

Articulations, ajustages : intonation expressive et système phonologique

Chantal Rittaud-Hutinet*

Université Sorbonne-Paris 3, 75231 Paris cedex 05 France

Résumé. Recourir au non-dit permet d'en dire plus qu'avec les mots, et souvent de se faire comprendre bien plus clairement qu'avec ces derniers, notamment parce que le sous-entendu, réalisé grâce à l'intonation expressive : échappe à un grand nombre des contraintes sociétales et relationnelles auxquelles la dimension verbale de la langue est soumise ; est à la fois rapide, discret ; dégage un signifié pragmatique riche et subtil. On nomme signes vocaux ces ensembles signifiants-signifiés opérant dans l'implicite de l'énonciation et émis volontairement. À partir d'une approche empirique, cet article : décrit les critères à respecter pour asserter que ces objets sonores correspondent à des unités linguistiques au même titre que les mots : expose les indices acoustiques de leurs signifiants (nature, usage) ; analyse leurs extensions ; sonde leur degré d'autonomie ou de dépendance face aux contraintes des composantes phonologique, phonétique et intonosyntaxique ; évalue l'importance du déploiement de leurs signifiés par rapport au discours lexical sur lequel s'actualisent leurs signifiants.

Abstract. Articulation between French prosodic cues and phonemic level. In our speeches we by-pass easily linguistic taboos with the implicit. Thanks to the discretion of the so called vocal signs, the speaker can transmit ideas, opinions, etc. It explains why the significant prosodic units holds a so important place among the resources of multimodality. Putting aside uncontrolled emotional phenomena, the paper investigates some of the ways the speaker manages to "say without saying" by voluntarily encoding opinions, emotions, sensations, judgments etc. through these prosodic cues superimposed on the verbal material expressed. The target is to point out: what they are made of (rules and the functions); what are their meanings; the relations between the phonemic, phonetic and phonopragmatic levels (combination, complementarity); how and in what directions their pragmatical effect works.

Introduction

Dans les échanges oraux de type conversation, ce que chacun cherche à transmettre à ses interlocuteurs, bien plus que des informations nouvelles, c'est tout ce qui l'a conduit à s'exprimer sur un sujet, et si possible à leur faire partager son point de vue. En même temps,

* chrit@wanadoo.fr

très souvent il veut prendre des précautions suffisantes pour être à l'abri d'un jugement ou de réactions négatives – toujours possibles. D'où la fréquence du recours aux intonations expressives [cf., entre autres, 1], qui permettent de demeurer dans l'implicite tout en faisant comprendre son intention réelle.

Au sein de la multimodalité de l'oral [cf. 2], les avantages de leur utilisation s'expliquent facilement : cette dimension du vouloir dire porte les allusions, arrière-pensées, insinuations, double sens et autres sous-entendus absents du sens dénoté des phrases articulées. De plus, notamment en cas de tension, ou quand les interlocuteurs marquent un mécontentement, l'énonciateur a toujours la possibilité de faire l'étonné, de nier avoir voulu produire le signifié incriminé. De ce fait, comprendre l'intonation expressive est essentiel à l'appréhension par les récepteurs de la totalité du message émis, c'est-à-dire y compris le non-dit de l'énoncé [cf. 3, 4, 5].

Soyons clair : toutes les variations acoustiques apparaissant dans le discours d'un énonciateur n'ont pas automatiquement une signification, au sens linguistique du terme : les instants où surgissent des marques sonores (associées ou non à des signaux visuels) trahissant de manière irrépressible des émotions, des inhibitions, des affects, n'en font pas partie. Celles-là, l'énonciateur ne les contrôle pas, on parle à leur sujet d'« émotions brutes », car elles traduisent des états d'être. Bien sûr elles donnent elles aussi lieu à interprétation, mais dans un autre cadre. Aussi, même si les formes sonores peuvent à l'occasion avoir les mêmes déclencheurs, seuls les signes émis volontairement et opérant dans le sous-texte peuvent prétendre à être des unités linguistiques au même titre que les mots.

On parle donc d'intonation expressive exclusivement quand l'énonciateur veut expliquer ses opinions, ses états d'âme, ou plutôt les commenter, et le faire au travers d'un méta-discours non-verbal [cf. 6]. « Intonation expressive » ne désigne ainsi que la part du sens qui dépend uniquement du vouloir transmettre, lorsqu'il se trouve hors des mots du discours de celui qui parle.

Après un bref rappel des critères qui autorisent à répertorier ces objets sonores comme des signes vocaux (ou prosodèmes) de l'intonation expressive, de leur nature et de leurs usages, nous détaillerons les résultats obtenus par l'analyse perceptuelle des données (méthode empirique) [cf. 7], concernant les outils acoustiques de la langue qu'ils mettent en œuvre, l'objectif étant de déterminer la façon dont ils opèrent par rapport à ceux qui sont exploités dans les composantes phonologique, combinatoire et intonosyntaxique. Puis nous examinerons quelques exemples pour mettre en lumière les relations qui s'établissent entre le signifiant du signe vocal et le signifié dégagé.

1. La notion de signe vocal en intonation expressive

1.1 Quand peut-on parler de signe vocal ?

Dans l'approche sémasiologique, on affirme la présence dans la parole d'unités linguistiques sonores – donc codées – quand l'analyse montre qu'à un signifié pragmatique est associé de façon régulière le même signifiant acoustique ; donc quand les conditions suivantes sont respectées :

- à une même combinaison d'indices sonores correspond toujours un même contenu de sens ;
- il faut et il suffit que cette partie phonique de la parole soit présente [cf. 8] pour que le destinataire soit en mesure de traduire le sens du signifié pragmatique qu'elle donne à l'énoncé, en en modifiant le sens dénoté (et ce même si elle peut également contribuer à la redondance, toujours utile à l'oral) ;
- l'utilisation de ces unités linguistiques sonores procède de la volonté du parleur ;

- elles ont des fonctions dans les plans discursif, dialogal, relationnel, conversationnel et interactionnel et leur signifié s'applique dans un de ces plans ou dans plusieurs ;
- leur effet peut s'étendre aux paroles qui les suivent ou à celles qui les précèdent.

1.2 Avantages et inconvénients

La présence de la prosodie signifiante explique pourquoi il arrive à l'allocutaire de s'insurger, de manifester son enthousiasme ou son doute, de l'ironie ou de l'impatience, alors que la partie explicite du discours de l'énonciateur était informative.

D'où les avantages que trouve l'énonciateur à en utiliser les signes :

- Ils permettent d'établir un climat de conciliation, de sympathie, de consensus, de connivence.
- Ils ont un rendement pluriel, car les signifiés entrent souvent dans plusieurs plans d'application.
- Ils ont pour eux la rapidité, car ils ne nécessitent aucun temps supplémentaire d'élocution, leur signifiant s'actualisant directement sur les mots.
- En cas de conflit entre les interlocuteurs, il est ardu pour le destinataire de les réfuter, car il lui est difficile d'en décrire les signes acoustiques, ou de les "copier", leur émetteur pouvant toujours prétendre qu'il s'agit d'un malentendu.
- Leur commodité se manifeste aussi dans le fait que, quand on veut verbaliser leur signifié, une longue périphrase est souvent nécessaire ; et on arrive rarement à être réellement satisfait du résultat.

Mais ces signes vocaux présentent des inconvénients potentiels :

- À l'instar de ce qui arrive pour les mots :
 - * ils peuvent être mal encodés (erreurs par distraction, étourderie, ignorance, inattention, effort inabouti) ;
 - * ils peuvent se révéler être un choix erroné par rapport au sens pragmatique visé.
- Leur efficacité dépend de la bonne volonté interprétative du destinataire, qui peut toujours faire comme s'il ne les avait pas entendus.
- De même que les autres signes linguistiques, ils peuvent être source de mésententes entre les sujets parlants, d'ambiguïtés et d'équivoques, d'oppositions.

2. Le signifiant

2.1 Les traits

Certains traits sont plus saillants que d'autres, aux oreilles des récepteurs non-spécialistes. En effet, ces derniers ne repèrent habituellement que ce qui concerne la hauteur tonale, et/ou la durée, et/ou l'intensité. Ainsi dans cet exemple, de 1662, que nous livrent Antoine Arnaud et Pierre Nicole dans *La Logique ou l'art de penser* [9] :

(1) « Personne ne trouveroit bon qu'un laquais qu'on reprend un peu fortement, répondît :

Monsieur, parlez plus bas, je vous entends bien : parce que le ton fait partie de la réprimande, et est nécessaire pour former dans l'esprit l'idée que l'on veut y imprimer. »ⁱ

Certains indices sonores entrent dans l'ordre du segmental, les autres dans celui du supra-segmental [cf. 10]. Quels sont-ils et de quelles façons se manifestent-ils ?

– La labialité de type vocalique est évalué par rapport à l'opposition phonologique : labial (lèvres arrondies par projection vers l'avant en même temps que commissures rapprochées l'une de l'autre) vs. non-labial (lèvres écartées par éloignement des commissures vers les oreilles en même temps que plaquées contre les dents).

Comme trait de signifiant de signe vocal, c'est la partie arrondissement-écartement de l'articulation qui est concernée : elle est soit « forcée », par un rapprochement partiel des

commissures, soit « contrariée », par un éloignement partiel des commissures. La partie projection des lèvres vs. placage contre les dents ne change pas : elle est toujours phonologiquement distinctive.

– Le voisement est mesuré par rapport à l'opposition phonologique voisé vs. non-voisé. Deux cas de figure : les sons correspondant aux phonèmes non-voisés sont partiellement voisés ; ceux qui correspondent aux phonèmes voisés sont partiellement dévoisés. Ils vont toujours de pair.

Mais ce trait est lié à la tension articulatoire. Il ne donc peut être pris en compte que quand il se trouve ailleurs qu'en fin d'énoncé (le schéma intonatif conclusif prédisposant à la diminution du voisement), ou qu'il n'est pas adjacent à un segment pouvant déclencher l'application de l'assimilation de sonorité.

– La quantité vocalique relève de la phonétique combinatoire. L'orthoépique dit que la voyelle est allongée dans les conditions suivantes : en raison de sa position dans le groupe accentuel – en syllabe tonique – ; en raison de la structure de la syllabe dans laquelle elle se trouve – fermée par une des consonnes [R z ʒ v] – ; et/ou parce que le timbre de la voyelle elle-même est [e ø o ɛ ɔ ɔ̃].

Par rapport aux durées de base du locuteur pour ses [V] et [V:] "canoniques" – c'est-à-dire respectant les règles de la combinatoire –, cinq durées sont actuellement répertoriés : abrégée [V̇] ; un peu allongée [V.] ; très longue [V::] ; supra-longue [V̂] ; auxquelles il faut ajouter [V] et [V:] quand la combinatoire demandait la quantité inverse.

– La prééminence est réglée par l'intonosyntaxe. Elle fonctionne comme indice quand, au lieu d'avoir un accent tonique de groupe accentuel, la dernière syllabe du groupe est atone.

– Le schéma intonatif de groupe accentuel est largement déterminé par l'intonosyntaxe [cf. 11] (pour le Fo syllabique, voir infra : hauteur tonale).

Il n'entre dans les constituants de signe vocal que pour autant que le contour [cf. 12] se réalise dans des positions non prévues en morpho-syntaxe. Trois cas sont élicités : schéma intonatif syntaxique conclusif sur un énoncé n'est pas terminé ; schéma bémolisé pour un schéma intonatif syntaxique continuatif ; schéma bémolisé pour un schéma intonatif syntaxique énumératif.

– La pause hors respiration, qui dépend de la morpho-syntaxe de la phrase – elle est en principe automatique par exemple pour : adjectif en apposition, proposition relative explicative. Elle n'agit comme indice de signifiant de signe vocal que quand elle n'est pas prévisible. Par rapport aux autres indices sonores du signifiant auquel elle est associée, elle peut être : initiale, ou interne. Dans ce dernier cas : soit elle intervient entre deux des groupes accentuels faisant partie du signifiant ; soit elle se trouve entre des syllabes marquées par les indices du signifiant.

Remarque : il semble ne pas y avoir une véritable régularité dans la réalisation des durées de la pause.

Les traits suivants ne font pas partie de la description acoustique du français.

– La hauteur tonale (Fo) syllabique (pour le Fo de groupe accentuel, voir supra : schéma intonatif syntaxique). On repère des Fo plats et des Fo modulés.

* Pour les Fo plats, neuf niveaux sont identifiés : supra-haut (SH) (= hors de la gamme courante du locuteur) ; très haut, (H⁺) ; haut (H) ; infra-haut (IH) ; moyen-haut (MH) ; moyen (M) ; moyen-bas (MB) ; bas (B) ; infra-bas (IB) (= hors de la gamme courante du locuteur).

* On nomme Fo modulés ceux qui montrent des changements intra-syllabiques. Par rapport à la gamme des Fo plats, ils partent de différentes hauteurs. Trois sortes sont actuellement élicitées : modulé simple montant ; modulé simple descendant ; modulé complexe montant-descendant.

Remarque : il est possible qu'il y ait des Fo modulés complexes descendants-montants.

– L'intensité. La perception de changement affecte une syllabe, ou un mot, ou plusieurs, ou un groupe accentuel entier. Pour certains signes vocaux, elle présente plusieurs degrés au

cours du passage marqué. On en distingue trois degrés : plus faible (dB⁻) ; plus forte (dB⁺) ; très forte à extrêmement forte (dB⁺⁺).

– Le débit. Il ne peut être considéré comme faisant partie des indices de signifiant que lorsque deux syllabes au minimum sont réalisées plus vite ou plus lentement que les autres. Est exclu aussi le ralentissement lié à l'allongement combinatoire de la voyelle de la dernière syllabe de groupe accentuel (voir supra quantité vocalique).

Deux possibilités : plus rapide, plus lent.

– Les consonnes épenthétiques. Trois se rencontrent dans les corpus analysés : la non-occlusive glottale non-voisée [h] ; la non-occlusive glottale voisée [ɦ] ; l'occlusive glottale non-voisée [ʔ].

Nous reviendrons (sous 3) sur les traits des signifiants vocaux qui ont un correspondant au plan phonologique, pour les comparer à ceux de celui-ci afin d'en vérifier et d'en mesurer les degrés d'efficacité.

2.2 Extension formelle, distribution et lieux d'actualisation

- Chaque indice acoustique de signifiant de signe vocal peut entrer dans plusieurs combinaisons et peut donc faire partie de la composition d'autant de signes vocaux. C'est le cas pour la très grande majorité d'entre eux.

- Certains indices acoustiques sont co-occurrents. D'autres apparaissent sur des syllabes différentes. Pour d'autres signifiants encore, ils sont en partie concomitants et en partie successifs, comme nous l'observerons dans l'exemple (2) (infra, sous 2.3).

- Le signifiant est matérialisé sur une syllabe, ou sur plusieurs, ou sur un groupe accentuel, ou sur plusieurs. Il s'avère que ce déploiement syntagmatique n'est conditionné ni par le découpage des unités lexicales, ni par la structuration syntaxique des énoncés. C'est pourquoi les signes vocaux sont à considérer comme des auto-segments.

2.3 Traits et indices

Le nombre de types acoustiques de traits va de un à cinq. Mais nous l'avons vu, certains types de traits ont deux termes, d'autres plus, et par ailleurs plusieurs syllabes peuvent être marquées par des traits. Il n'est donc pas rare que, par rapport au nombre de ses traits, le nombre d'indices d'un signifiant soit plus élevé. C'est notamment le cas pour les hauteurs tonales et pour l'intensité.

Un exemple marquant est celui du signe vocal abrégé en <Vi> : ses indices s'étendent sur tout l'énoncé marqué par le signifié. Ainsi, prenons la phrase :

(2) Le cortège officiel arrive dans une demie heure.

Dite comme une simple information, elle est réalisée avec les schémas intonatifs syntaxiques fondamentaux de l'énoncé déclaratif : continuatif + conclusif. Mais elle peut être marquée par exemple par le signe vocal <Vi> ; on traduirait alors son signifié à peu près par : « Dépêche-toi !! Plus Vite !! Je suis très en colère, tu es si lent à agir !! ». C'est-à-dire qu'avec ce signe vocal :

- › l'énonciateur (A) donne à son interlocuteur (B) l'ordre d'agir sur-le-champ et de se dépêcher – ce qui le met en situation de supériorité – ;
- › A fait reproche à B de sa lenteur, il est en colère envers B et même à la limite de l'explosion. À l'écrit, la phrase se terminerait par deux points d'exclamation ;
- › B est sommé de répondre par un faire.
- › avec ses mots, A n'exprime que la cause – l'équivalent de « parce que » – ; il sous-entend la conséquence – c'est-à-dire l'ordre, lequel renvoie soit à la situation de l'échange, soit à un énoncé antérieur d'un des interlocuteurs.

Le signe vocal <Vi> contient soit trois, soit quatre types de traits : intensité, hauteur tonale, quantité vocalique, et dans certaines conditions, consonne épenthétique [ʔ]. Mais il actualise huit ou neuf indices acoustiques, répartis de la façon suivante :

- la partie précédant le dernier groupe accentuel fonctionne comme une sorte d'amorce avec, pour chacun (quel qu'en soit le nombre) :
 - . dB⁺
 - . Fo plat IH sur chaque syllabe tonique ;
- pour le dernier groupe accentuel, on a :
 - * antépénultième (ou anté-antépénultième et antépénultième quand l'avant-dernier mot n'est pas monosyllabique) :
 - . dB⁺⁺
 - . Fo plat H (ou même SH)
 - . et, si cette syllabe commence par une voyelle, un [ʔ] épenthétique qui fait office de consonne initiale
 - * pénultième : Fo plat MB
 - * dernière syllabe :
 - . dB⁺⁺⁺
 - . Fo modulé simple descendant H⁺>MH (ou même SH>H)
 - . [V:]ⁱⁱ.

De telles configurations sont rares, mais témoignent de la complexité acoustique de certains signifiants qui, pourtant, ne posent aucun problème de diction à leur émetteur.

3. Phonologie et signes vocaux

Quel est le degré d'autonomie et de performance des traits des signifiants de signes vocaux quand ils sont déjà utilisés par les composantes phonologique, combinatoire et intonosyntaxique ? Autrement dit, quelles formes prennent-ils et quel rendement ont-ils par rapport aux contraintes et limitations générées par les autres composantes ?

3.1 Termes du trait et oreille

Pour une partie des traits des signifiants des signes vocaux, l'opposition n'est pas binaire, nous l'avons vu. Pour ceux-là, l'interprétation de leur fonction ne se réalise correctement que sur la base de ce que l'oreille linguistique du récepteur B a déjà "appris" des spécificités phoniques de la personne qui parle. Aussi, quand c'est sa première écoute de tel locuteur A, il faut un peu de temps à B pour établir l'échelle dans laquelle les indices qu'il entend se situent, à la fois entre eux et par rapport à leur usage non marqué chez A. C'est seulement ainsi qu'il sera capable de reconnaître à partir de quand ils sont significatifs chez A. Pour savoir en prendre la mesure, B doit donc avoir entendu plusieurs réalisations de ce trait pour pouvoir comparer, d'une occurrence à l'autre, les différences acoustiques de ses termes chez A. C'est seulement alors qu'il estime avec justesse la valeur de leur amplitude.

Voyons par exemple le trait de hauteur tonale : selon la tessiture de celui qui parle, le niveau du Fo : d'une part peut se situer plus ou moins haut dans l'échelle des Hz, et de l'autre peut couvrir un éventail d'Hz plus ou moins grand.

Le tableau en (3) compare la tessiture de sept locuteurs vis-à-vis de la localisation et de la gamme de leurs Fo :

On y observe que :

- › la tessiture de C est celle qui a l'amplitude maximale au sein de cette échelle ;
- › D et E ont la même amplitude, la plus petite des sept locuteurs ; D dans la partie haute et E dans la tranche moyenne ;

- › F montre une amplitude intermédiaire par rapport à celle des autres locuteurs, dans une partie assez basse de l'échelle ;
- › G a une amplitude un peu plus étroite que celle de F, et elle est située dans la partie haute ;
- › l'amplitude de H est à peine plus grande que celle de D et de E, un peu plus petite que celle de G, et ce dans une partie plutôt basse ;
- › J montre une amplitude plus petite que celle de C mais plus grande que celle de F, et ce dans la tranche moyenne.

(3)

	C	D	E	F	G	H	J
é	—				—		
c		≡			—		
h	—	≡≡≡			—		—
e		≡≡≡			—		—
l	—	≡≡≡		—	—		—
l	—		≡	—	—		—
e			≡≡≡	—	—		—
	—		≡≡≡	—		—	—
d	—		≡	—		—	—
e	—		—	—		—	—
s	—		—	—		—	—
	—		—	—		—	—
Hz	—						

Par rapport à l'exemple proposé dans ce tableau, il est nécessaire que B mémorise les particularités individuelles de chacun des sept locuteurs pour être en mesure d'apprécier si et dans quelles conditions la présence d'une variation de Fo participe chez chacun d'entre eux au signifiant d'un signe vocal.

3.2 Degré de rendement

L'effet des indices acoustiques spécifiques à l'intonation expressive semble automatiquement maximal, c'est le cas en (4) avec le signe vocal abrégé en <E> :

(4) Situation : au cours d'un sketch, l'humoriste Danny Boon chante le début de plusieurs chansons, et abandonne rapidement. Il tourne alors une nouvelle page de son cahier de partitions :

DB- \...\ 'hé 'oh c'est ai'gu 'ça hè / ((il chante)) « Hé'lène je m'appelle Hé'lène » / oh la 'la / ((il arrache la feuille et la jette loin de lui, simule un « chut », fait un geste de la main pour tenter de stopper les réactions du public)) oh 'non a'rrête / 'ah / ((il se gratte la tête)) arrête / ((il fait encore des mimiques gênées)) arrê'tez oh mais 'èh: / elle est sym'pa / sa 'mère /

le public- ((rires, sifflets, cris de joie pendant toute la séquence))

Avec <E>, on assiste à une sorte de traîtrise de A vis-à-vis de B : il l'entraîne à un contresens, pour ensuite le désabuser. En effet, ici, Dany Boon : pousse d'abord le public à une interprétation *a* de « elle est sympa » puis l'oblige dans un deuxième temps à lui donner un sens *b* qui le contredit totalement. Chaque spectateur effectue une première interprétation de cet énoncé, puis est obligé d'en faire une seconde analyse, ce qui l'amène à considérer : « elle est sympa », non pas comme un énoncé achevé – bien qu'il soit prononcé avec le schéma intonatif syntaxique conclusif –, mais comme le début d'un ensemble qui va jusqu'à : « sa mère ».

Pour arriver à ce résultat, Dany Boon détourne trois fonctionnements grammaticaux, ce que seule la présence de <E> rend possible :

- › il présente : « elle est sympa », comme un énoncé achevé – le schéma intonatif syntaxique conclusif, suivi de pause, ne s'utilisant "normalement" que pour une fin de phrase ou une fin de tour de parole – ;
- › puis il prononce : « sa mère », avec un schéma intonatif syntaxique incident, bien que ce type de séquenciation soit "interdit", parce que le schéma incident se trouverait en liaison avec une fin de phrase, donc réalisée avec un schéma conclusif ;
- › et il se permet un changement de sujet grammatical : « elle », renvoie d'abord à : « Hélène », et ensuite à : « sa mère ». Cette dernière devient alors la personne visée par l'adjectif « sympa » lequel, dans la première analyse, qualifiait : « Hélène ».

Le signifiant qui permet de dégager un tel signifié se développe sur deux groupes accentuels et combine les indices suivants :

- pause initiale (si <E> n'est pas en début de réplique)
- un groupe accentuel à schéma intonatif conclusif
- pause interne longue
- puis
- un groupe accentuel à schéma intonatif incident, et prononcé avec
 - * dB-
 - * dernière syllabe atone.

Mais les traits du signifiant d'un signe vocal peuvent faire déjà partie des couches phonologique, combinatoire, intonosyntaxique. Dans ce cas, jusqu'à quel point ont-ils une capacité différenciatrice [cf. 13]?

– Labialité de type vocalique. Qu'elle soit forcée ou contrariée, l'indice est toujours très efficace, car il ne concerne qu'une des deux parties du geste articulatoire complexe nécessaire au trait phonologique. De plus, elle est également visible : l'énonciateur sourit (labialité contrariée) ou fait la moue (labialité forcée).

– Voisement partiel, dévoisement partiel. L'action de l'indice est relativement faible. L'impact du dévoisement partiel est plus grand que celui du voisement partiel, car il affecte les consonnes ET les voyelles. Mais en tout état de cause, son audibilité est d'autant meilleure que : pour le dévoisement partiel, la séquence est riche en unités segmentales phonologiquement voisées ; pour le voisement partiel, la séquence est riche en unités segmentales phonologiquement non-voisées.

– Quantités vocaliques : [Ṽ], [V], [V.], [V:], [V::], [V̂]. Nous l'avons dit supra, elle est remarquable, donc significative, quand sa durée n'était pas prévue et donc contrevient à l'une des règles combinatoires concernant l'allongement des voyelles. Dès lors, quand l'allongement de la voyelle est déjà requis dans la même position par la combinatoire, la durée longue n'a aucune audibilité : il doit y avoir d'autres indices pour que le signe vocal prenne sens pour le récepteur.

– Plus surprenant, l'intensité, car elle n'a pas de fonction dans les autres composantes, en français : dB⁻, dB⁺, dB⁺⁺.

S'agissant de l'augmentation, l'importance de son rendement sur la consonne initiale de la syllabe marquée est adossée au mode d'articulation de cette consonne : c'est plus percutant avec les occlusives, et plus encore avec les occlusives non-voisées qu'avec les voisées. Et donc d'autant moins avec les non-occlusives non-voisées, et encore moins avec les non-occlusives voisées.

L'intensité fait partie des indices constitutifs de <Vi> (voir supra, sous 2.3, exemple (2), ou du signe vocal abrégé en <H>, comme en (5) (6) (7) :

(5) D- \...\ au contraire c'est très co'mmode / moi j'aime 'bien les gros porte-'clés parce que /
 tu peux j(e)ter tes clés au fond d(e) ton 'sac et:: même à mi'nuit / même bou'rré /
 <H> sur verbe _____ <H> sur 2 adverbes _____
 tu les r(e)trouves facil(e)'ment /

C- et S- ((perplexité, puis rire))ⁱⁱⁱ

(6) H- \...\ et puis j'e'ssaie surtout de 'voir de 'voir le 'prix parce que j(e) su(is) tout à fait 'con:tre de mettre des sommes euh:: / extravagantes dans des vêt(e)'ments /^v [14]

EF de <H>
AP de <H>

(7) F- \...\ mais:: j(e) 'crois qu(e) les gar'çons fon(t) extrêm(e)ment:: ?attention à la façon

EF de <H>
AP de <H>

dont i(ls) sont habi'llés et au'ssi qu'ils suivent la la 'mo.de / enfin:: j(e) sais pas y a:: \...\

Décrit dans les grammaires comme un « accent d'insistance », ou comme un « accent de force », nous avons avec <H> l'équivalent sonore d'un intensif : au sein d'une séquence, il focalise l'attention de B sur le mot que A considère comme important, quelle qu'en soit la raison. Ce signe vocal marque indifféremment un nom, un adverbe, un adjectif, un article, un verbe, un numéral, etc. et donne valeur de superlatif aux adjectifs qualificatifs axiologiques. Des équivalents lexicaux en sont, par exemple : « hyper, énorme, génial ! ».

Au sein des indices du signifiant de <H> (voir liste complète infra, sous 4.b), l'augmentation d'intensité s'entend mieux sur le -t- de : « extravagante », et sur le -?- de : « ?attention », que sur le -j- de : « jeter », et sur le -m- de : « même ».

Rappelons enfin que, comme plusieurs indices acoustiques concourent au signifiant d'un signe vocal, la faiblesse relative de l'un est compensée par la force des autres.

4. Relations forme-sens

Nous avons vu (supra, sous 2.2) les diverses dimensions syntagmatiques possibles des indices acoustiques d'un signe vocal. Examinons à présent les relations qui s'établissent entre son extension formelle (EF) et celle du sens pragmatique^{vi} qu'elle dégage. Quelle est l'étendue et quelles sont les limites de l'application – le champ d'action – du signe vocal ? Et quelles correspondances y a-t-il entre forme et sens [cf. 14] ?

Comme nous allons le constater, EF et application pragmatique (AP) ne correspondent pas toujours à la même portion d'énoncé. On observe des coïncidences totales, mais aussi partielles.

a) Très souvent il y a égalité entre les mots ou les syntagmes marqués par le signifié et l'actualisation du signifiant.

C'est ce qui se passe avec le signe vocal abrégé en <QRéf>. Avec ce signe vocal, l'énonciateur exprime une question préalable qui remet en cause la validité de la proposition qui vient d'être faite par l'interlocuteur, ou le statut que celui-ci prenait implicitement pour se permettre d'avancer la position qu'elle contenait. <QRéf> est donc utilisé pour contester ce qui vient d'être dit. L'énoncé, le plus souvent très bref, contient toujours un terme interrogatif. Question et réfutation sont indissociables ; mais la contradiction surpasse la question, qui en devient une sorte de simple suite logique [cf. 15].

En voici deux exemples, en (8) et en (9) :

(8) H- Alors j[e] vais t[e] poser une première ques'tion qu'est-ce que tu 'pen:ses / de la 'mode ac'tuelle /

F- Quelle 'mode / ((rire))
_____ <QRéf>

H- Euh:: ben la mode vestimen'telle: / vestimen'taire plu'tôt / sur euh:: la façon d(e) s'habiller des 'jeu:nes / ((euh ben)) des fi:lles / p(ar)ce que \...\^{vii}

Dans cet exemple, la question préalable de F oblige H à repartir à zéro : H s'expliquera et tentera plusieurs fois de s'excuser pour retrouver un consensus avec F.



Les deux phrases de cette affiche établissent un jeu avec les passants. Car pour comprendre ce que veut effectivement dire cette publicité, le sens dans lequel elle veut les pousser, ils doivent mentalement imaginer un dialogue, et opérer un choix de signe vocal pour : « Quelle vitre ? », plusieurs significations étant théoriquement possibles, par exemple : « De laquelle est-ce que vous parlez ? ».

Ici, la "bonne" interprétation pour ce qui a l'apparence d'une question se paraphrase : « Où y a-t-il une vitre ? Pourquoi parlez-vous de vitre ? Je ne vois pas de vitre, je vois seulement le ciel, quelques petits nuages et un flacon. ». Et elle conduit les lecteurs à conclure qu'on veut les amener à admettre ce que l'on pourrait traduire par : « "Ajax vitres" est tellement efficace comme produit de nettoyage que vous n'avez même vu pas qu'il y a une vitre sur cette photo et que voyez ce que vous voyez au travers de cette vitre ».

Les indices acoustiques de la question préalable transmise avec <QRéf> sont :

- sur le terme interrogatif :

* Fo très élevé : de H à SH

* dB⁺

- la suite de l'énoncé – toujours très courte – est en plage basse, mais jamais aussi basse qu'un groupe syntaxique conclusif car il demeure au niveau MB.

b) Le signifié s'étend en aval du signifiant (EF plus petite qu'AP). Les limites de l'AP sont faciles à tracer quand le signifiant ne s'actualise que sur un mot.

C'est le cas pour le signe vocal abrégé en <H> des exemples (5) (6) (7) (supra, sous 3.2) : seule la première syllabe de type [CV] du mot marqué voit l'actualisation de son signifiant, mais étant donné le signifié, on déduit facilement que son AP est égale au mot^{ix}.

Dans le cas général (la syllabe est de type [CV]), comme en (5), les indices de <H> sont les suivants :

- dB⁺

- [V], quel que soit le type syllabique ;

- Fo plat H à SH.

Mais la première syllabe peut être de type [V]. Là, deux variantes se font jour :

- soit l'ensemble des traits est déplacé à droite et s'applique sur la première syllabe de type [CV] du mot, comme en (6) ;

- soit – ce qui est beaucoup plus courant aujourd'hui – à l'initiale de la première syllabe est actualisée une consonne épenthétique, l'occlusive glottale non-voisée [ʔ], ce qui, en donnant : [ʔV], comme en (7), redonne à la syllabe le type [CV].

Les autres traits demeurent alors sur la première syllabe.

c) Il est plus souvent difficile de préciser jusqu'où s'étend l'AP par rapport à l'EF.

Ainsi pour le signe vocal abrégé en <Rh0>. Avec ce signe vocal, A demande à B un accord avec sa propre position, mais ne veut pas en avoir l'air. <Rh0> commuant un énoncé grammaticalement déclaratif en question exigeant une réponse, grâce à lui, A peut interpellé discrètement B et faire en sorte de l'orienter vers la réponse qu'il désire entendre – l'équivalent d'un « non » si l'énoncé est une affirmation ; d'un « oui » si l'énoncé est une dénégation. Donc, alors que A fait mine de questionner B, en réalité il porte un jugement de valeur favorable vis-à-vis de l'objet de son propre discours, et fait comprendre à B qu'il refuse de croire à la véracité de l'information que ce dernier lui a délivrée ; conséquence, A veut que B revienne sur son affirmation car pour A, à l'évidence, elle est fautive. De plus, A veut de B une réponse brève, c'est-à-dire seulement une réplique, ce qui signifie que, dès qu'il l'aura obtenue, il entend continuer son tour de parole.

C'est ce qu'on trouve en (10) :

(10) F- Je crois au'ssi qu(e) ça donne une in? des indica'tions mais:: / juste'ment / le dan'ger
c'est de:: / de classifier les gens comme 'ça /

EF de <Rh0>

H- On a trop l'habi'tude de de classer les 'gens:: de fa? de par la fa'çon \...\^x

Le signifiant de <Rh0> est constitué des éléments suivants :

- sur tout le groupe accentuel – en (10) « de classifier les gens comme ça » – :
 - * labialité de type vocalique contrariée (donc, que les segments soient ou non phonologiquement labials)
 - * dévoisement partiel des sons phonologiquement voisés
 - * voisement partiel des sons phonologiquement non-voisés
 - * débit ralenti
- sur les trois dernières syllabes du groupe, Fo respectivement : B, MB, MH.

Mais à quelle portion d'énoncé correspond l'AP de <Rh0> vers la gauche, étant donné que son EF ne couvre que le dernier groupe accentuel ? Faut-il remonter jusqu'à « le danger » ? ou plus loin ?

Et là, exceptionnellement, l'écrit a un avantage sur l'oral, car la description écrite a l'avantage de ne pas poser au chercheur la question de la relation entre l'EF et l'AP du signe vocal. quand le narrateur explique la façon dont le personnage a prononcé une réplique, ses commentaires affectent automatiquement l'ensemble de l'énoncé (sauf s'il en limite expressément la portée), comme en (11) :

(11) « “Vous avez reçu une lettre anonyme ?” Elle affirmait en questionnant, avec un faux air d'être certaine de ce qu'elle avançait »^{xi}

Pour nous faire comprendre le message réel de la femme, le narrateur glose^{xiii} le "ton" sur lequel elle a dit : « Vous avez reçu une lettre anonyme ? », pour lui donner un signifié pragmatique différent du sens transmis par l'apparence interrogative de sa phrase.

5. Bilan et éléments de conclusion

Les indices acoustiques de signes vocaux qui se servent de traits absents des plans phonologique, combinatoire ou intonosyntaxique se développent sans contrainte (nous avons vu que le nombre de formes pour un trait peut monter jusqu'à neuf).

En revanche, ceux qui utilisent des traits déjà présents dans ces composantes sont souvent pré-orientés : ils font appel aux moyens que celles-ci laissent inutilisés.

En toute logique, l'application d'un trait acoustique de signe vocal se fait après celle de ceux des couches phonologique, combinatoire ou intonosyntaxique. Par suite, pour ceux dont les termes sont les mêmes au plan de l'intonation expressive, quand l'application d'une des règles précédentes commande leur présence au même endroit du discours, la capacité différenciatrice de l'indice de signe vocal est nulle, car : soit le terme est le même, auquel cas sa fonction est déjà déterminée par l'une des couches précédentes ; soit c'est le terme opposé, auquel cas il ne peut tout simplement pas être réalisé (conflit de programmation).

On dira donc que l'efficacité propre des indices de traits de l'intonation expressive déjà présents dans les autres couches est inversement proportionnelle à la qualité acoustique qu'imposent ces dernières sur le ou les sons du passage concerné.

L'actualisation d'un signifiant de signe vocal n'a d'effet que dans les lieux où un au moins de ses indices constitutifs n'est pas déjà utilisé au même endroit avec une fonction distinctive ou combinatoire ou syntaxique dont le terme est le même, ou lui est opposé. Autrement dit, pour qu'un signifiant soit performant, c'est-à-dire pour que le signifié pragmatique du signe vocal soit perçu, il faut qu'au moins un de ses indices acoustiques ne soit utilisé que par lui là où il est réalisé.

Ainsi, les marques vocales des signes vocaux fonctionnent à la fois en opposition et en synergie avec les autres couches : les impératifs et les contraintes régis par ces couches, mais

aussi leurs laxités et leurs sous-spécifications, tout ce qu'on pourrait nommer leurs "zones interdites" offre à l'intonation expressive ses possibilités en même temps que ses limites. De ce fait, on peut raisonnablement penser qu'à l'occasion l'énonciateur choisira tel signe vocal de préférence à tel autre – quand il a un synonyme –, pour tenir compte de la présence de l'une de ces servitudes.

Rappelons pour conclure que l'énonciateur dispose du lexique et des signes vocaux, qu'il peut les utiliser conjointement ou non, mais que, sans l'intonation expressive, il ne transmettrait que des squelettes d'énoncés, des atomes informatifs. Car c'est avec l'intonation expressive que surgit le vrai message, avec sa finalité et ses arrière-pensées – instigation, jugement, connivence, réserve, suggestion et autres –, qui, lui, transmet à l'interlocuteur ce qu'on dit à propos de ce qu'on dit – opinion, sensation, goût, émotion, jugement, sentiment, etc.

Conventions d'écriture

X-	=	émetteur ayant le tour de parole
(Y-)	=	émetteur en position de récepteur
\...\ Abd ef	=	le tour de parole a commencé avant/se poursuit après le fragment
(bla)	=	la majuscule signale le début d'une prise de parole
(())	=	son/syllabe non prononcé(e)
(())	=	fragment illisible/réaction non-langagière
?	=	consonne occlusive glottale non-voisée
V. V: V:: V̄	=	allongement de voyelle : faible ; normal ; très long ; supra-long
Ṽ	=	abrègement de voyelle
‘	=	accent tonique (groupe accentuel)
/	=	pause
<W>	=	abréviation de signe vocal
_____	=	souscrit : délimitation de signe vocal
[V] [C] [CV]	=	toute voyelle du système phonologique ; toute consonne du système phonologique
[CV]	=	syllabe de type consonne + voyelle

Références bibliographiques

1. Ch. Rittaud-Hutinet (2019). *Intonation expressive et français langue étrangère*. Paris : L'Harmattan
2. O. Siezov, J. Wildfeuer (éds) (2017). *New Studies in Multimodality. Conceptual and Methodological Elaborations*. London: Bloomsbury Publishing
3. R. Carston (2009). The explicit/implicit distinction in pragmatics and the limits of explicit communication. *International review of pragmatics*, 1, 35-62
4. S. Berbinski (éd.) (2016). *Le Dit et le Non-Dit. Langage(s) et traduction*. Bern : Peter Lang
5. G. Planchenault, L. Poljak (eds) (2021). *Pragmatics of Accents*. John Benjamins, Pragmatics & Beyond New Series 327
6. Ch. Rittaud-Hutinet (2014). Sans les mots. *Actes : Autour des formes implicites*. Limoges (à paraître 2022)
7. K. Kosecki, J. Badio (éds) (2015). *Empirical Methods in Language Studies*. Bern : Peter Lang, Lodz Studies in Language, Band 37
8. Ch. Rittaud-Hutinet (2021). Comment comprendre les non-dits ?. In É.Castagne, Ph. Monneret (éds.). *Intercompréhension et analogie*. Bruxelles : de Boeck supérieure, collection Champs linguistiques, 197-215

9. A. Arnaud, P. Nicole ([1662] 1978). *La logique ou l'art de penser*. Paris : Flammarion
10. J. Caspers, Y. Chen, W. Heeren, J. Pacilly (éds) (2014). *Above and Beyond the Segments. Experimental linguistics and phonetics*. Amsterdam : John Benjamins
11. M. Rossi, A. Di Cristo, D. Hirst, Ph. Martin, Y. Nishinuma (1981). *L'intonation. De l'acoustique à la sémantique*. Paris : Klincksieck, études linguistiques XXV
12. M.-A. Morel (1997). Stratégies intonatives et syntaxe du discours dans l'oral spontané en français. *TAL, Actes de la journée d'études ATALA "Prosodie et syntaxe"*. Paris : C.N.R.S., 38-1, 1-19
13. S. Frota, G. Elordieta, P. Prieto (éds) (2011). *Prosodic Categories : Production, Perception and Comprehension*. New York : Springer, Studies in Natural Language and Linguistic Theory Series
14. J. Cosnier, C. Kerbrat(éds) (1987). *Décrire la conversation*. Lyon : Presses universitaires de Lyon
15. F. Liedtke, C. Schulze (éds) (2013). *Beyond Words. Content, context, and inference*. Berlin : De Gruyter Mouton, Mouton Series in Pragmatics 15
16. M. Urgelles-Coll (2010). *The Syntax and Semantics of Discourse Markers*. New York : Continuum International publishing group, Continuum Studies in Theoretical Linguistics Series

ⁱ p. 131

ⁱⁱ quelle que soit la structure de la syllabe : ouverte, fermée par une ou par deux consonnes.

ⁱⁱⁱ conversation entre collègues, à la sortie d'un restaurant, Chambéry 24 mars 2001

^{iv} corpus « Les jeunes et le vêtement » [14]

^v id. *ibid.*

^{vi} Faute de place au sein de ce développement, je laisse de côté le problème de l'orientation pragmatique^{vi} (OP) du signifié, laquelle décrit la direction dans laquelle : l'énonciateur, en utilisant un signe vocal, espère obtenir un résultat chez le protagoniste ; s'effectue le lien de cause à effet direct dont le signe vocal est la conséquence et dont la cause est soit des paroles ou des manifestations non-verbales du destinataire, soit une difficulté d'ordre cognitive chez l'énonciateur.

En effet, l'OP est en partie liée au fait que les signes vocaux ont souvent un effet dans plusieurs plans d'application, dont il faudrait donc d'abord faire la description.

^{vii} corpus « Les jeunes et le vêtement », *op. cit.*

^{viii} reproduction d'une affiche publicitaire vue en 2006

^{ix} On pose ici une convention cellulaire : tout signifié de signe vocal dont les indices acoustiques s'actualisent sur une partie de mot marque le mot entier.

^x corpus « Les jeunes et le vêtement », *op. cit.*

^{xi} Simenon Georges. *L'ombre chinoise*. ([1932] 2003), Omnibus "Tout Simenon" 17, p. 21

^{xii} À l'écrit en effet, un écrivain n'a pas d'autre ressource que de décrire les marques phoniques, et/ou de commenter l'objectif, et/ou de signaler l'effet produit sur le récepteur.