

Laurent FAURÉ  
Praxiling, UMR CNRS 5475 Discours, textualité et production du sens  
Université Paul-Valéry – Montpellier III  
Laurent.Faure@univ-montp3.fr

---

## **Transcrire les données vocales : en quoi les corrélats prosodiques des interjections sont-ils notables ?**

Chercher à rendre compte des interjections vocaliques (*ah, oh, ouh, eh*, etc. : désormais « Voc ») à l'oral, c'est se heurter, d'emblée, à un problème de transcription à la fois pratique et théorique<sup>1</sup>. D'une part, en effet, la représentation usuelle qui en est donnée renvoie, pour l'ensemble des interjections — en particulier onomatopéiques —, à un niveau de réalité vocale : l'instabilité de leurs fonctionnements phonologiques (s'agit-il toujours de phonèmes de la langue observée ?) rendrait chaque occurrence foncièrement plastique et dépendante du contenu expressif de sa réalisation phonétique, délicate à restituer sous une stylisation graphique. On peut d'autre part se demander si la valeur sémantico-pragmatique de ces mots du discours suffit à les élever au rang d'*énoncés*. À cette catégorisation correspond effectivement un ensemble de données intonatives, à distinguer d'aspects phoniques propres à une réalisation individuelle qui serait nourrie par l'affect ou sa représentation expressive. On le voit : trancher le problème passe par une décision de transcription fondée sur une analyse fine des données observées au plan vocal. Ce questionnement induit l'intérêt pour la valeur que le contour intonatif des Voc peut revêtir en discours, au même titre que pour les autres segments énonciatifs.

---

1 Cette contribution a bénéficié des stimulantes suggestions de J.-M. Barbéris et d'un relecteur anonyme que je tiens à remercier au même titre que B. Bonu, pour sa confiance et ses conseils dans la phase préparatoire du présent travail.

On restituera la problématique à l'analyse conversationnelle qui tend à replacer dans une perspective séquentielle et temporellement située la forme interjective d'enregistrement d'un élément contextuel. Les options de transcription s'avèrent dès lors cruciales, non seulement au plan de la divulgation des données analysées, pour l'exemplification, mais dès le départ, pour tenter de mettre en forme des observables et en rendre compte systématiquement. Un double projet anime ainsi le présent travail : il s'agit dans un premier temps de justifier de choix de transcription en relation avec l'analyse de données vocales, dans la visée spécifique d'observer un objet linguistique et interactionnel particulier ; de montrer ensuite en quoi la transcription, loin d'être un produit stabilisé, doit être conçue comme une pratique stratifiée et inscrite dans des aller-retours séquentiels entre l'analyse et son objet. Les étapes de cette pratique<sup>2</sup>, sont à envisager pleinement comme autant de procédures de l'analyse, non seulement pour leur caractère de mise en forme des données mais encore de dégagement des observables induisant l'interprétation. En réfutation d'une expressivité définitoire des Voc, la problématique reposera en l'occurrence sur la notion de *corrélats* prosodiques des seuls *ah* et *oh* du français. Les exemples seront prélevés d'un corpus sonore d'interactions en situation de service dans des bureaux de poste<sup>3</sup> de la région de Béziers (Hérault). Ils ouvriront à des études de cas visant à illustrer l'intérêt analytique qu'il y a à transcrire la matière fine des fonctionnements interactionnels qui fait l'objet de l'exposé.

### **1. Une option de transcription de l'intonation : le système CATS... adapté à toutes fins pratiques**

**1.1.** Les phénomènes prosodiques qui intéressent l'analyse de conversation (désormais : AC) sont aussi ceux qui s'avèrent *pertinents* pour les participants (cf. Local 1996 : 177-179, Léon et de Fornel 1997 : 103), lesquels y recourent comme à autant de ressources (cf. Léon 1999 : 37 *sqq.*).

---

2 Elle relève des catégorisations théoriques que Lafont (1978) regroupe sous la notion de *praxis de linguistique*. C'est à cette dimension analytique de la transcription que sont précisément vouées les contributions du présent recueil.

3 Respectivement dans les villages de Sérignan (Sér1) et Thézan (Thé1 et 2).

Cela induit des pratiques de transcription qui mettent l'accent sur la perception<sup>4</sup>, l'analyste ayant à charge de rendre compte de ce qui se passe pour les interactants, sur la foi de leur comportement attesté et non de ce qui extérieurement paraît se produire. Le système CATS (Conversation Analysis Transcription System) initié par Gail Jefferson et explicité par Atkinson et Heritage (1984 : IX-XVI) choisit d'intégrer dans sa notation (orthographique) un certain nombre d'informations perceptuelles relevant de la prosodie (*cf.* en particulier Couper-Kulhen et Selting 1996 : 39-47). Cette démarche offre toutefois un cadre suffisamment souple pour ne pas exclure *a priori* l'ajout de précisions utiles à l'AC et qui nécessiterait, par exemple, une partition entre données verbales et vocales. Ainsi, dans ses travaux récents, Jacqueline Léon propose-t-elle des conventions de transcription qui reposent sur une dissociation de la ligne orthographique (selon le système CATS) et d'une notation de la prosodie (inspirée de Mertens 1990). Ce dispositif permet une représentation efficace des données vocales perçues et transcrites. Sans négliger cette approche perceptive (et surtout ses attendus méthodologiques du point de vue conversationnaliste<sup>5</sup>), ma propre problématique me conduit toutefois à recourir plus systématiquement aux tracés d'analyseurs du signal. Plusieurs considérations m'y amènent.

- 
- 4 Dans la reconnaissance des faits vocaux, concéder à l'oreille des aptitudes supérieures à celle des analyseurs instrumentaux — dont les réglages sophistiqués demeurent encore délicats — repose sur une réalité adaptative, notamment en conversation naturelle. Cette option justifie d'ailleurs de reconduire l'hypothèse du rôle *intégrateur* de la prosodie (au sens de ressource globale pour l'activité séquentielle) dans la gestion de la prise de tour, défendue d'un point de vue conversationnaliste par J. Léon (1999 : 37 *sqq.*). En effet, l'examen de la parole en interaction favorise particulièrement un tel choix, dont procède fondamentalement notamment la dimension contextualisatrice. La prudence exigible du non-spécialiste m'impose cependant de souligner la relative (mais attestée par les recherches dans le domaine) complémentarité avec les caractères de l'intonation généralement recensés par la littérature spécialisée : démarcatif, identificateur ou expressif, en évitement des exclusives mutuelles. Bien d'autres fonctions sont au reste assignables à l'intonation (*cf.* en particulier Lacheret-Dujour et Beaugendre 2000).
- 5 En recourant à la notion de perception par les participants des faits verbo-vocaux pertinents, l'AC vise la dimension actionnelle de la séquentialité et non son seul aspect linguistique, auquel je restreins ma présente investigation. Cette décision fédère les options de dégageant d'unités discrètes qui n'oblitérent certes pas la réalité continue des fonctionnements intonatifs à d'autres niveaux d'analyse.

On peut être, d'abord, en droit de relativiser le bien-fondé de l'option strictement perceptive, dans la mesure où les conditions matérielles de la transcription ne sont jamais celle du contexte des échanges enregistrés, même lorsque le transcrip-teur/analyste y a pris part. Son interprétation, déjà faillible dans ce cas précis, peut, plus généralement, s'avérer reconstitution erronée de ce qu'auront proprement interprété les interactants (le processus de catégorisation est, en particulier, intimement lié à celui de la perception, lui-même indissociable de l'interprétation empirique). La pratique analytique, si on lui reconnaît bien, comme le fait l'AC, de relever des mêmes procédures fondamentales que celles auxquelles recourent les participants est au risque de tels écueils, du moins dans la phase — cruciale — de représentation des données<sup>6</sup>. Aussi, quand elle est possible, la vérification instrumentale systématique des données vocales observées apparaît indispensable. Le tracé d'un analyseur peut pallier en partie la faillibilité d'une oreille moins exercée<sup>7</sup> que celle d'un phonéticien. À l'inverse, invitant à rendre compte des distorsions entre l'oreille et la machine, il permet également la confrontation aux cas perçus par les participants qu'analyseurs et intensimètres ne savent pas distinguer et qui s'avèrent autant de composantes de la réalité à observer (le *chantonnement*, étudié par de Fornel et Léon 1997 relève de cet ensemble<sup>8</sup>).

---

6 Toute démarche analytique comporte du reste un investissement subjectif d'autant plus lié à l'affect que l'objet d'étude en est une représentation attendue (c'est le cas de la notion d'interjection). C'est que l'observateur n'est pas « neutre », *a fortiori* dans une telle situation (Kappas 1991) et a une tendance naturelle à se mettre en phase avec les sujets observés. Cf. aussi Auchlin 1998. Par ailleurs, l'AC relie la notion d'observable (en particulier vocal) à ses implications pratiques dans l'interaction, sur le comportement même des participants, leur synchronisation interactionnelle et les enchaînements qu'ils produisent (approche locale). Mais intégrer le caractère dynamique de cette réalité séquentielle n'ouvre pas d'emblée à l'abandon du formatage discret, propre à la langue et qui sert de support à cette plasticité évolutive (cf. *infra* 2.1).

7 Un cas qui vaut malheureusement pour l'auteur de ces lignes...

8 Le caractère problématique des distorsions conduit à les mettre en relief par des notations spécifiques en les dissociant clairement des autres phénomènes. Cet aspect, assez souvent pointé, n'est pas systématisé dans la littérature spécialisée : ainsi, la notation musicale, en dépit de sa consensualité et de la fréquence du type de fait qu'elle met au jour, n'est-elle que *parfois* usitée, pour représenter la mélodie singulière d'un discours.

1.2. Rapporter les données acoustiques au comportement situé des participants peut ainsi faire émerger des ressources interprétatives du point aveugle où se niche parfois l'excès d'évidence analytique. La double approche permet de dégager par croisement des régularités qui peuvent dès lors être consignées et entrer dans des opérations de balisage de formes pour des analyses systématiques<sup>9</sup>. De plus, le raisonnement sur des objets récurrents en discours telles que les Voc implique l'étude comparative entre leurs caractères définitoires et différentiels, qui reposent (non seulement mais) notamment sur des propriétés prosodiques discrètes. C'est cette dimension — dont peut attester la systématisme des résultats de la machine — qu'il reste à encoder grâce à un jeu de graphèmes suffisamment précis et symboliques, voire iconiques, pour restituer ce qui est pertinent pour l'analyse. Il est de plus désormais possible de rendre aisément disponibles les tracés issus des analyseurs numériques du signal pour discuter des interprétations à partir de leur objet.

Dans le même sens, l'analyse de l'oral, à partir de corpus recueillis en situation ordinaire, nécessite, pour progresser réellement, l'extension significative des bases de données (en particulier en français) sur lesquelles elle repose. Or la recherche de régularités<sup>10</sup> propres à la textualité orale ne peut être établie sans la notation systématique et affinée de la relation entre vocal et verbal. Divers systèmes informatisés comme ceux

---

9 Il reste évident que, de façon plus générale, la restitution de *tous* les faits n'est pas cruciale pour l'analyste. Toute prétention à l'exhaustivité serait du reste vouée à l'échec, comme le souligne Ochs (1979). Du moins, la précision et la rigueur de la transcription en amont assurera-t-elle de certaines précautions analytiques.

10 On sait la prévention d'un certain nombre de conversationnalistes vis-à-vis du recours systématique aux statistiques. Du point de vue ethnométhodologique, cette réserve provient essentiellement du caractère globalisant des catégorisations dans les corrélations de phénomènes auxquelles fait appel la méthode statistique : ce traitement a tendance, en effet, à omettre le caractère local, donc sa diversité située — au fondement de la perspective analytique de l'ethnométhodologie —, sans pour autant assurer de son objectivité. L'argument vaut d'être médité. Toutefois, outre le fait qu'il puisse exister aussi des observables mesurables et quantifiables, en ce qui concerne la présente investigation, les récurrences relevées ont émergé de l'étude affinée des données, non de présupposés catégorisants dans la mesure où ce type de fonctionnements n'avait précisément pas fait l'objet de corrélations antérieures.

développés dans le cadre des projets EAGLES<sup>11</sup> (qui prend appui notamment sur les modèles phonétiques SAME et INTSINT<sup>12</sup>) ou CHILDES<sup>13</sup> ont été conçus depuis quelques années pour tenir compte de cette série de contraintes. Les besoins suscités par l'étude du corpus m'ont pour ma part conduit à adopter une solution (sans doute) transitoire, qui est elle-même le produit d'une stratification dans les tentatives de transcription<sup>14</sup>. Elle repose sur la notation d'un double niveau et use de caractères phonétiques pour noter le vocal, sur la base des options du CATS, partiellement adaptées<sup>15</sup>. L'usage de cette représentation symbolique est destinée à faciliter les recherches de concordances informatiques<sup>16</sup>. Je ne puis développer ici l'importante problématique relative à ce dernier aspect ni les raisons techniques de mes propres choix typographiques, en partie distincts de ceux des programmes évoqués ci-dessus. Mon option découle en fait simplement de l'étude des données, dont les propriétés ne

- 
- 11 Les travaux du regroupement international EAGLES (*Expert Advisory Group in Language Engineering Standards*) sur la standardisation des grands corpus transcrits dans le cadre du réseau NERC (*Network of European Reference Corpora*) sont actuellement poursuivis sous le sigle ISLE (*International Standards for Language Engineering*) : <[http://lingue.ilc.pi.cnr.it/EAGLES96/isle/ISLE\\_Home\\_Page.htm](http://lingue.ilc.pi.cnr.it/EAGLES96/isle/ISLE_Home_Page.htm)>.
- 12 *INTERNational System of INtonation Transcription*. On trouvera une description récente et détaillée de ce système par ses concepteurs dans Di Cristo et Hirst (1999).
- 13 Attendus et contenu du programme CHILDES (*CHild Language Data Exchange System*) sont consultables sur le site : <<http://cnts.uia.ac.be/childes/>>.
- 14 L'élaboration d'une application d'aide informatisée à la transcription est ainsi en cours, dans le cadre du laboratoire de l'équipe *Praxiling*, à Montpellier III.
- 15 J'ai conscience que l'option retenue me situe au milieu d'un gué épistémologique : pour l'AC, la transcription est un outil évolutif — qui peut donc s'avérer instable — adapté à l'analyse même, au contraire des linguistiques de corpus qui visent à une représentation aboutie des données recueillies comme préalable à leur analyse. Les deux démarches sont fondamentalement distinctes, sinon en tout point opposées. Pourtant, je plaide bien pour une complémentarité raisonnée des deux perspectives, en ce qu'elle me semble être dans le droit fil des développements de l'analyse de discours et de la linguistique interactionnelle. Ici même et dans des travaux ultérieurs, la praxématique, située en carrefour de ces deux tendances, devrait m'aider à surmonter la contradiction initiale.
- 16 On peut se reporter à l'annexe 1 pour le listage des conventions de transcription. Les catégories en italiques sont des suggestions d'ajout au système CATS. Pour une présentation de trois systèmes recourant au principe de « niveaux » de transcription, cf. Bilger 2001.

correspondent pas toujours aux modes d'étiquetage et d'indexation informatiques proposés.

**1.3.** Le cadre de l'échange spécialisé dont procède la présente livraison des *Cahiers de praxématique* permet, enfin, d'avancer la notion de *plasticité* de la représentation éditoriale des données orales. Plutôt que d'en appeler à la déformabilité de la transcription elle-même en fonction de l'objet d'étude, il semble en effet préférable de recourir à un système ouvert et susceptible à la fois d'accueillir le maximum d'informations en amont et de susciter validation et échanges des corpus entre chercheurs. Il est ainsi désormais possible de représenter diverses strates (dont relèvent les phénomènes gestuels et vocaux de synchronisation avec le verbal) sur des lignes de transcription distinctes. De ce corpus encore brut mais surtout caractérisé par la disponibilité des faits qu'on en retire et toujours associé à l'étude des enregistrements proprement dits, on extraiera bien sûr les observables utiles à l'analyse. C'est dans la deuxième étape que la lisibilité intervient sous l'apparence d'une nouvelle mise en forme des données en fonction de divers impératifs éditoriaux et d'exemplification, en particulier.

Je plaiderai donc pour une déformabilité de la transcription assumée en amont du principe de *fidélité*, qui doit relever des strates de dégagement des observables et rendre disponibles les données pour la recherche (*cf.* utilité des corpus informatisés, susceptibles de restituer en banque de données texte, son, image). Dans cette perspective, la *lisibilité* relève ainsi d'un tout autre niveau de représentation, légitimement plus déformable que le précédent.

## **2. Vocalisations et prosodie**

### **2.1. Position du problème et orientations de l'analyse**

L'approche conversationnaliste s'est distinguée, en ce qui concerne la prosodie, d'une tradition plus structurale, en quête d'invariants identifiables. L'AC, on l'a vu, s'efforce de reconstituer les catégories pertinentes sur lesquelles s'alignent d'authentiques interactants (Couper-Kuhlen et Selting 1996 : 48). Tenues pour des formes indexicales participant de la production d'inférences, les marques prosodiques relèvent de la construction discursive des interactions. Leur étude s'est longtemps centrée sur leur fonctionnement séquentiel, comme formes constitutives

notamment des unités de construction de tour (*cf.* notamment Auer 1996<sup>17</sup>). Et le domaine français n'a guère fourni d'analyses purement conversationnelles et systématiques des indices intonatifs. À ce constat un peu frustrant, on opposera une mise en relation avec des marques du niveau verbal : le cas des vocalisations permet d'ouvrir quelque peu l'angle d'approche. À cette fin, on postulera le caractère interactionnel que revêtent les Voc pour tenter d'en définir l'ordre fonctionnel. Cette propriété passe par un recours direct à la langue, dont relève leur sémantisme. En discours, les Voc se présentent indéniablement comme plus plastiques, dans la mesure où elles donnent à inférer divers effets de sens. Leur évidente déformabilité emprunte directement à la dimension vocale. Mais loin de la considérer comme première, je pose celle-ci comme développant ou bloquant un trait instructionnel sous-jacent, dans le cadre des contraintes induites par les normes interactives et la construction du contexte par les interactants. Seront ainsi conçues comme autant de *corrélats* les formes vocales associées aux procédures sémantiques impliquées par le recours à l'interjectif. Ces corrélats vocaux relèvent de ce que l'on entend par *indices de contextualisation* et en partagent effectivement la vocation à produire des inférences interprétatives au cours de l'interaction située (Gumperz 1982/1989). En évitement provisoire d'une difficulté évidente (statut de la notion d'indice *vs* celle de signal en phonétique), la souplesse de ce cadre définitoire (dont le maillage analytique large a parfois été critiqué) permet d'intégrer l'idée que les fonctionnements vocaux reposent aussi sur une signalétique procédurale, propre à chaque langue, donc en amont de l'interprétation contextuelle. Cette signalétique ne peut toutefois être signifiante qu'en conjonction de paramètres (Morel et Danon-Boileau 1998 : 17-20 ; Selting 1996). On posera donc qu'*actualisée*<sup>18</sup>, l'association (Voc+corrélats) correspond bien à un

---

17 Une partie significative des travaux du recueil de Couper-Kuhlen et Selting (1996) dément évidemment cette tendance, en étendant la notion de fonction dialogale de la prosodie.

18 Au sens praxématique, l'actualisation consiste en la transition vers le discours des programmes capitalisés en langue *via* les opérations menant la pensée à se constituer en acte de langage. Pour Bres (1995) l'émergence interjective correspond au degré initial  $\emptyset$  de l'actualisation phrastique. Barbéris (1995, 1998) y voit, à un autre niveau actualisateur (celui des marques du sujet en langue), l'expression d'une subjectivité dialogique, sur une position de co-orientation, dite en *Idem* analogique. À noter d'autre part

faisceau d'*indices contextualisateurs* tout en reposant sur un jeu de contraintes phono-syntaxiques et interactionnelles plus complexes. C'est ce dernier encodage qui, sous les régularités qu'il exhibe, nécessite l'adaptation des conventions de transcription.

## 2.2. Contours intonatifs et tendances des Voc : premiers balisages

### 2.2.1. Les leçons de l'étude du *oh* anglais

Le rapport entre données vocales dans l'interaction et recours aux Voc, a été en partie examiné par Local (1996). Son étude — inscrite plus largement dans l'élaboration d'une phonétique conversationnelle — met en évidence de façon systématique les multiples modulations vocales (phonétiques et mélodiques) et l'emploi situé de l'interjection anglaise *oh*<sup>19</sup>. L'auteur recourt à la double transcription : celle de Jefferson dans le corps du texte, pour l'analyse des exemples (extrait d'un corpus de conversations téléphoniques) et, en annexe, une représentation phonétique plus précise, assortie des courbes mélodiques, qui exhibe les données obtenues auditivement et celles issues d'un analyseur électronique. Sont étudiés quatre types d'emploi conversationnel de l'occurrence *oh* (libre et isolée, associée à d'autres segments dans un même tour de parole, en réponse à une demande d'information et comme marque de surprise). Il ressort de cette analyse que la tendance préférée dans le recours à la Voc isolée correspond au patron intonatif nettement descendant. La principale variation relevée pour ce type d'occurrence dans le corpus de Local concerne la tenue de la voyelle. Dans les trois autres cas de contexte, on peut relever, en revanche, la neutralité ou l'alignement anticipateur de la mélodie du *oh* sur la courbe du segment qu'il précède, dans le même tour.

L'auteur plaide surtout contre une tendance traditionnelle des études intonatives à affecter un sens à un contour. C'est plutôt le comportement participatif des interactants qui permet d'inférer la valeur qu'eux-mêmes attribuent, en contexte, à telle ou telle forme prosodique. Diverses régu-

---

que, reprise en phonétique, la notion d'actualisation, y désigne notamment une fonction « assumée par l'intonation pour guider l'auditeur dans l'interprétation sémantique des énoncés » (Lacheret-Dujour et Beaugendre 2000 : 276). On peut rapporter cette conception au processus de contextualisation défini par Gumperz. Cette importante problématique fera l'objet d'un travail ultérieur.

19 Elle correspond approximativement au *ah* français.

larités apparaissent, dans ce cadre, pour chaque type d'emploi. En lien avec les aspects séquentiels, leurs propriétés prosodico-phonétiques divergent plus ou moins du modèle de base. Elles procèdent par ailleurs de relations entre l'émission des Voc et les autres segments discursifs de l'énoncé, du tour, de la séquence.

D'un apport considérable par leur minutie, de tels travaux s'inscrivent délibérément dans le cadre de l'AC et du système CATS qui « absorbe » la majeure partie des informations étudiées. Toutefois, on observera que le recours à des outils analytiques et à des conventions de notation spécialisés s'avère crucial dans le cadre d'une phonétique conversationnelle de la parole ordinaire, encore très programmatique. Celle-ci suppose un espace de transcription qui rende compte de la gradualité et des formes prosodiques elles-mêmes : niveaux relatifs, intensité vocale, restitution discrète des contours sont les premiers éléments d'un ensemble voué à en comporter bien d'autres (comme le grain de la voix, et les marques de l'ethos vocal, notamment). Par ailleurs, bien que mon objet présent ne soit pas d'établir de comparaison systématique, les observations de Local permettent aussi de souligner la différence de modulation vocale entre langues (anglais et français, en l'occurrence).

### 2.2.2. Tendances du *ah* français

Le patron intonatif le plus commun correspond, pour le *ah* initial conversationnel français, à une courbe non modulée, légèrement ascendante ou descendante, en niveau bas de la tessiture vocale du locuteur<sup>20</sup>. Prélevées de mon corpus, quelques illustrations de cette configuration

---

20 Prenant appui sur une tradition phonologique de discrétisation des faits prosodiques, j'opte pour la distribution en plages mélodiques chez un même locuteur, contrairement à la procédure utilisée par Local. Cette configuration (issue de la représentation discrète de l'intonation donnée en anglais par Pike et, en français, par P.-R. Léon ou Rossi ; cf. aussi, plus récemment, Morel et Danon-Boileau 1998) présente quatre niveaux de hauteur distribués de part et d'autre d'une médiane (que je note : *i*) et constituent des intervalles déterminés de façon équidistante des points cruciaux hauts (4 : *à*) et bas (1 : *ü*) de chaque locuteur. On peut rapprocher ces derniers des *tons absolus* (haut, bas, medium) de la théorie linéaire-hiérarchique (cf. Hirst et Di Cristo 1999) qui décrit également les mouvements mélodiques en *tons relatifs* (+bas, +haut, identique, rehaussé ou abaissé) entre points-cibles d'un même contour. Je ne puis conserver de cette approche que des notations (sommairement) adaptées à ma problématique (cf. annexe 1).

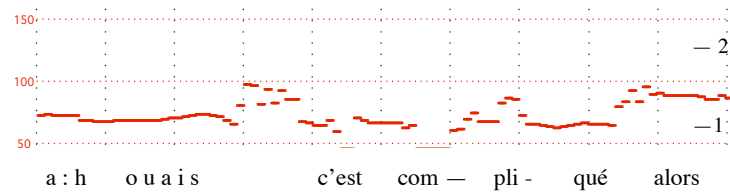
dans des co(n)textes variés donneront un premier aperçu de cette tendance en français ordinaire :

a) en initiale de tour de parole et associée à un segment consécutif  
 [1] Sér 1. 190. B<sup>21</sup> : B, agent en formation, a demandé une précision technique à A, guichetier confirmé en exercice, qui lui annonce une réponse complexe.

186.B et (!)l'ouverture(!) du compte là quand elle t'a porté tout à l'heure<sup>22</sup>  
 [...]

189.A alors là c'est pas par... là ouh là c'est compliqué

→ 190.B a:h ouais c'est compliqué alors m'en parle pas

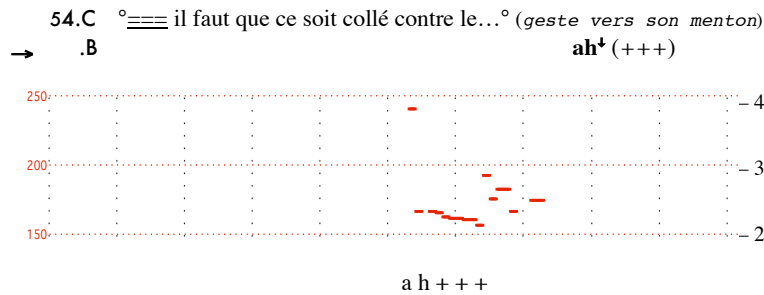


21 Les tracés mélodiques présentés ont été obtenus à l'aide de l'analyseur informatique Spectro (mis au point par F. Longchamp au laboratoire de phonétique de l'Université de Nancy II) que la robustesse et l'ergonomie rendent efficient pour l'analyse de la parole ordinaire en milieu bruyant. Les chiffres de gauche mesurent le ton fondamental en Hertz (Hz), ceux de droite indiquent les niveaux relatifs.

22 B parle de l'opération précédemment soumise par l'agent de contrôle au guichetier.

b) isolée en canal de rétroaction

[2] Sér 1. 54.C : l'enquêteur C signale à l'enquêtée qu'elle doit rajuster son micro contre son larynx.



Cette faible modulation mélodique, majoritaire dans mes données pour les sites d'occurrence a) et b<sup>23</sup>) ne doit toutefois pas masquer les variations (plus ou moins sensibles à l'audition) entre les différentes manifestations de ce type de *ah*, généralement conçu comme marque interactionnelle d'enregistrement d'une information (*news-receipt* au sens de Heritage 1984). On en a du reste un exemple dans la légère modulation qui distingue, par exemple les *Voc* représentées en [1] plane et allongée et en [2], plus brève et qui présente une pente intonative douce (notée : <sup>4</sup>). La distribution des indices est loin d'être radicale : si la tenue vocalique en [1] peut correspondre à une marque de tenue de tour (et n'a donc pas lieu d'être manifestée en [2]), la légère inflexion observée dans le deuxième exemple se rencontre également dans des *ah* en initiale de prise de parole.

Par ailleurs, il existe d'autres occurrences de la même *Voc* dont le format prosodique diffère nettement des cas précédents. L'extrait suivant en fournira un exemple :

23 Ce schème mélodique vaut pour 52 % des occurrences du corpus Sér1 examinées.



Cette dernière association d'une Voc avec un intonème spécialisé du français<sup>25</sup> met en évidence la dissociation en langue du signifiant interjectif et d'une marque modalisatrice, sous laquelle il apparaît, au sens de Bally, actualisé comme *énoncé*. J'ai choisi de noter cette courbe intonative au moyen du graphème spécialisé pour le registre exclamatif. Une représentation fort imparfaite puisqu'elle ne retient que l'un des traits<sup>26</sup> des énoncés habituellement répertoriés comme expression de cette modalité (la Voc n'est, par exemple, pas spécialement marquée en intensité). Cette approximation peut satisfaire cependant à un dépouillement et à une analyse globale de l'extrait. Pour l'étude des fonctionnalités vocales dans le même corpus, cette première approche (perceptive) devra être précisée par une reproduction plus proche de sa réalité phonétique et du dessin stylisé de sa courbe mélodique, codée à l'aide de signes spécialisés : †↑††. Rapportée à la lisibilité éditoriale, une telle précision n'est pas souhaitable quand il s'agit d'exhiber des passages importants du corpus. Elle s'avère, en revanche, déterminante pour l'analyste, notamment s'il veut rendre compte de la récurrence des phénomènes dits d'expressivité et de la typologie de leurs manifestations dans l'interlocution : elle concerne le chercheur intéressé à la gradualité du marquage de la surprise, de l'expression des modalités ou de la représentation des émotions en discours. Les données proprement vocales participent, dans l'exemple observé d'un faisceau de traits prosodiques qui associent à l'entité interjective le statut d'*indice d'expressivité* (cf. notamment Drescher 1998) qui on va le voir, n'est cependant ni suffisant ni nécessaire pour la décrire.

---

25 J'entends par *intonème* un patron intonatif fonctionnel composé de plusieurs paramètres acoustiques. On en rencontre régulièrement le couplage avec les Voc non seulement dans mes données mais aussi dans d'autres types de corpus : cf. en particulier Morel 1995, Morel et Danon-Boileau 1998. Il importe bien entendu d'étudier contrastivement les autres formes d'association relevées. Quelques illustrations en sont proposées dans Fauré (1997). Un dépouillement plus approfondi du corpus a permis d'en préciser les caractéristiques et les sites d'occurrence. Mon objet actuel, en amont d'un essai statistique de typologie actualisée qui sera présentée ultérieurement, se restreint à montrer pourquoi et comment celle-ci est à envisager.

26 Selsing montre que, comme le français, l'allemand parlé exhibe un tel faisceau de traits prosodiques dont la combinaison partielle en situation permet le marquage conversationnel de l'étonnement ou de la surprise (1996 : 239).

### 2.3. Des propriétés prosodiques des Voc à la notation des corrélats

Les spécificités relevées ordonnent bien un traitement distinguant clairement les objets Voc de leurs variations intonatives. Or même chez un phonéticien conversationnaliste comme Local, le fait n'est pas acquis : il nécessite en effet que l'on pose une dimension sémantique préalable à l'unité Voc au plan verbal qui ne lui est pas plus associée chez un auteur comme Heritage (1984)<sup>27</sup>. Du coup, oubliée du sémantisme verbal des Voc, l'AC n'a pas à problématiser leur rapport au vocal. Même si elle atténue une conception purement indexicale des ressources linguistiques, ma proposition n'y est cependant pas opposée. Elle déplace simplement l'interprétation contextuelle sur la sélection et les propriétés prosodiques des Voc. Et la représentation discrète de celles-ci devient un enjeu lors qu'elles procèdent d'opérations de marquages discursifs. Au même titre que dans le cas envisagé en [3], l'allongement syllabique sur la voyelle sélectionnée /a/, /ɔ/, /u/ est déjà une manière de les élever au statut d'énoncé ou *a minima* de segment énonciatif puisqu'il s'agit, par le passage de la micro à la macro-mélodie (soit de la contrainte phonétique au choix énonciatif d'un intonème par le locuteur) d'une affectation de contour à un support segmental (un phonème, une syllabe ou un morphème). On notera du reste que la tenue vocalique, en association avec un contour (*cf.* en particulier Di Cristo 1999) est un démarcatif que les phonéticiens décrivent sous l'accent primaire du français, dont le rôle consiste à mettre en saillance le « mot prosodique ». Je relierai cette proposition à l'observation de certaines grammaires et de quelques linguistes selon laquelle le graphème « h » note déjà cet allongement<sup>28</sup>, relu en expressivité. Que les Voc aient déjà fait l'objet d'une stylisation à l'écrit enregistrant cette spécialisation accentuelle et mélodique, nul doute, donc. Il va de soi que cette représentation orthographique ne rend

27 Il semble du reste qu'une propriété du français, qui donne à observer une relative rareté de variation de timbre aux plans diphtongal (exceptionnelle) mais aussi au monophthongal (quelques exemples), l'oppose à l'instabilité des vocoïdes exhibée pour les interjections anglaises. Peut-on y voir une source de divergence analytique ?

28 Stylisation à ne pas confondre avec l'annotation des effets de soufflerie : *hh* et *·hh*. Pour un exposé critique des choix conversationnalistes en la matière, *cf.* O'Connell et Kowal (1994). Par souci de cohérence orthographique, j'ai choisi par ailleurs de ne pas ajouter de symbole au « h » en cas de simple allongement (propriété du morphème Voc, comme en [2]) par rapport à la voyelle dans d'autres morphèmes.

pas compte, cependant, des variations dans la durée elle-même. Enfin, l'autre dimension, phonologique, nécessitée par la recherche, consiste à codifier de façon discrète les courbes et les informations prosodiques pertinentes.

#### 2.4. Options analytiques et retombées pour la transcription des données

##### 2.4.1. Les traces de l'affect ou la portée d'un préjugé tenace

Si la notation de ces différentes informations revêt une telle importance lors du balisage des données du corpus, c'est qu'elle soustrait l'analyste à d'éventuelles fausses évidences. Elle permet notamment de constater à quel point une certaine conception de l'oral peut être induite, d'une part par la tradition de l'écrit et, d'autre part, par une conception relativement mécaniste des relations de congruence entre syntaxe et prosodie, en évitement d'une problématisation du rapport grammaire/interaction. En considération d'un tel risque, j'ai choisi d'écarter comme postulat définitoire l'association entre l'occurrence d'une Voc et un marquage d'expressivité. C'est à ce préjugé en effet qu'empruntent les réflexions grammaticales et linguistiques s'étant essayées à rendre compte de l'interjectif comme émanation émotionnelle ou affective. En gardent trace les conceptions phonétiques qui n'ont retenu comme manifestation de la Voc que l'énonciation du cri poussé. Ainsi, en phonéticien, D. Autesserre considère-t-il « l'exclamation de peur ou de surprise "Ah !" », comme une *ingressive pulmonique*<sup>29</sup>, « production phonique qui ne se rencontre qu'occasionnellement dans les langues » (1974/1993 : 178-179).

Cette définition inspire une double réflexion : d'une part, elle réduit la Voc à l'émission d'un son « hors système » — non seulement en français, mais « dans les langues » — ce qui l'assigne dans le meilleur des cas en spécificité expressive, ou, au-delà, en *échappée* du sujet. Elle ne rend pas compte, de plus, de la réalité du support vocalique, oblitérée par la manifestation affective acoustiquement marquée. La *réalisation phonétique* devient à la fois la constante d'un marquage émotif et la représentation qui, phonologiquement, subsume toutes les autres occurrences

---

29 Ce phénomène correspond à un gonflement des poumons pour attirer l'air dans les conduits vocaux. Je n'en conteste évidemment pas la réalité mais simplement sa généralisation à tous les « pseudos-criis » *ah*. Il reste d'ailleurs à étudier scrupuleusement la relation entre ces jeux de soufflerie et les fonctionnements interactionnels.

du type. Du coup, la circularité de la description peut induire des réflexes de fausse évidence pour la transcription.

Les jeux intonatifs permettent précisément de comprendre que, non seulement ce postulat ancien (issu des préjugés scriptistes de la grammaire traditionnelle et de l'usage que fait des Voc le texte littéraire, en particulier théâtral) n'est pas applicable à toutes les occurrences du corpus, mais qu'il n'en est, de plus, même pas représentatif. On précisera l'orientation en discutant une citation récente de Morel et Danon-Boileau (2001 : 78) :

*De transcription orthographique difficile (« ah », « ha », « a:::ah »), ces interjections en « a » témoignent toutes d'un mouvement de sidération, de surprise, coïncidant avec le surgissement d'un fait imprévu. Leur durée et la hauteur de  $F_0$  permet toutefois de différencier le mouvement expressif que chacune d'elles traduit.*

Cette position, qui concentre des questionnements plus approfondis sur la valeur discursive et interlocutive de ces « petits mots du discours » (qu'on trouvera notamment dans Morel et Danon-Boileau 1998), s'écarte du naturalisme sous-jacent au point de vue d'Autesserre. Elle présente toutefois d'autres difficultés. D'une part, les divers effets de sens associés au recours en co(n)texte de chaque type de Voc, correspondent à l'interaction des données verbo-vocales et sont plus généralement liées à leur définition par les paramètres indexicaux. La *sidération*, si l'on entend par là le délai médiat par lequel la survenue d'un Fait interactionnel est signalée verbalement (sous le support de la Voc : cf. Morel 1995 : 68) comme transitant par la pensée du locuteur, marque un comportement réactif du sujet-locuteur. Mais elle ne rend pas compte de la réalité interactionnelle du recours à la Voc et de ce que, fonctionnellement, il donne à lire à l'interlocuteur pour lequel elle est proférée.

Pour tenir compte de cette dimension, je poserai que, du fait de son orientation intersubjective<sup>30</sup>, la Voc suture l'antériorité conversationnelle et le dire à venir du locuteur en exercice (cf. Fauré 2002). La vocalisation

---

30 Dans le sens, relevé par Goffman, où les *self-talk* (pseudos-soliloques exclamatifs), émis en / pour un public, le sont aussi ; on trouvera quelques essais pour montrer en quoi les Voc relèvent, dans ce cadre, de la coénonciation dans Fauré 2000 et Fauré et Verine (à paraître).



accompagnent le choix des Voc (et contribuent, en ce sens, à les actualiser). Ils ne sauraient être associés directement à la représentation affective du locuteur. Il importe également de préciser que le modèle traditionnel des Voc comme signaux expressifs fait également l'impasse — tout en relevant leur grande variété prosodique — sur la teneur de la différenciation entre elles (*ah*, *oh*, *ouh* n'ont pas la même valeur discursive) et, du coup, rend extrêmement floue la frontière entre les niveaux vocal et verbal (comme unité de langue).

#### 2.4.2. Dimension interactionnelle de la variable [Voc+corrélats]

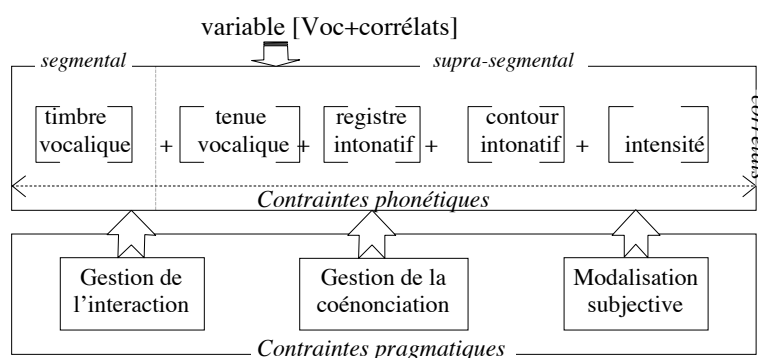
La Voc, unité linguistique recelant des propriétés sémantiques, entre en composition notamment avec des éléments d'autres niveaux communicatifs (vocaux et gestuels), pour former une *unité interactionnelle*. Les retombées pratiques de cette postulation (mise au jour par le feuilletage de la pratique de transcription) et des remarques qui précèdent conduisent à repartir, pour l'analyse, de deux observations fondamentales :

(a) les interjections sont des unités émises en réaction à un phénomène contextuel ou discursif. Elles sont donc généralement observées en début de prise de parole ou en piste de guidage. Une première distinction utile ici car on connaît la différenciation vocale usuelle (donnée par l'intensité du signal) des deux positions au plan de l'interlocution : attaque forte en position haute et peu marquée en position basse. En dehors des cas particuliers, on constate que la Voc ne peut ainsi être, en tant que telle, une simple unité de prise de tour mais relève, interactionnellement, d'un autre niveau de fonctionnement.

(b) La difficulté à transcrire le vocal provient d'un trait justement spécifique aux Voc : leur caractère massivement continu et relatif, qui contrevient aux méthodes statistiques de classement de faits discrets comme à leur représentation graphémique usuelle. À des fins analytiques, la discrétisation (après stylisation) d'indices prosodiques actualisateurs est cependant possible.

Tout l'enjeu de la présente démonstration consiste à mettre en avant que les corrélats ordonnent le recours aux Voc, dans des co(n)textes différenciés. Postulat qu'on schématise ainsi : la sélection par le timbre de la voyelle interjective s'effectue dans le cadre d'une situation pragmatique induisant des contraintes interactives, coénonciatives, ou modalisatrices. Celles-ci entrent en relation avec des règles de niveau

phonologique et phonétique pour déterminer les traits prosodiques associés à la Voc. Ce qu'on peut figurer comme suit :



La transcription fine des associations permet, quand elle est possible, de dégager des tendances d'utilisation dans un corpus donné. Sans préjuger de la valeur de l'unité ainsi constituée, un balisage systématique des données restituera à l'analyse son potentiel interprétatif.

### 3. La notation des corrélats acoustiques des Voc : des outils pour l'analyse

Ce dernier volet sera consacré à l'étude des Voc dans une séquence du corpus transcrite selon la démarche envisagée.

#### 3.1. Transcrire le continu : la co-construction de l'interaction

Préciser les corrélats des Voc c'est aussi montrer *comment* ils constituent, en association avec des morphèmes donnés, des instruments de la gestion de l'intersubjectivité, tant au plan de la progression thématique qu'à celui, plus englobant, des processus de synchronisation entre participants<sup>32</sup>. L'examen de l'exemple [5] va permettre de développer ces aspects de la *contextualisation*.

<sup>32</sup> Il est bien entendu que mon objet est limité aux formes verbo-vocales dont je ne peux que postuler ici qu'elles contribuent, dans une perspective multimodale et située, à l'instar des autres types de phénomènes, à la construction interactive.

[5] Thé 1 : *au cours d'une transaction de service qui a nécessité l'impression préalable d'un relevé de compte, le client (X<sub>18</sub>) observe le caractère malcommode du procédé.*

- 252.X<sub>18</sub> à une époque au Crédit Agricole (*raclement de gorge*) y a un moment de ça (*raclement de gorge*) comme i(l)s envoyaient pas de lettre / chaque fois qu'on faisait une opération quand on passait au bureau là-bas / y avait un appareil comme ça (*geste à l'appui*)
- A mh mh  
clac clac clac / et i(l) vous sortait un... / un relevé ↑ de compte ↑ ///
- A mh
- 253.A **a::h**
- 254.X<sub>18</sub> voyez ↑
- 255.A ui
- 256.X<sub>18</sub> et y avait un relevé de compte ↓  
au lieu de vous l'envoyer
- 257.A et ça existe plus ↑
- 258.X<sub>18</sub> **\*a:::h\*** c'est que le Crédit Agricole  
j'ai laissé tomber / y a un moment que j'ai laissé tomber
- -A **\*a::h\***
- 259.A mais vous  
êtes ↑ postier /// vous êtes un ancien postier vous non ↑ vous m'avez dit ↑
- 260.X<sub>18</sub> ui j'av... oui je mais moi c'est au Crédit Agricole qu'i(l)s  
faisaient ça
- 261.A (!) ui (!) ui ui
- 262.X<sub>18</sub> voyez ↑ moi j'avais été au Crédit Agricole  
pasqu'on m'avait prêté des sous pour construire ↑ la première fois ↓
- A ui ui
- et puis après là j'ai chose là moi j'ai dit moi je largue / là dedans /  
voilà ui ui ui
- A  
mais c'était bien on passait / et clac / on sortait un... / un relevé  
eh ui ui ui
- 263.A **oh** mais y a pas si longtemps que ça même ↑
- 264.X<sub>18</sub> **o:::h** j'y suis plus 'allé là: peut-être y a vingt ou trente ans
- -A **ah**
- 265.A **\*a::h\*** ouh alors en effet /

### 3.1.1 Le recours aux Voc dans la séquence

Cette interaction peut être analysée sous divers aspects. Je me contenterai d'examiner les niveaux dont relèvent les fonctionnements vocaux des interjections étudiées<sup>33</sup>. La séquence est initiée par l'introduction d'un topique conversationnel par le client : *<commodité d'un automate délivrant des relevés de compte dans une banque concurrente de la Poste>*. Dans le cadre institutionnel donné, avancer un tel thème induit une comparaison implicite avec les services offerts par le système postal, voire par ses agents eux-mêmes. Pour chaque participant, c'est de la gestion même de ce topique que procède le recours aux Voc et à leurs variations locales. Noter ces dernières relèvent donc non d'une prétention à l'exhaustivité mais bien d'une ressource analytique, comme peut l'illustrer la représentation schématique<sup>34</sup> de ces enchaînements :

X (client)	A (guichetier)
→ Introduction du topique →	<b>a::h</b> Enregistrement + validation
[...]	[...]
Enregistrement + Réponse <b>*a::::h*</b> [...] y a un moment que j'ai laissé tomber	← <i>ça existe plus ?</i> Sommutation (demande d'explicitation) ←
↻ Explicitation →	<b>*a::h*</b> Enregistrement + validation marquée
[...]	[...]
→ Auto-justification <mini-récit> au passé →	<b>oh</b> Sommutation <i>mais y a pas si longtemps ?</i> ↻
Réponse <b>o::h</b> j'y suis plus allé	← (mise en débat de l'ancienneté des faits)
↻ Amorce d'évaluation <i>peut-être</i> →	<b>ah</b> Enregistrement
↻ Réponse (suite) y a vingt ou trente ans →	Enregistrement marqué <b>*a::h* ouh</b>

33 Les paramètres acoustiques observables sont portés dans le tableau récapitulatif en annexe 2.

34 Les flèches figurent le sens et la hiérarchie de la continuité séquentielle.

La séquence révèle un jeu d'échos entre les Voc, leur profération et leurs émetteurs respectifs. À diverses étapes, elles sont placées à l'interface des activités discursives successives, recouvrant, selon leur placement, une double facette réactive (enregistrement) et anticipatrice (préface à la réponse). En réagissant d'abord au souvenir du client puis à la localisation de son récit dans un passé relativement éloigné alors que l'instrument évoqué paraît plus récent, A fait un emploi plus systématique des Voc. Il les utilise principalement en canal de rétroaction ou en emploi isolé, donc comme signaux de guidance inférentielle (sauf en 263 et 265). À un moment donné, il apparaît que X répond aux réactions de A comme à autant de demandes indirectes (actes de sommation) quant à son ancien statut professionnel (agent de la poste) et sa situation actuelle vis-à-vis du Crédit Agricole<sup>35</sup>. Ainsi placé en position basse, dans les deux cas où X recourt aux Voc, il duplique, en l'allongeant, celle que vient d'émettre son interlocuteur.

### 3.1.2. Catégorisation et gradualité : timbre et signifiant vocal

L'écoute minutieuse de la bande et l'examen du spectrogramme révèlent un phonétisme intéressant de la Voc (264) : elle tend vers l'émission d'un [ɔ:::] (soit un «a» postérieur arrondi, prononcé sur le mode du « craquement vocal » et qui partage ses autres traits avec [ɔ]). Son allongement peut être rapporté à la durée du *ah* émis précédemment par le même locuteur (Voc 258 : [a:::]). Le choix catégorisant de la transcription « oh » correspond d'une part à mes premières perceptions d'auditeur (participant) et de transcripteur (locuteur francophone de la région languedocienne<sup>36</sup>) et d'autre part à une attestation systématique de la présence attendue de cette vocalisation en discours. Pour problématiser cette option, j'ajouterai que je n'ai avancé l'hypothèse d'une opposition sémantique en contexte des deux Voc<sup>37</sup>, qu'à une date ultérieure, éloignée de celle de la transcription. Il me faut donc rendre compte à la fois de cette conviction de locuteur, de son appréhension théorique même et de la raison pour laquelle la différence [ɔ] vs [ɔ] ne suffit à en faire ni

---

35 Organisme bancaire privé.

36 Ceci expliquant évidemment cela.

37 On tente en effet de montrer ailleurs (Fauré 2000), qu'en français, l'émission de *oh* par L2 infère la mise en débat rétrospective des présupposés de L1.

un *ah* approximatif ni un *oh* raté<sup>38</sup>. J'avancerai l'hypothèse suivante, dans le droit fil des observations précédemment émises sur la pratique de transcription. En fait, l'alignement intersubjectif sur la vocalisation précédente relève du plan vocal et de l'interlocution tandis que la sélection verbale du morphème *oh* (*vs ah*) correspond à une saisie par le sujet et à son engagement interactionnel. Or ce dernier plan l'emporte dans la pratique analytique des participants (et celle du transcripteur dans un premier temps, du moins) qui rejettent en inconscience linguistique<sup>39</sup> les traces du processus pour retenir les produits pertinents. Dans le cas présent, la mimétique a réalisé la syntonisation sans atteindre tout à fait au phonème signifiant (cas du ratage). Aux limites de la catégorisation linguistique, la pratique de transcription rencontre ici la modélisation théorique de l'objet d'étude et des niveaux (interlocutif/interactif) de son appréhension. D'autre part, loin d'être idiosyncrasique (un indice sociolinguistique d'identification du locuteur, par exemple<sup>40</sup>), cette émission relève d'un fonctionnement qui trouve sa cohérence à la fois dans sa situation en discours et, dans une certaine mesure, en langue, dont elle intercepte des tensions conflictuelles. Le corrélat prosodique guide interactionnellement l'interprétation locale de la Voc (en dépit de l'altération de son timbre) dont la sélection et le placement correspondent à une procédure sémantique.

Le cas relevé éclaire les questionnements ouverts plus haut. Analyser la réalisation, l'étirement et la forme de la courbe mélodique de ces segments interjectifs permet d'une part de rendre compte du niveau d'actualisation de l'unité Voc comme variable et, d'autre part, de relier cette actualisation à la construction discursive. Celle-ci en fait, dans le

38 Pour l'étude d'un cas problématique de glissement phonique *a-oh* dans le même corpus, cf. Fauré 2000 : 87-89. Un projet d'étude de la réalité diphtongale — avérée en anglais (cf. Local 1996 : 206) mais problématique en français — des Voc et des interjections du français, est en préparation.

39 La notion empruntée à Lafont (1978 : 58-59) réfère à la mise en économie cognitive pratique par le locuteur des opérations langagières de production du sens. Sa mise au jour constitue précisément l'un des objets de la praxis de linguistique (cf. *supra* note 2) qui en est la remontée.

40 Bien que le français méridional ouvre tendanciellement le vocoïde /o/, il n'est donc pas utile de faire recours à une explication relevant fondamentalement d'une variation régionale ou individuelle, que n'attestent pas du reste les autres proférations du locuteur.

cours même de l'interlocution, une unité interactionnelle, porteuse de fonctionnements opératifs (prédicatifs en particulier), un support signifiant qu'on peut rapporter à un énoncé-occurrence. Le caractère situé de cette dimension (son indexicalité) est ici relatif à une forme vocale antérieure, dans le cotexte : l'allongement du *ah* ( $\pm 400$  ms :  $\downarrow^{\uparrow\uparrow} \rightarrow \downarrow$ ) précède celui du *oh* ( $\pm 250$  ms :  $\downarrow^{\uparrow\uparrow} \rightarrow \downarrow$ ) et contribue à lui donner du sens sous l'écholalie prosodique<sup>41</sup>. De même, le *ah* du guichetier (en 265 : [ $?a::$ ]  $\downarrow^{\uparrow} \rightarrow \uparrow \downarrow$ ) correspond à une validation de l'intégralité de la séquence discursive émise par X (alors que celui émis en canal de rétroaction est partiel), dont les implications sont également validées. Son orientation sémantique (l'instruction qu'il active) est distincte, ce qui explique le choix du timbre (différentiel) et la spécificité intonative (niveau mélodique, forme de la courbe :  $\downarrow^{\uparrow} \rightarrow \uparrow \downarrow$  ; et intensité marquée :  $^+a::h^+$ ) de l'unité émise. Pour rendre compte de la valeur discursive et interactionnelle de ces unités, et, préalablement dans le balisage intonatif du corpus, il est donc bien indispensable de signaler à la fois ces similarités et ces importantes différences vocales.

### 3.2 Synthèse

Dans l'analyse des Voc, donner le primat, non à leur dimension affective (qu'il s'agisse de marqueurs émotionnels bruts ou d'indices stylisés d'expressivité), mais à leur statut d'unités linguistiques (verbales) ayant une fonctionnalité interactionnelle, permet de les assigner en variables interactives : [Voc+corrélats].

- (a) Celles-ci, comme unités de langue, interceptent différemment une visée actualisatrice<sup>42</sup> dans leur passage en discours. On peut, à ce titre, attribuer une instruction distincte à chaque Voc, correspondant fondamentalement à une inférenciation conversationnelle (cf. Fauré 2000).

41 Ce type de copie intonatoire se laisse rapporter aux observations de Couper-Kuhlen (1996) sur les reprises de registres lors d'imitations vocales par des animateurs de jeu radiophoniques des réponses de leurs candidats. Cf. aussi le cas du chantonement dans les réponses-échos selon de Fornel et Léon (1997 : 112-116).

42 C'est-à-dire atteignent à un seuil d'actualisation correspondant à une marque linguistique propre au niveau impliqué : cf. *supra* note 18.

- (b) L'actualisation discursive de cette instruction est précisée (ou modifiée) par les fonctionnements vocaux ayant également une valeur signifiante. Le co(n)texte contraint à la fois ces derniers, le placement de la Voc (et donc sa sélection) mais aussi le rapport entre intonation et [Voc+corrélats] actualisée : de même que certains couplages de Voc avec d'autres unités sont inadéquats, toutes les formes prosodiques du français ne sont pas envisageables pour toutes les Voc (ni, d'ailleurs dans l'absolu, pour une seule d'entre elles).

La différentialité des Voc détermine le fonctionnement essentiel de leur sémantisme interactionnel. C'est à elle que je rattache par conséquent la stratégie de transcription (spécialisée pour ces unités) pour laquelle j'ai opté. Cette dernière permet non seulement de décrire plus précisément les oppositions entre *ah* et *oh*, mais encore se veut ouverte au repérage de marques vocales qui participent spécifiquement d'une textualité orale interactionnelle, en donnant à lire les « points de capiton » par lesquels les points de vue des participants élaborent progressivement celle-ci. Le cas des Voc participe de la sorte du renouvellement du regard sur les faits verbo-prosodiques, rapportés à l'examen des placements interactionnels.

On peut ainsi mesurer la portée *locale* des Voc actualisées, c'est-à-dire leur orientation discursive vers la transition ou le tour à venir. Le fait qu'elles constituent des marques réactives est à relier aux rôles même qu'assume alternativement un interactant : parleur ou écouteur. Qu'elles surviennent en attaque de tour ou en piste de co-pilotage, les Voc se donnent à lire comme autant d'anti-signaux de place transitionnelle, indiquant la prise de tour ou, au contraire, l'encouragement à poursuivre (*continuers* pour Schegloff 1982).

### Conclusion

Quel est le degré d'encodage linguistique des corrélats vocaux ? En quoi relèvent-ils d'ailleurs de représentations affectives ou iconiques ? Leur fonctionnement est-il strictement indexical et contextualisateur ou correspond-il à la valeur énonciative régulière de leur forme prosodique rapportée à un énoncé (les intonèmes répertoriés par les descriptions phonologiques) ? Les problèmes rencontrés au cours des opérations de transcription et de balisage restent largement ouverts et conduisent à

s'interroger sur ce qui, généralement, constitue le solde des rapports entre analyse perceptive et instrumentale soit celui des relations entre verbal et « co-verbal ». Or, surtout si cette différence présente des régularités, l'investigation doit s'y attacher. Dans le cas des corrélats prosodiques des Voc, cette réalité a été en partie occultée par divers préjugés grammaticaux et scripturaux qui ont longtemps eux-mêmes participé d'un présupposé relatif à la question du langage affectif. Sauf à reconduire involontairement de telles conceptions, l'AC et la linguistique interactionnelle assument à l'inverse d'avoir à charge l'étude systématique et affinée des données vocales dans une conception intégrée des fonctionnalités langagières et pragmatiques. Cela nécessite une mise au point sur la teneur des observables analysés. Le réglage est corrélé à l'activité pratique et constructiviste de transcription pour laquelle voudrait plaider le présent travail. Ce dernier, sans autre objet que de participer, de façon encore exploratoire et sur un corpus limité, à cette démarche, se veut un essai pour montrer qu'il existe des constantes fonctionnelles dans les jeux prosodiques attachés à des formes significatives de portée instructionnelles telles que les Voc, qui, actualisées en interaction, assurent la remontée du local au global. C'est à ce titre qu'il est possible d'envisager ces formes d'association comme de véritables unités interactionnelles. Une catégorie d'objets discursifs et donc d'analyse qu'il faut transcrire, tant dans leurs caractères définitoires que dans leurs variations différentielles, pour révéler un aspect essentiel de la construction interactive des relations intersubjectives. La transcription comme instrument d'analyse des interactions orales doit rendre compte de ces menus faits, qui relèvent de la cohérence discursive et conversationnelle, sous la forme des sutures inférentielles dont les formats associés [Voc+corrélats] relèvent. Loin d'être des stigmates de l'oral, versés au compte d'un solde expressif, ces marques verbo-prosodiques s'avèrent des indices essentiels de la construction du sens que l'analyste a à charge de mettre au jour, en dépit des filtres métalinguistiques qui conditionnent parfois catégories d'observables et formatages interprétatifs. La notation affinée de ces couplages verbo-vocaux (et, en perspective, de leur dimension gestuelle), se révèle capitale au plan méthodologique dans la mesure où elle touche à l'abandon d'un certain impressionnisme dans la réflexion sur l'indexicalité des ressources linguistiques, bien au-delà des vocalisations.

### RÉFÉRENCES BIBLIOGRAPHIQUES

- Auchlin A., 1998, « Grain fin et rendu émotionnel subtil dans l'observation des interactions : sur le caractère "trans-épistémique" des attributions d'émotions », in J. Cosnier, M. Doury et V. Traverso, *Actes du colloque « Les émotions dans les interactions »*, Chap. 10, Lyon, GRIC, Université Louis Lumière, CD-Rom PUL et Kymé Éditions.
- Auer P., 1996, « On the prosody and syntax of turn-continuations », in *Prosody in Conversation : Interactional Studies*, E. Couper-Kuhlen et M. Selting (eds.), 57-100, Cambridge, Cambridge University Press.
- Autesserre D., 1974/1993, « Ingressive », in G. Mounin, *Dictionnaire de la linguistique*, 1974, rééd., Paris, P.U.F., Quadrige, 178-179.
- Barbérís J.-M., 1995, « L'interjection : de l'affect à la parade, et retour », *Faits de Langue*, 6, 93-104.
- Barbérís J.-M., 1998, « Pour un modèle de l'actualisation intégrateur du sujet », in J.-M. Barbérís, J. Bres, P. Siblot, *De l'actualisation*, Gap, CNRS Éditions, Sciences du Langage, 199-218.
- Blanche-Benveniste C., Jeanjean C., 1987, *Le français parlé. Édition et transcription*, Paris, Didier érudition.
- Bilger M., 2001, « Petite typologie des conventions de transcription à l'oral. Quelques aspects pratiques et théoriques », in M. Bilger (ed.), *Linguistique sur corpus. Études et réflexions*, *Cahiers de l'Université de Perpignan* 31, 77-92.
- Bres J., 1995, « “-Hóu ! Haa ! Yrrââ” : interjection, exclamation, actualisation », *Faits de Langue*, 6, 81-91.
- Couper-Kuhlen E., 1996, « The prosody of repetition : on quoting and mimicry », in E. Couper-Kuhlen et M. Selting (eds.), *Prosody in conversation*, 366-405.
- Couper-Kuhlen E., Selting M., 1996, « Towards an interactional perspective on prosody and a prosodic perspective on interaction, in E. Couper-Kuhlen et M. Selting (eds.), *Prosody in conversation*, 11-56.
- Couper-Kuhlen E., Selting M. (eds.), 1996, *Prosody in conversation*, Cambridge, Cambridge University Press.

- Di Cristo A., 1999, « Le cadre accentuel du français : essai de modélisation », première partie, *Langues*, 3, vol. 2, 184-205.
- Drescher M., 1998, « Les manifestations de l'expressivité en français parlé : le cas de l'interjection », in Rufino, G. (ed.), *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza*, vol. IV, *Le strutture del parlato — Storia linguistica e culturale del Mediterraneo*, Tübingen, Niemeyer, 101-109.
- Fauré L., 1997, « Les interjections à l'oral : quelles valeurs pour les vocalisations ? », *Cahiers de praxématique* 28, 127-148.
- Fauré L., 2000, « Actualisation et production interjective du sens : le cas de la forme vocalique *oh* », *Cahiers de praxématique* 34, Université Paul Valéry — Montpellier III, 77-105.
- Fauré L., 2002, « *Ah bon !, oh non !, ah bè ...* Qu'implique la "prédication" interjective ? », S. Leroy et A. Nowakowska, *Aspects de la prédication, Actes du II<sup>e</sup> colloque de Jeunes Chercheurs-Praxiling*, Université de Montpellier III.
- Fauré L, Verine B., 2001, « Authentifier un discours autre en y mettant du sien : les vocalisations *ah* et *oh* en frontière de discours rapporté direct à l'oral », communication présentée au colloque : *Le discours rapporté, problèmes de frontières*, Ci-Dit, Bruxelles, novembre 2001, à paraître dans les Actes.
- Fornel de M., Léon J., 1997, « Des questions-échos aux réponses-échos. Une approche séquentielle et prosodique des répétitions dans la conversation », *Cahiers de praxématique* 28, 101-126.
- Goffman E., 1981/1987, *Forms of Talk*, trad. frçse. *Façons de parler*, Paris, Minuit.
- Grosjean M., 1995, « Voix et contextualisation », in D. Véronique et R. Vion (eds.), *Modèles de l'interaction verbale*, Aix-en-Provence, Publications de l'Université de Provence, 461-472.
- Gumperz J. J., 1982/1989, *Discourse strategies*, London, Cambridge University Press.
- Heritage J., 1984, « A change-of-state token and aspects of its sequential placement », in J. M. Atkinson et J. Heritage, *Structures of Social Action. Studies in Conversation Analysis*, Paris, Cambridge, Maison des Sciences de l'Homme et Cambridge University Press, 299-345.

- Hirst D., Di Cristo A.,  
1999, « A survey of intonation systems », in D. Hirst et A. Di Cristo (eds.), *Intonation Systems. A survey of twenty languages*, Cambridge, Cambridge University Press, 1-44.
- Kappas A.,  
1991, « The illusion of the neutral observer : the communication of emotion », *Cahiers de linguistique française* 12, 153-168.
- Lacheret-Dujour A., Beaugendre F.,  
1999, *La prosodie du français*, Paris, CNRS Éditions, Langage.
- Lafont R.,  
1978, *Le Travail et la Langue*, Paris, Flammarion.
- Léon J.,  
1999, *Les entretiens publics en France. Analyse conversationnelle et prosodique*, Paris, CNRS Éditions (Sciences du langage).
- Local J.,  
1996, « Conversational phonetics : some aspects of new receipts in everyday talk », in E. Couper-Kuhlen et M. Selting (eds.), *Prosody in conversation*, 177-230.
- Mertens P.,  
1990, « Intonation », in C. Blanche-Benveniste et alii (eds.), *Le français parlé. Études grammaticales*, Paris, Éditions du CNRS (Sciences du langage), 159-176.
- Morel M.-A.,  
1995, « L'intonation exclamative dans l'oral spontané », *Faits de Langue*, 6, 63-70.
- Morel M.-A., Danon-Boileau L.,  
2001, « Les productions sonores de l'écouteur du récit : coopération ou subversion ? », *Revue québécoise de linguistique*, 29, 1, Montréal, UQAM, 71-95.
- Morel M.-A., Danon-Boileau L.,  
1998, *Grammaire de l'intonation, l'exemple du français*, Bibliothèque de Faits de Langues, Paris, Ophrys.
- Müller F.-E.,  
1996, « Affiliating and disaffiliating with continuers : prosodic aspects of reciprocity », in E. Couper-Kuhlen et M. Selting (eds.), *Prosody in conversation*, 131-176.
- Ochs E.,  
1979, « Transcription as theory » in E. Ochs et B. B. Schieffelin (eds.), 1979, *Developmental Pragmatics*, New York, Academic Press, 221-301.
- O'Connell D., S. Kowal,  
1994, « Some current transcription systems for spoken discourse : a critical analysis », *Pragmatics* 4, 1, 81-107.

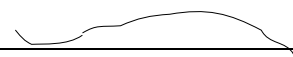
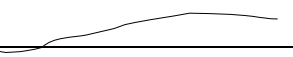
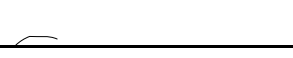


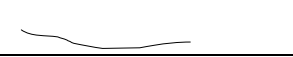

- Schegloff E., 1982, « Discourse as an interactional achievement : some uses of 'uh huh' and other things that come between sentences », in D. Tannen (ed.) *Analyzing Discourse : Text and Talk*, Georgetown University, Round Table on Languages and Linguistics, 1981, Washington, D.C., Georgetown University Press, 71-93.
- Selting M., 1996, « Prosody as an activity-type distinctive-cue : the case of so-called 'astonished' questions in repair initiation », in E. Couper-Kuhlen et M. Selting (eds.), *Prosody in conversation*, 231-270.

Annexe 1 : Conventions de transcription<sup>43</sup>

Adaptation de quelques catégories retenues d'après le système CATS original	Correspondance signalétique	
	Jefferson (1979)	Adaptations proposées
Allongement syllabique	:	:
Pause (± 400 ms)	(.) (0.0)	/
Intonation descendante	.	↓
Intonation continue	,	→
Inflexion croissante	?	↑
Montée ( <i>globale</i> )	↑	↑
Chute ( <i>globale</i> )	↓	↓
Niveau supra haut (Point haut) (4)		⌈
Niveau haut (3)		↑
Niveau médian (2, 5)		⊥
Niveau bas (2)		↓
Niveau infra bas (Point bas) (1)		⌋
Ton animé – exclamation	!	(!) (!)
Intonation rieuse		☺ ☺
Voix chuchotée		* *
Relief (selon la mélodie)	<u>e:r</u> , <u>e:r</u> , e:r	
Chevauchement	[]	<u>alignement</u>
Élévation du volume	MAJ	++
<i>Pic d'intensité faible</i>		△ △
Passage « plus calme »	° °	° °
Inspirations/ expirations audibles	(hhh) / (·hhh)	h, hh / ·h, ·hh
Bruit vocal ou phénomène non analysés	((toux;contexte))	(toux;contexte)
Bloqueur glottal		?
Accélération / décélération du rythme	> / <	> / <
Passage manquant	...	[...]
syllabe inaudible		=
Bruit gênant l'audition		+++
Incertitude	()	{ }

43 Je ne puis détailler ici les choix opérés, ni, *a fortiori* les symboles retenus pour des raisons de commodité éditoriale, en particulier. On aura reconnu un certain nombre d'emprunts aux conventions phonétiques de l'API (révisé 1993) et au système INTSINT. Pour d'autres règles de transcription du français parlé, cf. également Blanche-Benveniste et Jeanjean (1986) ou encore Morel et Danon-Boileau (1998).

## Annexe 2 : Table des exemples analysés

Voc	Occurrence	Réalisation phonétique	Courbe mélodique
<b>Oh</b>	Thé2.99X <sub>3</sub> *o:o:::h <sup>+</sup> pardon	[o:o:::] ↓↑→↑↓	
	Thé1.264X <sub>18</sub> o:::h j'y suis	[ɔ : : :] ↓↑↑→↓	
	Thé1.263A oh mais	[ɔ] ↓↑→↓	
<b>Ah</b>	Sér1.95X <sub>6</sub> ah d'accord	[a] ↓↑↑↑	
	Thé1.258X <sub>18</sub> *a:::h <sup>+</sup> c'est	[a:::] ↓↑↑→↓	
	Thé1.258X <sub>18</sub> (A) (*a::h <sup>+</sup> )	[a : :] ↓↑→↓	
	Thé1. 265.A *a::h <sup>+</sup> / ouh	[?a::?] ↓↑→↑↓	

Laurent FAURÉ

**Transcrire les données vocales : en quoi les corrélats prosodiques des interjections sont-ils notables ?**

La transcription du paramétrage intonatif des interjections vocaliques françaises *ah / oh / ouh* (désormais : Voc) induit directement leur analyse comme objets discursifs et interactionnels. Le paramétrage intonatif permet effectivement, à la perception, de lever l'ambiguïté de leur orientation sémantique, fréquemment rencontrée à la lecture de données transcrites. On postule que ces caractéristiques suprasegmentales, en actualisant discursivement les Voc., déterminent partiellement les effets de sens qu'elles supportent en contexte. En retour, observer l'association de cette dimension vocale à de telles unités permet de préciser la valeur en langue des Voc., grâce, notamment, aux formes d'associations préférentielles qu'elles présentent vocalement. On propose en conséquence certains choix méthodologiques — rapportés au système CATS (Conversation Analysis Transcription System) initié par G. Jefferson —, dans la transcription de ces particules, à l'interface du vocal et du verbal. On vise à rendre compte des inférences formatant l'activité interprétative des interactants. En appui sur ces conventions de transcription, l'étude systématique des Voc, sur un corpus de langue parlée (notamment en situation d'interaction de service) permet de dessiner un peu plus précisément le contour sémantico-pragmatique des Voc comme unités interactionnelles.

**Transcribing vocal datas : are the interjections prosodic correlates noteworthy ?**

*Transcribing the French primary interjections ah, oh, ouh (foreward : Voc) infers to analyze them as discourse and interactional units. Intonative cues leads to their perceptual interpretation. But these one often seems quite ambiguous when one reads transcripts. We can postulate that suprasegmental features "actualize" Voc in discourse. In so far, these characteristics partly define discursive meaning effects of the Voc in conversational context. Conversely, these particles vocal combinations (and notably their preference system) reveal more accurately the vocalisations semantic schematic value. This leads to discuss a number of options referred to the Jefferson's CATS, ordering to transcribe Voc particles, that stand up between vocal and verbal levels. Our own methodological decisions aim to show how the interjections affect the participants interpretative activity. Several examples of vocalisations, drawn of service encounter spoken datas, are analysed in order to display how transcription's conventions help to define semantical and pragmatival meaning of these primary interjections as interactional units.*