus de parole spontané sous Praat-MOMEL (Boersma 2001 ; Hirst et al. 1993) prononcé sous l'émotion de l'agacement contenu puis de l'ironie : "Non, c'est des matières sur l'étude de l'agronomie [pause] comme c'est intéressant"... De bas en haut, le signal de parole, la courbe de F0, le texte de l'énoncé, le codage des séquences tonales relatives aux unités linguistiques, les cibles tonales codant les valeurs de F0 exprimées en demi-tons sur la ligne du dessous.

En résumé, le concept de mélisme a été défini pour décrire plus objectivement sous l'angle mélodique 1° le processus d'émotion-racine, d'expression lexicale, avec la précision et la mobilité nécessaire pour capter ce phénomène éphémère et hautement variable, 2° ses fonctions phonologiques, linguistiques, pragmatiques.

Par ailleurs, du fait que cette méthode se fonde sur l'amplitude maximale du registre du locuteur, exprimée dans une échelle tonale, cette méthode autorise toute comparaison intralocuteur ou interlocuteur, au sein d'une même langue ou entre plusieurs, (langues à accent libre, à accent lexical, à tons), s'appliquant aussi aux idiolectes et variantes dialectales. Elle peut tout aussi bien décrire des modes d'expression vocale diversifiés, tels que parole, poésie chantée, ou chant. Une autre application possible concerne aussi les dysfonctionnements langagiers.

4.3. FONCTION DES MELISMES

Pour des études plus complètes sur cette question montrant des expérimentations, on peut se reporter aux études antérieures (Caelen-Haumont et Bel, 2000 ; Caelen-

Haumont et Auran, 2004 ; Caelen-Haumont, 2004 ; Caelen-Haumont 2005) qui montrent que dans la parole ou le chant, les mélismes sont révélateurs des valeurs psychologiques propres au locuteur et/ou motivées par la tâche ou le but en cours. Des études récentes (Morange, 2005 ; Vuillet, 2005) ont confirmé par ailleurs avec d'autres corpus et d'autres contextes pragmatiques, nos premiers résultats. En effet, s'investir dans son discours est coûteux en énergie, car cela demande de produire des notes fréquentielles hors du registre ordinaire du locuteur, prononcées souvent avec un ralentissement du débit, et avec une énergie soutenue. Ainsi le locuteur ne s'investit pas sans motivation, ni à propos de n'importe quoi.

Dans le cadre de ce chapitre, on se contentera de montrer deux exemples d'investissement de la part de deux locutrices. Le premier exemple (figure 4 ci-dessous) produit par le même logiciel Praat-MOMEL et INTSMEL, a été réalisée par une locutrice âgée (HV1) de Cussac Fort-Médoc.

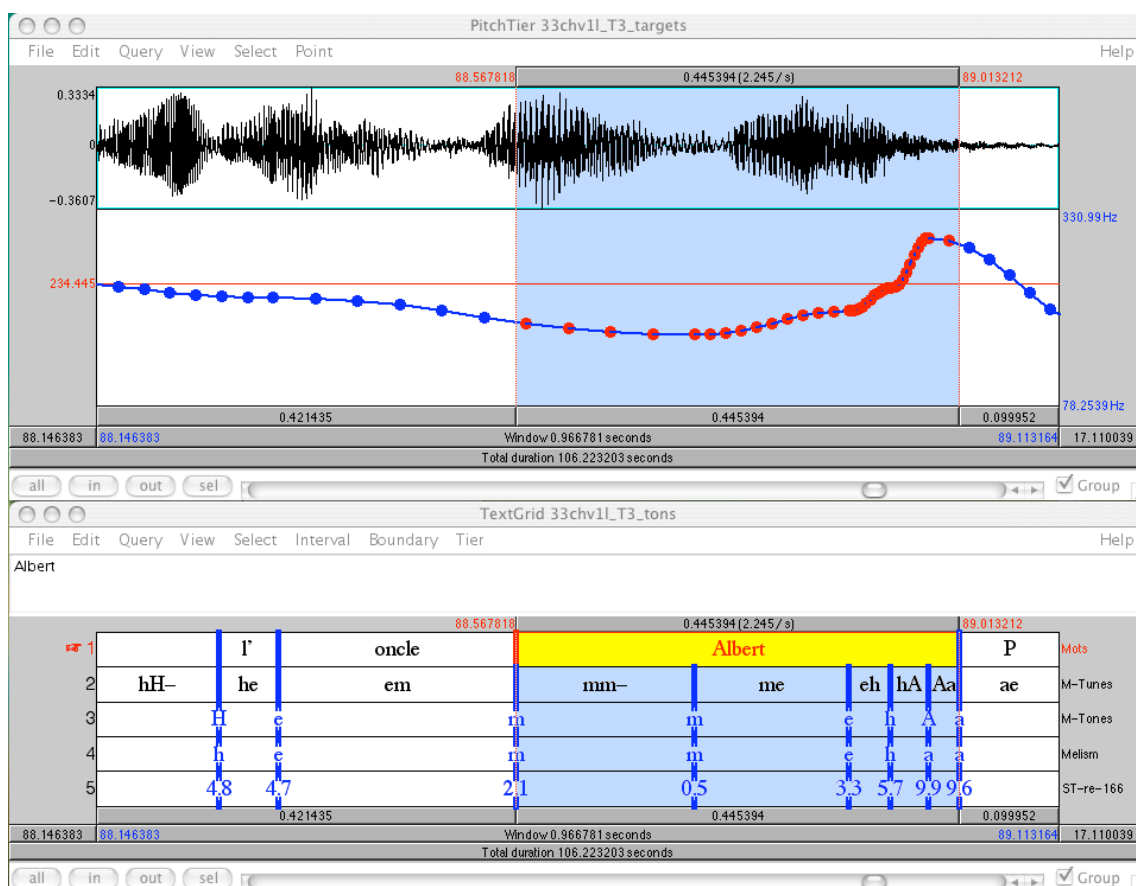


Figure 4 : Mélisme extrait sous Praat-MOMEL (Boersma 2001 ; Hirst et al. 1993), prononcé sur le nom Albert. On note les valeurs très aiguës du niveau a qui signalent l'investissement affectif de la locutrice.

Dans cet extrait le prénom *Albert* est mélismé. La locutrice HV1 raconte en effet que toute jeune mariée, elle a dû s'occuper dans son foyer de quatre hommes (mari, beau-père, père et « l'oncle Albert »), et ce souvenir est resté très vif chez elle. A chaque évocation de chacun de ces hommes qui lui étaient proches, elle réalise un mélisme. Un exemple est donné concernant « l'oncle Albert ». On note que la fréquence est très élevée (niveau A).

Un autre exemple concerne sa petite fille, LR1, âgée d'une quarantaine d'années, et qui est aux commandes, avec son entourage familial, d'une propriété vinicole dans le Médoc. Pour cette famille qui n'a pas bougé depuis des générations, l'entreprise

familiale et la profession représentent des valeurs psychologiques très fortes.



Figure 5 : Mélisme extrait sous Praat-MOMEL (Boersma 2001 ; Hirst et al. 1993), prononcé sur le nom profession. On note également les valeurs très aiguës du niveau a qui signalent l'investissement affectif de la locutrice.

L'extrait ci-dessus (figure 5) montre un mélisme extrêmement délié du terme « profession », possédant huit syllabes tonales, avec une amplitude maximale de 9.8 demi-tons.

Dans ces deux exemples, comme dans les études plus approfondies que nous avons menées sur des corpus entiers, dans les études menées aussi par ailleurs (Morange 2005 ; Vuillet, 2005), il s'avère que les mots mélismés expriment des valeurs personnelles profondes, ou sont le résultat de stratégies de parole bien intégrées, qu'il s'agisse d'adultes ou de très jeunes enfants, toutes étant l'expression de croyances bien ancrées dans la psychologie.

Ainsi selon Vuillet 2005,

les mélismes apparaissent principalement dans les phases de recherche, notamment dans les reformulations¹ qui ont une fonction principale de relance et de valorisation de la parole des enfants. Ils visent à manifester et faire partager un intérêt [...]. Ils se combinent généralement dans ce qu'on peut appeler des séquences affectives —comme le note S. Lebovici (1983), à la différence de l'émotion au sens classique du terme, l'affect a besoin de se représenter—, qui singularisent ce corpus par rapport à ceux étudiés précédemment par G. caelen-Haumont. [...]. Les mélismes apparaissent aussi dans les phases de régulation de l'interaction avec [...] une particulière efficacité. Vuillet 2005 p. 242-3.

¹ Les mots en caractères gras et/ou en italiques appartiennent au texte de J. Vuillet 2005.

Et encore :

par ailleurs, l'étude de ce type de données, non seulement confirme **l'importance quantitative et fonctionnelle des mélismes**², mais elle accentue un aspect très important des mélismes qui avait été signalé dès 2000 par G. caelen-Haumont et B. Bel : les mélismes se confirment être non seulement des marques de la subjectivité, des signes d'une identité, mais aussi une **voie privilégiée pour l'intersubjectivité, la construction du lien, la construction d'un soi groupal** [...] le mélisme n'est pas seulement lié au mot en tant que niveau de l'analyse linguistique, mais au mot en tant que lieu de la présence de l'autre, en tant que sceau posé sur la réalité, en tant que lieu d'hypothèses et de projections, en tant que noeud de mémoire, en tant que niveau peut-être fondamental du fonctionnement cognitif, en tant que palimpseste de l'expérience corporelle. Vuillet 2005, p. 250-1.

Ces croyances fonctionnent en fait comme des véritables sources d'énergie pour la prosodie, ce qui entraîne la production de valeurs fréquentielles élevées, mélodiquement très mobiles, avec des ruptures de niveaux bien marquées dans l'enchaînement mélodique.

5. CONCLUSION

Comme on l'a précisé dans le chapitre précédent, la fonction pragmatique du mélisme est ainsi d'exprimer une croyance ou de faire-croire, c'est-à-dire de partager sa croyance avec son ou ses interlocuteurs. Les mélismes par le registre aigu (*a, s et h*) exercent indépendamment du linguistique, une fonction perlocutoire par l'action sur autrui, volontaire, consciente ou pas. Ce sont sans doute ces notes aiguës qui permettent le repérage auditif d'une vérité du locuteur, d'une subjectivité, de cette émotion-racine à laquelle l'auditeur peut adhérer ou qu'il peut réfuter. En tous les cas, d'une manière ou d'une autre, l'auditeur est lui aussi conditionné par le registre affectif qu'il perçoit.

Comme nous avons pu le constater et le comprendre au cours de ce chapitre, le domaine de l'émotion-racine est l'une des expressions-clef du locuteur dans sa parole ou son chant spontanés. C'est l'état prépondérant dans lequel se trouve le locuteur qui se sent impliqué dans un contexte de discours qui le motive, quelle qu'en soit la raison, et qui de ce fait engendre dans son énoncé les indices prosodiques spécifiques qui portent la marque de son investissement. En retour, ces indices au-delà de la composante linguistique, ont toutes les chances d'être perçus par son ou ses interlocuteur(s) comme tels, ce qui provoque en chaîne des réactions affectives et les indices prosodiques qui leur sont corrélés.

A ce titre, quels que soient le mode d'expression orale, chant ou parole spontanés, et la culture, les espaces lexical et mélodique du mélisme sont à ce titre un lieu de rencontre avec l'autre, de communication privilégiée de l'intersubjectivité, voire du passionnel, en un mot du *sens vivant*.

² cf note 1.