



**HAL**  
open science

# Analyse prosodique de marques de l'impolitesse dans une perspective interactionnelle et didactique

Cécile Couteau-Boddaert

► **To cite this version:**

Cécile Couteau-Boddaert. Analyse prosodique de marques de l'impolitesse dans une perspective interactionnelle et didactique. Sciences de l'Homme et Société. 2014. dumas-01084071

**HAL Id: dumas-01084071**

**<https://dumas.ccsd.cnrs.fr/dumas-01084071>**

Submitted on 18 Nov 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.







## *Remerciements*

Je voudrais ici remercier les nombreuses personnes sans qui ce travail n'aurait pu aboutir.

Il s'agit au premier chef de mes deux directrices, Claudine Moïse et Laura Abou Haidar, pour avoir accepté le sujet de ce mémoire, et m'avoir laissé m'y confronter; pour leur suivi et leurs conseils.

Je remercie aussi Cristelle Cavalla pour son dynamisme, et le soutien qu'elle a su apporter à chacune des étudiantes du tout nouveau "Master pour professionnel en un an".

Un merci particulier à Isabelle Rousset du Lidilem pour ses aides et "tuyaux" divers, à Solange Rossato du Gipsa-Lab, qui a pris sur son temps pour m'initier aux arcanes de Praat, à Roxane Bertrand, du Laboratoire Parole et Langage d'Aix-en-Provence, pour ses conseils bibliographiques et sa gentillesse, à Lu Yan, doctorante au Gipsa-Lab, pour avoir partagé avec moi ses tout premiers articles, à Chara Kornilaki enfin, ma collègue et amie grecque, qui a bien voulu corriger mon travail sur les énoncés en langue grecque.

Je remercie le personnel de la bibliothèque du Département des Sciences du langage pour sa gentillesse et sa disponibilité, ainsi que Manon Carrissimo-Bertola, doctorante au Gipsa-Lab et "gardienne" du centre de documentation du Département Parole et Cognition.

Mes remerciements vont enfin à tous ceux sans la bonne volonté de qui ce travail n'aurait pu même commencer : i) les locuteurs de langues française, grecque et arabe, qui ont accepté de ... parler, pour me fournir un corpus conversationnel qui serait le support de mon expérimentation puis de mes analyses. ii) L'ensemble des testeurs sur les tests perceptifs, pour leur temps et leur souci de bien faire, notamment Caroline, mon cobaye de la première heure.

Un remerciement spécial à Valérie pour son accueil et son amitié, sans lesquels cette année de master aurait été bien différente.



**DECLARATION**

1. Ce travail est le fruit d'un travail personnel et constitue un document original.
2. Je sais que prétendre être l'auteur d'un travail écrit par une autre personne est une pratique sévèrement sanctionnée par la loi.
3. Personne d'autre que moi n'a le droit de faire valoir ce travail, en totalité ou en partie, comme le sien.
4. Les propos repris mot à mot à d'autres auteurs figurent entre guillemets (citations).
5. Les écrits sur lesquels je m'appuie dans ce mémoire sont systématiquement référencés selon un système de renvoi bibliographique clair et précis.

NOM : COUPEAU BODDAERT PRENOM : Cécile

DATE : 05/10/14 SIGNATURE

***MOTS CLES : prosodie, impolitesse, analyse interactionnelle, didactique de la prosodie, interculturel.***

## ***RESUME***

Dans quelle mesure est-il possible de réduire l'apparition de malentendus communicationnels et/ou interculturels par l'apprentissage en langue seconde, ici le français, d'intonations (im)polies?

La recherche menée dans ce mémoire s'inscrit dans un domaine au croisement de la phonétique - plus particulièrement de la prosodie en ce qu'elle s'intéresse aux variations de la fréquence fondamentale - et de la pragmatique et didactique de l'oral - plus précisément de l'expression de la politesse et de l'impolitesse. Sa visée ressortit aussi au domaine de l'interculturel et des contacts de langues, en ce qu'elle pourra constituer le point de départ de recherches portant sur l'élucidation de malentendus possiblement dus à l'intonation, malentendus communicationnels et/ou interculturels entre locuteurs de langues maternelles et locuteurs de langues allophones. Ce mémoire s'appuie sur l'hypothèse qu'il est possible de déterminer un certain nombre de traits ou marques prosodiques intrinsèquement impolis en français. Il s'agira, à partir d'énoncés produits en langue française par des locuteurs natifs du français, de tenter d'identifier et d'isoler des marqueurs acoustiques pertinents de l'impolitesse, pour envisager les retombées éventuelles dans les champs de recherche concernés et faire part de quelques pistes didactiques dans les domaines conjoints de l'intonation et de l'interculturel.

***KEY WORDS : prosody, impoliteness, interactional analysis, didactics of prosody, cross-cultural interactions.***

## ***ABSTRACT***

Is it possible to lower communicative and intercultural misunderstanding by learning impolite intonational patterns in L2 - here French L2?

The field of this essay falls within the crossing of phonetics - and within it of prosody as far as it is concerned with fundamental frequency changes - pragmatics and second languages didactics - especially politeness and impoliteness vocal expression. This research might be a first step towards the elucidation of intonational related misunderstandings between L1 and L2 speakers, and therefore comes under the field of intercultural and languages contacts.

This work is based on the idea that it is possible to identify prosodic characteristics and tokens for genuine impoliteness in French, which was made from ecological conversational sets produced by French native speakers. Possible consequences in the above fields of research are considered as well as a few didactic trails in both intonational and intercultural fields.



# TABLE DES MATIERES

TABLE DES MATIERES	1
TABLE DES TABLEAUX, FIGURES ET SCHEMAS	7
INTRODUCTION.....	10
<b>1. CADRES THEORIQUES ET METHODOLOGIQUES .....</b>	<b>14</b>
<b><u>1.1 Etat de l'art.....</u></b>	<b>15</b>
1.1.1 Politesse, impolitesse, apolitesse : où en est-on? .....	15
1.1.1.1 Le concept de politesse... ..	15
1.1.1.2 ... Le concept d'impolitesse...	18
1.1.1.3 ... Et leurs avatars : apolitesse, hyperpolitesse, polirudesse	19
1.1.1.4 Conclusion	20
1.1.2 Emotions et attitudes	21
1.1.2.1 Définitions	21
1.1.2.2 Difficultés terminologiques	22
1.1.2.3 Modèles théoriques	23
1.1.2.4 Les affects : un sujet d'actualité	24
1.1.3 Prosodie	25
1.1.3.1 Présentation	25
1.1.3.2 Difficultés	25
1.1.3.3 L'intonation dite "neutre"	26
1.1.3.4 Les fonctions de l'intonation	27
1.1.3.5 Où en est-on aujourd'hui?	30
1.1.4 Prosodie et pragmatique	31
1.1.4.1 Organisation du discours et tours de parole	31

1.1.4.2	<i>Retombée dans le champ de la prosodie</i>	32
1.1.4.3	<i>En analyse des conversations</i>	32
1.1.4.4	<i>La phonopragmatique (Rittaud-Hutinet 1995, 2005) : un cas à part</i>	33
1.1.4.5	<i>Evolution récente</i>	34
<b>1.1.5</b>	<b>Prosodie des émotions et des attitudes</b>	<b>35</b>
1.1.5.1	<i>Introduction</i>	36
1.1.5.2	<i>Deux approches qui ont marqué la recherche en prosodie de l'expressivité des émotions et attitudes</i>	36
1.1.5.3	<i>Evolution récente</i>	38
1.1.5.4	<i>Liens entre expression des affects et analyse du discours en interaction</i>	38
1.1.5.5	<i>Liens entre prosodie des affects et impolitesse</i>	39
<b>1.1.6</b>	<b>Prosodie, politesse et interculturel</b>	<b>40</b>
1.1.6.1	<i>L'apprentissage/enseignement de la prosodie</i>	40
1.1.6.2	<i>De la nécessité d'enseigner l'intonation pour faciliter la communication interculturelle</i>	41
1.1.6.3	<i>Exemples</i>	42
1.1.6.4	<i>Conclusion</i>	44
<b>1.1.7</b>	<b>Synthèse</b>	<b>44</b>
1.1.7.1	<i>Incertitudes et questionnement</i>	44
1.1.7.2	<i>Hypothèses et principes théoriques</i>	44
1.1.7.3	<i>Retombées possibles</i>	45
	<b><u>1.2 Choix méthodologiques</u></b>	<b>46</b>
1.2.1	<b>Filtrage sémantico-lexical des énoncés</b>	<b>46</b>
1.2.2	<b>Contexte</b>	<b>46</b>
1.2.3	<b>Mimique faciale et gestuelle</b>	<b>47</b>
1.2.4	<b>Nature des données</b>	<b>47</b>

1.2.5 Paramètre d'analyse pris en compte dans la recherche : les variations de la fréquence fondamentale	49
1.2.6 Comment comparer des voix différentes?	50
1.2.7 Emission ou réception des expressions émotives?	50
1.2.8 Démarche de recherche	51
<b>2. TESTS PERCEPTIFS</b>	<b>54</b>
<hr/>	
<b><i>2.1 Présentation des corpus</i></b>	<b><i>55</i></b>
<hr/>	
2.1.1 Enregistrements : cadre, prise de son	55
2.1.2 Identification et durée des corpus	56
2.1.3 Procédure d'organisation des données	56
2.1.4 Les locuteurs et leur phonostyle	57
2.1.4.1 Identification des locuteurs	57
2.1.4.2 Phonostyles des locuteurs	57
- Voix féminines	58
- Voix masculines	60
2.1.5 Sélection des énoncés, choix des modalités	61
2.1.5.1 Impolitesse directement ou indirectement adressée	61
2.1.5.2 Sélection des énoncés	64
2.1.5.3 Modalités	64
<b><i>2.2 La méthode d'expérimentation</i></b>	<b><i>65</i></b>
<hr/>	
2.2.1 Les tests de perception auditive sur le corpus français par des locuteurs français natifs	65
2.2.1.1 Les auditeurs-testeurs natifs français	65
2.2.1.2 Passation des tests	65
2.2.1.3 Entretiens après les tests	67

2.2.2 Retour réflexif	70
2.2.2.1 Enregistrement	70
2.2.2.2 Formulation des consignes	70
2.2.2.3 Filtrage des données	71
2.2.2.4 Limites d'exploitation d'un corpus sonore spontané	71
<b><i>2.3. Résultats des tests perceptifs</i></b>	<b>73</b>
<hr/>	
2.3.1 Décompte	73
2.3.1.1 Résultats d'ensemble	73
2.3.1.2 Enoncés interrogatifs	74
2.3.1.3 Enoncés déclaratifs	76
2.3.2 Impolitesse directement ou indirectement adressée	78
2.3.2.1 Répartition	78
2.3.2.2 A qui l'impolitesse s'adresse-t-elle directement?	79
2.3.3 Prolongements	81
2.3.3.1 Pour d'autres analyses	81
2.3.3.2 Ecart entre les résultats effectifs et les résultats attendus	81
<b>3 Analyses acoustiques</b>	<b>86</b>
<hr/>	
<b><i>3.1 Procédure</i></b>	<b>87</b>
<hr/>	
3.1.1 Logiciel Praat	87
3.1.2 Mesures de F0 retenues	82
3.1.3 Mesures autres	82
3.1.4 Identification des éléments d'une fenêtre Praat	83
<b><i>3.2 Résultats significatifs</i></b>	<b>84</b>
<hr/>	

3.2.1 Rappels	84
3.2.2 Comparaison des indices acoustiques des Q° polis <sup>1</sup> et impolis	86
3.2.2.1 Augmentation brutale de F0 sur la dernière syllabe de l'énoncé	86
3.2.2.2 Augmentation et/ou diminution brutale de F0 en cours d'énoncé	88
3.2.2.3 Augmentations/diminutions compensatoires successives	89
3.2.2.4 Variation brutale de F0 dans une même syllabe	91
3.2.2.5 Chute brutale de F0 i) sur un segment de plusieurs syllabes en fin d'énoncé ii) sur la dernière syllabe de l'énoncé	92
3.2.2.6 Conclusion de l'analyse acoustique des énoncés Q°	95
3.2.3 Comparaison des analyses acoustiques des D° polis <sup>2</sup> et impolis	96
3.2.3.1 Description et analyse des indices acoustiques des énoncés D° jugés "polis"	96
3.2.3.2 Description et analyse des indices acoustiques des énoncés D° jugés "impolis"	98
3.2.3.3 Conclusion de l'analyse acoustique des énoncés D°	105
<b><i>3.3 Conclusion</i></b>	<b><i>105</i></b>
<hr/>	
3.3.1 Mise en regard des observations faites sur les corpus Q° et D°	105
3.3.2 Pour un usage de la prosodie en analyse interactionnelle	107
3.3.3 Limites	109
<b>CONCLUSION(S) ET PERSPECTIVES</b>	<b>111</b>
<b>BIBLIOGRAPHIE</b>	<b>117</b>
<b>ANNEXES</b>	
A- Fiche support des tests perceptifs	
B- Transcription des corpus	
C- Tableau de conversion des fréquences en Hertz en tons musicaux	
D- Tableaux des mesures de F0 par locuteur	

<sup>1</sup> "Q°" est ici mis pour "énoncés interrogatifs".

<sup>2</sup> "D°" est ici mis pour "énoncés déclaratifs".

- E- Capture d'écran des fenêtres d'analyses acoustiques (logiciel Praat) et tableaux des indices acoustiques par énoncé des Q° "impolies"*
- E'- Capture d'écran des fenêtres d'analyses acoustiques (logiciel Praat) et tableaux des indices acoustiques par énoncé des Q° "polies"*
- F- Capture d'écran des fenêtres d'analyses acoustiques (logiciel Praat) et tableaux des indices acoustiques par énoncé des D° "impolies"*
- F'- Capture d'écran des fenêtres d'analyses acoustiques (logiciel Praat) et tableaux des indices acoustiques par énoncé des D° "polies"*
- G- Tableaux excel des réponses aux tests perceptifs par énoncé et par locuteur*
- H- Actes de langage des énoncés*
- I- Tableau des phénomènes acoustiques affectant les énoncés (im)polis (et des autres phénomènes acoustiques (durée, intensité))*
- J- Énoncés sous format numérique wave*
- K- Fichiers Praat par taux d'(im)politesse (format numérique)*

## **TABLE DES FIGURES, SCHEMAS ET TABLEAUX**

- <i>Figure I : Répartition des réponses pour l'ensemble des questions par catégorie</i>	69
- <i>Figure II : Répartition des réponses pour l'ensemble des déclarations par catégorie</i>	69
- <i>Figure III : Variations individuelles d'interprétation des intonations polies et impolies des énoncés interrogatifs. En abscisse, le nombre de testeurs</i>	70
- <i>Figure IV : Variations individuelles d'interprétation des intonations polies et impolies des énoncés déclaratifs.</i>	70
- <i>Figure V : Répartition des questions impolies par taux d'accord</i>	71
- <i>Figure VI : Répartition des questions polies par taux d'accord</i>	71
- <i>Figure VII : Répartition des déclarations impolies par taux d'accord</i>	72
- <i>Figure VIII : Répartition des déclarations polies par taux d'accord</i>	73
- <i>Schéma I : Identification des éléments d'une fenêtre Praat</i>	82
- <i>Schéma II : Capture d'écran de la fenêtre d'analyse acoustique de Praat pour la Q°16 I_79.4%</i>	94
- <i>Schéma III : Capture d'écran des fenêtres d'analyse acoustique de Praat pour la D°20_91.2%. et la D°41_85.3%</i>	102
- <i>Schéma IV : Capture d'écran des fenêtres d'analyse acoustique de Praat pour la D°71_91.2%. et la D°41_94.3</i>	104
- <i>Tableau 1 : Tableau synthétique de présentation du corpus</i>	52
- <i>Tableau 2 : Tableau synthétique des valeurs relevées pour les voix féminines et leur conversion en tons musicaux</i>	56
- <i>Tableau 3 : Tableau synthétique des valeurs relevées pour les voix masculines et leur conversion en tons musicaux</i>	57
- <i>Tableau 4 : Difficulté ressentie par les testeurs locuteurs natifs du français dans la réalisation des consignes du 1er test (corpus français)</i>	63
- <i>Tableau 5 : Questions jugées impolies à plus de 75%</i>	71
- <i>Tableau 6 : Questions jugées polies à plus de 80%</i>	71
- <i>Tableau 7 : Déclarations jugées impolies à plus de 75%</i>	72
- <i>Tableau 8 : Déclarations jugées polies à plus de 80%</i>	72
- <i>Tableau 9 : répartition des réponses "énoncés impolis" par taux d'accord</i>	73
- <i>Tableau 10 : Répartition des réponses "énoncés polis" par taux d'accord</i>	74
- <i>Tableau 11 : Pourcentage d'énoncés impolis directement, pseudo-directement et indirectement adressés, respectivement DA, PSEUDO DA et IA</i>	75

- Tableau 12 : Synthèse des énoncés directement adressés (DA), pseudo directement adressés (PSEUDO DA) et indirectement adressés (IA) à des enfants	76
- Tableau 13 : indices acoustiques de l'énoncé D°18_P_61.8%	78
- Tableau 14 : indices acoustiques de l'énoncé Q°28_I_91.2%	87
- Tableau 15 : indices acoustiques de l'énoncé Q°23_P_85.3%	88
- Tableau 16 : indices acoustiques de l'énoncé Q°18_I_97.1%	89
- Tableau 17 : indices acoustiques de l'énoncé Q°12_I_97.1%	90
- Tableau 18 : indices acoustiques de l'énoncé Q°25_P_91.2%	90
- Tableau 19 : indices acoustiques de l'énoncé Q°34_I_94.1%	92
- Tableau 20 : indices acoustiques de l'énoncé Q°1_I_97.1%	93
- Tableau 21 : indices acoustiques de l'énoncé Q°16_I_79.4%	94
- Tableau 22 : illustration par les indices acoustiques d'un énoncé D° poli de type a) standard (D°43)	96
- Tableau 23 : illustration par les indices acoustiques d'un énoncé D° poli de type b) descendant (D°22)	97
- Tableau 24 : indices acoustiques de D°15_I_91.2%	98
- Tableau 25 : indices acoustiques de D°17_I_79.4%	99
- Tableau 26 : indices acoustiques de D°66_I_82.4%	100
- Tableau 27 : indices acoustiques de D°20_I_91.2%	101
- Tableau 28 : indices acoustiques de D°41_I_85.3%	101
- Tableau 29 : indices acoustiques de D°111_I_88.2%	103
- Tableau 30 : indices acoustiques de D°71_I_94.1%	104
- Tableau 31 : Nombre d'énoncés D°I et D°P affectés par les différents types d'indices acoustiques	105
- Tableau 32 : Moyenne du nombre d'indices acoustiques par taux d'impolitesse	105
- Tableau 33 : indices acoustiques de D°109_I_97.1%	107
- Tableau 34 : indices acoustiques de D°3_P_91.2%	109



*Qu'ils soient prolos ou profs, les gens qui parlent vite et sans réfléchir - être volubiles, garruleux, prolixes, bavocheurs, ceux qui se servent de quatre mots là où un seul suffirait - me sidèrent, toujours. Aujourd'hui encore, il m'arrive d'éprouver vis-à-vis de la parole orale un degré de self-consciousness qui relève de la torture. Plus la situation est formelle et intimidante, plus je suis susceptible d'avoir la langue qui fourche, me faisant dire un mot pour un autre ou commettre d'ahurissantes fautes de grammaire. D'où, sans doute, ma préférence pour l'écriture : là, au moins, j'ai droit au remords, à la biffure ou à l'incise. Là, de plus, mon accent ne s'entend pas.*

*Nord perdu*, Nancy Huston (1999), p. 38 (Ed. Actes Sud, coll. *Babel*)

## ***INTRODUCTION***

L'année universitaire 2012-2013 de "Master FLE (français langue étrangère) pour professionnels en un an" m'a permis de découvrir l'existence de deux champs d'études linguistiques et des problématiques qui leur sont associées. Il s'agit du plurilinguisme et des contacts de langues d'un côté, et de l'analyse interactionnelle de l'autre. Deux aspects des relations humaines qui, finalement depuis l'adolescence, ne cessaient de m'interroger, à travers principalement la question de la discorde entre les individus : pour quels motifs l'Autre, en tant qu'appartenant à un groupe différent du nôtre - que les limites situées de ce groupe soient extrêmement lointaines, proches, ou sur le continuum entre les deux extrêmes - pour quels motifs donc, l'Autre, l'Etranger, est-il rejeté? Et s'il ne l'est pas, quels sont les éléments qui, souvent, font échouer les tentatives de communication?

Par ailleurs, au cours de voyages ou d'échanges professionnels, le hasard m'a fait rencontrer de multiples langues, voire simplement des variations régionales accentuelles du français, dont l'intonation de certains énoncés a pu me heurter : dans une perception intuitive et ethnocentrée, je me suis d'abord demandé pourquoi telle question ou telle déclaration "sonnait" agressivement, ou pour mieux dire, impoliment, à mes oreilles... Il y avait là une source première, fondamentale, à malentendu interculturel!

Le sujet de ce mémoire s'est imposé à moi lorsque, au cours de mes lectures en analyse des interactions, j'ai pu croiser ces interrogations avec l'affirmation de Kerbrat-Orecchioni, qui m'a fortement interpellée : lorsque celle-ci se propose de souligner à quel point la prise en compte du contexte dans l'analyse des interactions est importante (2005 : 207), et qu'il s'agit de définir "l'effet-de-politesse (ou d'impolitesse) produit par un énoncé", elle cite bien sûr parmi les différents facteurs "le 'ton' (notion aussi importante que difficile à définir)". Plus loin (*ibid.* 210), elle renchérit en parlant du "fameux 'ton', dont tout le monde<sup>3</sup> s'accorde à reconnaître le rôle décisif (...) mais dont personne ne sait encore à vrai dire en quoi il consiste exactement : le mystère reste à peu près entier sur ce qu'est un 'ton poli' (ou impoli)".

Pourtant, si les interactionnistes évoquent régulièrement le "ton" voire la "prosodie", ils ne les mentionnent qu'en passant, pris dans les énumérations d'éléments contextuels, dont on signale à juste titre qu'ils modifient l'aspect (im)poli de l'énoncé... On parle

---

<sup>3</sup> Et de citer Grevisse (1986 : 1 594), Tannen (1986 : 54) ou Blum-Kulka (1990 : 270).

"d'accompagnement prosodique" (Kerbrat-Orecchioni 2002 : 11; 2010 : 57), d'énoncé "renforcé[e] par l'intonation" (Kerbrat-Orecchioni 2007 : 11<sup>4</sup>; 2010 : 47), ou de "la valeur insultante [d'un énoncé qui] dépend à la fois de son contenu sémantico-pragmatique, de sa réalisation linguistique, prosodique et mimo-gestuelle..." (Kerbrat-Orecchioni 2010 : 55).

A ce constat, l'étonnement et la perplexité m'ont saisie : personne donc, dans le champ de l'analyse interactionnelle ou dans le champ de la prosodie - pour mieux dire à mi-chemin entre les deux - ne s'était attelé à la description des tons impolis? Mon intuition me disait pourtant que c'était là un des éléments majeurs de l'(im)politesse... Nombre de gens sont en effet très sensibles au(x) ton(s) sur le(s)quel(s) on leur parle; la langue courante nous rappelle qu'il vaut mieux ne pas "parler sur ce ton", ou encore en guise de menace qu'on a "intérêt à changer de ton". Nombreux sont ceux parmi nous qui ont fait l'expérience de produire un enchaînement lexico-sémantique suave sur le ton exacerbé de la colère ou de la menace (ou inversement). Ne fait-on pas jouer, dans les symphonies ou autre pièce de musique, tel rôle aux instruments, selon que leur timbre est adapté à tel personnage menaçant, imposant ou au contraire fragile etc.? N'a-t-on pas, dans telle comédie<sup>5</sup>, supprimé la partie lexico-syntaxique de l'énoncé pour ne garder que l'intonation jouée par un sifflet, ce qui n'empêche pas, grâce au contexte il est vrai, de comprendre le message? D'ailleurs, «les résultats démontrent que même lorsque le contenu des questions est rendu inintelligible, une grande part de la signification affective reste dans le signal acoustique.» (Scherer, Johnstone & Sangsue 1998).

Pourtant depuis longtemps, l'on sait que l'intonation joue un rôle de premier ordre dans la communication - « Comme on sait, rien qu'en changeant l'énonciation, les mêmes mots indiquent, affirment, critiquent, nient, s'étonnent, s'indignent, interrogent, ridiculisent, déprécient » (Quintilien 90 : 3-4) - et que tout locuteur natif sait intuitivement, pour la majorité des interactions quotidiennes<sup>6</sup>, si l'on s'adresse à lui sur un ton (im)poli, qui l'enthousiasme ou l'ennuie, le blesse ou l'agresse, étrangement, "le support acoustique des mots et des phrases est jusqu'ici resté presque totalement inexploré" (Rittaud-Hutinet 1995 : 10). Or, de même que l'on reconnaît aujourd'hui sans difficulté les variations culturelles de l'(im)politesse (Kerbrat-Orecchioni 2002 : 1 ou 16 par exemple), de même on sait qu'il existe des variations culturelles de l'intonation : "language differences between cultures may have a

---

<sup>4</sup> Dont on signale toutefois que le "contour est nettement plus ascendant".

<sup>5</sup> *Le Gendarme en balade*, de Jean Girault, 1970 <http://www.youtube.com/watch?v=WZAdenw74e0>

<sup>6</sup> On reviendra cependant plus loin sur la part éminemment subjective de la perception auditive des énoncés : les variables d'appréciation sont différentes d'un individu à l'autre, et donc responsables d'un certain nombre de malentendus.

very powerful impact on vocal, particularly prosodic, parameters involved in the expression of emotion" (Banse & Scherer 1996 : 634). Bien auparavant, Fónagy (1976 : 43) affirme que "la valeur stylistique du son dépend (...) des normes phonologiques d'une langue donnée", et plus loin, que "les produits de la transformation émotive ne peuvent être correctement interprétés sans la connaissance du code de la langue en question."

La **problématique** de ce mémoire est donc la suivante : dans quelle mesure est-il possible de réduire l'apparition de malentendus communicationnels et/ou interculturels par l'apprentissage en langue seconde, ici le français, d'intonations (im)polies?

Cette recherche s'inscrit dans un **domaine** au croisement de la phonétique - plus particulièrement de la prosodie en ce qu'elle s'intéresse aux variations de la fréquence fondamentale - et de la pragmatique et didactique de l'oral - plus précisément de l'expression de la politesse et de l'impolitesse. Sa visée ressortit aussi au domaine de l'interculturel et des contacts de langues, en ce qu'elle pourra constituer le point de départ de recherches portant sur l'élucidation de malentendus possiblement dus à l'intonation, malentendus communicationnels et/ou interculturels entre locuteurs de langues maternelles et locuteurs de langues allophones.

Ce mémoire s'appuie sur **l'hypothèse** qu'il est possible de déterminer un certain nombre de traits ou marques prosodiques intrinsèquement impolis en français. Il s'agira, à partir d'énoncés produits en langue française par des locuteurs natifs du français, DE tenter d'identifier et d'isoler des marqueurs acoustiques pertinents de l'impolitesse, pour tâcher de déterminer le plus précisément possible des patterns caractéristiques.

Après avoir présenté les théories et notions utiles dans les champs de la pragmatique, de l'expression des émotions et de la prosodie, je présenterai le travail que j'ai mené : 1) tests perceptifs opérés par des locuteurs natifs de français sur des énoncés conversationnels en langue française, afin d'en déterminer le caractère (im)poli. 2) mesures acoustiques de la fréquence de la voix des locuteurs, comparaison des différents profils de variation de la fréquence fondamentale (appelés aussi "courbes", "patrons" ou "patterns") et mise en relation avec le caractère poli ou impoli précédemment perçu. Je dégagerai alors si possible des caractéristiques prosodiques, notamment intonatives, d'énoncés (im)polis, pour finalement envisager les retombées éventuelles dans les champs de recherche concernés et faire part de quelques pistes didactiques dans les domaines conjoints de l'intonation et de l'interculturel.

# **PARTIE I.**

## **Cadres théoriques et méthodologiques**

*(...) dans le commerce des humains, bien souvent les mouvements du corps, les intonations de la voix et les expressions du visage en disent plus long que les paroles (...)*

*Un mot pour un autre, préambule, Jean Tardieu (1951)  
(Ed. Folio Junior).*

## 1.1 Etat de l'art

### 1.1.1 Politesse, impolitesse, apolitesse : où en est-on?

#### 1.1.1.1 Le concept de politesse...

Il faut attendre les années 1970 pour que les linguistes s'intéressent à la description des marques de la politesse dans les relations humaines lors des conversations quotidiennes. Auparavant, le terme de politesse renvoyait uniquement au savoir vivre en société, aux règles de maintien, aux "bonnes manières" etc.; sur ce point, la littérature, prescriptive, est assez abondante depuis le 18<sup>e</sup> siècle.

Les précurseurs de ce changement d'approche de la politesse furent :

1) Grice (1975) qui pose les "Principes de Coopération" sous la forme de quatre règles ou maximes visant avant tout l'efficacité de la communication. Ce sont les maximes i) de quantité : il s'agit de ne pas trop en dire, ni trop peu; ii) de qualité : le locuteur doit se limiter à dire ce dont il a la preuve, ce qui implique qu'il ne doit pas mentir; iii) de pertinence : il faut parler à propos; iv) de manière : on se doit de s'exprimer avec clarté, de manière brève et ordonnée. A ces règles, Grice ajoute la notion d'"implication" (traduction de l'anglais *implicature*), qui consiste pour les interlocuteurs à induire les éléments manquants d'un énoncé. Evidemment les implications sont à même de causer nombre de malentendus puisque le déroulement harmonieux de l'interaction va reposer sur des inférences. La notion d'implication, bien que les pragmaticiens n'y fassent dans l'ensemble guère référence, me semble tout à fait intéressante : n'est-ce pas là en partie l'implicite que l'analyste des conversations traque derrière les actes de langage explicites, implicite sur lequel se construisent entre autres les énoncés ironiques, une certaine forme d'humour ou de reproches, etc.? On peut se demander en outre, **étant donnée la problématique de ce mémoire, comment l'intonation, ajoutée au comportement mimo-gestuel, peut aider le destinataire à percevoir les implications du destinataire.**

2) Lakoff (1972, 1973), qui pose la problématique de la variation linguistique à travers les cultures, depuis l'usage de certaines particules en passant par celui de formes verbales, jusqu'à l'expression de la politesse, dans un double objectif : i) montrer que les phénomènes linguistiques dépendants du contexte peuvent être identifiés et décrits pour toutes les langues;

ii) contribuer à l'amélioration de l'enseignement/apprentissage des langues secondes (Lakoff 1972 : 908), **ce qui intéresse particulièrement la problématique de ce mémoire.**

Mais la notion de politesse linguistique a réellement été théorisée dans un ouvrage qui a fait date, celui de Brown & Levinson (1978, 1987), et qui est au fondement des recherches dans ce domaine. S'appuyant à la fois sur les travaux de Searle (1969) sur les actes de langage (*Speech Acts*), et ceux de Goffman (1973) sur la notion de faces, il reste à ce jour le plus célèbre et le plus exploité, mais aussi le plus critiqué ou aménagé.

L'interaction est vue par Goffman comme une pièce de théâtre, dans laquelle les interactants sont des acteurs qui jouent un rôle, dans un cadre ou décor donné. Chaque locuteur, par besoin d'approbation, d'estime, de reconnaissance, revendique pour lui-même au cours de l'interaction une image sociale positive : il s'agit du *face-want*. Il prend donc soin de se montrer à son avantage et essaie de cacher ce qu'il croit pouvoir être interprété comme un défaut ou une faiblesse<sup>7</sup>. Chaque nouvelle interaction remet en jeu cette image positive des interactants qui n'est jamais définitivement acquise. L'image sociale positive à protéger prend deux formes : i) le territoire, rebaptisé "face négative" par Brown et Levinson, désigne l'étendue physique (corps, espace, biens) ou symbolique (opinions, croyances, sentiments) que l'individu souhaite préserver du contact, du regard ou du jugement d'autrui; ii) la face, dite "face positive" par Brown & Levinson, désigne quant à elle l'image psychologique, professionnelle ou hiérarchique positive que l'individu a de lui-même et qu'il souhaite voir respectée ou approuvée par autrui. Préserver cette image positive, ne pas "perdre la face" et ne pas la faire perdre à l'autre, "car la perte de face est une défaite symbolique qui peut mettre à mal l'ordre de l'interaction" (Kerbrat-Orecchioni 2005 : 195), tels sont les enjeux de toute rencontre sociale; pour y parvenir, les interactants ont à leur disposition un certain nombre de stratégies, désignées par le terme de "travail de figuration des faces" (*face-work*), qui permettront de maintenir le bon déroulement de l'interaction. Les procédés "adoucisseurs" (*softeners*) font partie de ces stratégies : il s'agit par exemple de prendre des précautions liminaires avant de déranger quelqu'un ("*Je peux* vous demander quelque chose?"), de désarmer une possible réaction de refus ("*Je ne voudrais pas* vous déranger/gêner/interrompre, mais..."), de minimiser le dérangement occasionné ("*Tu peux m'attendre cinq petites minutes?*"), ou d'amadouer l'interlocuteur en échange d'une requête

---

<sup>7</sup> Ici, le mot "personne" rejoint son étymologie : la personne est un masque. (Personne < *persona* (lat. emprunté à l'étrusque) : "masque").

("Sois gentil, laisse-moi tranquille", "Toi qui es si habile, tu peux...?")<sup>8</sup>. Ont aussi été identifiés des "intensifieurs", qui renforcent l'acte de langage ("C'est vraiment trop gentil!").

C'est précisément cette panoplie de procédés que Brown & Levinson appellent "politesse", alors que le *Face Threatening Act (FTA)*, ou "Acte menaçant pour la/les face(s)", mettra en péril le bon déroulement de l'interaction. Pour Brown & Levinson (1987 : 69, cité par Culperer 1996 : 356), le locuteur a à sa disposition cinq formes de FTAs : 1) *Bald on record* : le FTA est ouvertement proféré ("Ton dessin ne ressemble à rien!") - 2) *Positive politeness* : le locuteur va tâcher de compenser l'atteinte qu'il commet à la face positive de son interlocuteur ("Il faudrait faire un effort sur le coloriage, *mais* le dessin est bien exécuté!") - 3) *Negative politeness* : même démarche pour la face négative de l'interlocuteur ("*Quel dessinateur habile!* Mais en ce qui me concerne, je trouve les coloris un peu vifs".) - 4) *Off-record* : le locuteur n'assume pas directement le FTA ("On dirait que tu as utilisé toute ta boîte de crayons de couleurs...") - 5) *Withhold the FTA* : le FTA n'est pas réalisé, le locuteur ne dit rien (ni FTA ni compliment). Il existe aussi pour les FTAs des intensifieurs, plutôt appelés "durcisseurs" dans ce cas, qui accentuent alors notablement l'impolitesse ("Tu es vraiment idiot!"). Il faut ici préciser que, du point de vue de Brown & Levinson, tout acte de langage est potentiellement menaçant pour les interactants en présence, ce qui met en danger permanent le désir universellement partagé de *face-want*.

On a reproché au modèle brown-levinsonien i) de prétendre à l'universalité. Dans une approche résolument interculturelle, Watts, Ide & Ehlich (1992-2003), notamment, proposent un modèle alternatif à celui de Brown & Levinson, en mettant en avant l'extrême variabilité, dans le temps et dans l'espace, du concept de politesse et de ses manifestations (2003 : xi sq.). Cependant, si l'on adhère à l'argumentation perspicace de Kerbrat-Orecchioni (2002 : 6 sq.), la prétention à l'universalité imputée à Brown & Levinson découle d'une lecture erronée; de plus, Brown & Levinson (1978 : 288) eux-mêmes sont très clairs sur la question : "The application of the principles differs systematically across cultures and within cultures across subcultures, categories and groups." ii) de proposer une vision excessivement pessimiste des relations sociales humaines. C'est pourquoi Kerbrat-Orecchioni (2002, 2005, 2010) propose de l'aménager en y faisant une place aux *FFAs (Face flattering acts)*, partant de l'observation qu'un certain nombre d'actes de langage sont tout de même intrinsèquement positifs ou valorisants.

---

<sup>8</sup> Pour une liste plus détaillée, voir Kerbrat-Orecchioni 1992 : 2<sup>e</sup> partie, chap.2.

Depuis la fin des années 70 donc, et l'ouvrage de Brown & Levinson, l'intérêt pour la question de la politesse s'est grandement accrue, donnant naissance à un nouveau champ de recherche dont on peut dire avec Preisler & Haberland (1994 : 227, repris par Kerbrat-Orecchioni 2005 : 190) que depuis un vingtaine d'années, il "constitue l'un des domaines de recherche les plus importants et productifs en pragmatique et sociolinguistique." Un certain nombre de cadres théoriques ont depuis vu le jour proposant des angles d'approche divers : "la politesse doit-elle traitée comme 'un langage de connotation' (Pernot), en termes de 'travail des faces' (Goffman, Brown & Levinson), de 'coût/bénéfice' (Leech), ou de 'respect de contrat conversationnel' (Fraser)?" (Kerbrat-Orecchioni 2002 : 1-2). **Pour ma part, je me rangerai dans la lignée des théoriciens du travail des faces.**

#### 1.1.1.2 ... *Le concept d'impolitesse* ...

L'impolitesse, concept apparu plus récemment dans le domaine interactionnel, se présente quant à elle comme le revers de la politesse, ou pour reprendre Culperer, le premier interactionniste à théoriser le concept (1996), "the parasite of politeness" (*Ibid.* : 355); ce champ investit les stratégies que les locuteurs mettent en œuvre lorsqu'ils souhaitent, à l'inverse de celles utilisées dans le cadre de la politesse, provoquer une rupture dans le contrat conversationnel, mettre les faces de leur(s) interlocuteur(s) en péril, ce qui ne va pas sans blessure narcissique : "such behaviours always have or are presumed to have emotional consequences for at least one participant, that is, they cause or are presumed to cause offence." (Culperer 2010 : 3233).

Culperer (1996 : 350) signale quelques prédécesseurs dans le domaine, dont aucune des études ne se sont pourtant données pour objectif d'en comprendre de manière exhaustive le fonctionnement ou d'en assurer les fondements théoriques. Il reprend ensuite les différentes stratégies de politesse proposées par Brown & Levinson (1987) pour en donner le pendant dans le domaine de l'impolitesse, en les classant de la plus menaçante à la moins menaçante : i) *bald on record impoliteness*, ii) *positive impoliteness*, iii) *negative impoliteness*, iv) *sarcasm or mock impoliteness*. (*Ibid.* : 356). Il énumère (*Ibid.* : 357-358) les différentes stratégies i) d'impolitesse positive - celle qui attaque la face positive de l'interlocuteur - : ignorer sa présence, le mettre à l'écart, parler une langue qu'il ignore, rechercher le désaccord,

etc.; ii) d'impolitesse négative - qui s'en prend à la face négative de l'interlocuteur - : le menacer, le ridiculiser, envahir son espace physique ou symbolique etc<sup>9</sup>.

Le concept, considéré alors comme un champ d'investigation à part entière, a ensuite été principalement complexifié par Bousfield (2007a, 2007b, 2008, 2010), Culperer, Bousfield & Wichmann (2003), Bousfield & Locher (2008). Leur recherche prend comme support d'étude certains types d'interactions à fort potentiel de confrontation, telles les interactions en milieu militaire, policier, ou juridique, c'est-à-dire finalement dans un cadre où les rapports dissymétriques entre les interactants régissent les interactions. Parallèlement, à partir de 1996 pour le monde anglo-saxon, et de 2008 pour la France (Moïse *et al.* 2008), l'intérêt s'est porté sur la problématique des conflits et de la violence verbale, qui apparaissent notamment dans les espaces publics et urbains ou en milieu scolaire. Il s'agit là de cerner les mécanismes de "la montée en tension verbale" puis du "dénouement de la crise". Dans la lignée de cette dernière approche, **ma recherche tente de cerner ce qui dans la prosodie, et notamment dans les variations de la fréquence fondamentale, peut aider à identifier la violence verbale et participer de "la montée en tension verbale"**.

#### *1.1.1.3 ... Et leurs avatars : apolitesse, hyperpolitesse, polirudesse*

D'autres concepts connexes sont aussi apparus, ceux d'"apolitesse", d'"hyperpolitesse", et de "polirudesse".

L'apolitesse, ou non-politesse, est caractérisée par l'"absence 'normale' d'un marqueur de politesse" (Kerbrat-Orecchioni 2005 : 209). Il s'agirait d'une sorte de degré zéro de l'(im)politesse de l'interaction, un comportement neutre que l'on ne pourrait qualifier ni de "poli" ni d'"impoli" (Lakoff 1989 : 103 cité par Kerbrat-Orecchioni 2002 : 14, et Oprea 2012 : 145). La difficulté consiste toutefois à déterminer *dans quelle situation* l'absence de politesse est considérée comme normale, donc acceptable... Cela dépendra assurément de critères (inter)culturels, sociologiques, individuels et situationnels, ce qui pourra provoquer bien des malentendus, notamment interculturels (Kerbrat-Orecchioni 2005 : chapitre 4).

L'hyperpolitesse se signale par la présence excessive de marqueurs de politesse par rapport à la norme socialement attendue (*Ibid.*). Elle est interprétée soit comme une marque

---

<sup>9</sup> Ma traduction.

de respect ostensible notamment en situation d'interaction hiérarchique, soit comme l'un des outils favoris de l'implicite, (ironie, humour, reproche etc.), et donc parfois de l'impolitesse. On peut ici poser la même question que précédemment à propos de cette "norme socialement attendue", qui va forcément varier d'une culture à l'autre, d'un milieu social à l'autre, d'une personne à l'autre, etc.

La polirudesse enfin, mot forgé sur l'anglais "*polirudeness*", renvoie à une catégorie complexe d'énoncés à la fois polis et impolis (Kerbrat-Orecchioni 2010 : 39-40). On en trouvera des exemples dans Kerbrat-Orecchioni (2007 : 9-12) qui analyse les marques de pseudo politesse au cours d'un débat télévisé opposant Nicolas Sarkozy à Jean-Marie Le Pen : forme polie ("*Bonjour Monsieur Le Pen*"), reproche implicite (*vous ne m'avez pas dit bonjour*) en même temps que manœuvre interactionnelle (*tout ce qui a été dit auparavant est considéré comme nul et non avvenu, puisque les salutations, qui ouvrent d'ordinaire les échanges, n'ont pas eu lieu*).

#### 1.1.1.4. Conclusion

1) Les concepts d'(im)politesse sont bien entendus centraux dans ma démarche, puisqu'il s'agira i) de catégoriser, au moyen de tests perceptifs effectués par des auditeurs naïfs, un certain nombre d'énoncés comme "impolis" au regard de l'intonation. ii) de mettre en évidence au moyen de l'analyse acoustique des traits ou marques prosodiques propres à l'impolitesse (ou à des formes d'impolitesse).

Si l'on peut cerner assez facilement les marques d'(im)politesse (Kerbrat-Orecchioni 2005, 2007) pour ce qui est du contenu verbal des interactions (termes axiologiques, langue familière ou vulgaire)<sup>10</sup>, qu'en est-il du contenu prosodique? Quelles sont les marques d'une intonation (im)polie, attendues ou non, (in)consciemment utilisées par le locuteur, de manière à maintenir/rompre le contrat interactionnel? A fortiori, comment déterminer "l'absence 'normale'" ou "la présence excessive" d'intonation polie? Le champ n'est abordé que depuis très récemment, sous l'angle de l'étude des émotions et des attitudes<sup>11</sup>.

---

<sup>10</sup> Encore que la méthodologie d'analyse soit grandement complexifiée par les variations culturelles (Kerbrat-Orecchioni 2005 : chapitre 4), et socioculturelles (Culperer 1996 : 352 sq.) et ci-dessous en 2).

<sup>11</sup> Voir en 1.1.5.

2) Par ailleurs, comme le souligne la définition de l'impolitesse proposée par Culperer (2010 : 3233),

*Impoliteness is a negative attitude towards specific behaviours occurring in specific contexts. It is sustained by expectations, desires and/or beliefs about social organisation, including, in particular, how one person's or group's identities are mediated by others in interaction. Situated behaviours are viewed negatively when they conflict with how one expects them to be, how one wants them to be and/or how one thinks they ought to be. (...) Various factors can exacerbate how offensive an impolite behaviour is taken to be, including for example whether one understands a behaviour to be strongly intentional or not.*

Il faut ainsi insister sur le caractère éminemment subjectif du ressenti de l'(im)politesse, qui est doublement encodé chez i) le locuteur (*Ibid.* 3237), ii) l'interlocuteur (Culperer 2010 : 3234 et Leech 1983 : 23-24, cité par Culperer 2010), par le processus de socialisation. Les représentations en sont donc très variables d'un individu à l'autre, puisqu'elles se construisent entre autres choses au moyen de données familiales, socioculturelles, ethniques et/ou nationales<sup>12</sup>. Elles sont en outre situées historiquement et géographiquement. **Mon travail dans le cadre de ce mémoire tâche de montrer que le malentendu, interculturel ou non, peut prendre son origine dans un double encodage prosodique.**

### 1.1.2 Emotions et attitudes

L'impolitesse a partie liée avec les émotions et les attitudes<sup>13</sup>, c'est pourquoi il convient de s'attarder un peu sur ce sujet.

#### 1.1.2.1 Définitions

Pour Darwin (cité par Léon 1993 : 113; repris par Lacheret-Dujour 2011b : 4; Scherer *et al.* 2001 : 76), l'émotion<sup>14</sup>, ou "état émotionnel", est la manifestation physiologique d'un état psychologique, "une réponse de l'organisme à une situation donnée". Dans la lignée de cette définition, on considère aujourd'hui en neurophysiologie les émotions comme l'ensemble

<sup>12</sup> Pour les variations interculturelles de l'(im)politesse, voir aussi 1.1.6.

<sup>13</sup> Voir en 1.1.6.

<sup>14</sup> A noter qu'étymologiquement (émotion < \**ex-movere* (lat.) : "mettre en mouvement, remuer"), le terme "émotion" renvoie effectivement à un mouvement, ici appliqué au corps, une agitation plus ou moins importante qui s'opposerait à un état d'immobilité, de tranquillité, de repos.

des réponses physiologiques, somatiques (accélération du rythme cardiaque, sueurs etc.), provoquées par l'environnement. Pour les interlocuteurs, les émotions se manifestent à la fois dans le contenu verbal, la voix, les gestes et attitudes corporelles, et les expressions faciales (*Ibid.*). Ces réactions, "généralement brusques et momentanées" (Petit Robert 1994 : 744) ne doivent pas être confondues avec les sentiments,

*"qui correspondent à une étape dérivée plus élaborée des processus émotionnels et qui nous intéressent plus particulièrement, nous linguistes, puisqu'ils donnent lieu à la verbalisation. En d'autres termes, au stade du sentiment, une émotion d'origine somatique et inconsciente est transformée en image mentale, cette transformation associée à la conscience de soi/de l'autre, convoque (...) l'intervention des processus cognitifs de haut niveau, en particulier le langage."*

Lacheret-Dujour 2011b : 4

On peut rapprocher cette distinction entre "émotions" et "sentiments" de celle opérée par les psychologues, qui ont différencié l'"émotion brute", "émotion choc" ou simplement "émotion", imprévisible, *accidentelle*<sup>15</sup> et par laquelle l'individu peut être submergé, de l'"émotion socialisée", qui correspond à un comportement contrôlé, plus généralement appelée "attitude" par les linguistes (Léon 1993 : 113). Les attitudes<sup>16</sup>, *social affects* en anglais, renvoient à un état psychologique actualisé (para)linguistiquement; citons par exemple l'admiration, l'agressivité, la déception, l'incrédulité, l'insolence, la provocation, la résignation, etc.

### 1.1.2.2 Difficultés terminologiques

Se pose alors le problème de la terminologie des émotions et attitudes, des "étiquettes", c'est-à-dire aussi de la délimitation entre des émotions/attitudes connexes, et de celui de l'identification et de la neutralisation<sup>17</sup> des synonymes. Qu'appelle-t-on par exemple "colère"? Banse & Scherer (1996) et Scherer (1986, cité par Scherer, Johnstone & Klasmeyer 2003 : 433) ont montré que, dans différentes expérimentations, ce terme pouvait désigner aussi bien une légère irritation qu'un violent accès de rage. En effet, jusqu'à encore récemment (Keltner & Ekman 2003 : 411),

<sup>15</sup> Du latin *accido* < \**ad-cado* : "tomber sur/ à la rencontre de", ce qui donne l'image d'un événement inattendu (cf l'expression familière "tomber sur le coin de la figure").

<sup>16</sup> On ne confondra pas le sens du terme "attitude" ici proposé avec celui qu'on lui attribue généralement en sociolinguistique, à savoir "pratique langagière".

<sup>17</sup> C'est-à-dire ne pas utiliser deux mots voisins pour un même affect, "peur" et "crainte" par exemple.

i) la typologie des émotions était bien réduite : la classification de Woodworth (1939, cité par Léon 1993 : 115) donne six émotions *de base* (amour, surprise, peur, colère, dégoût, mépris), réaménagées par Scherer *et al.* (2001), Lacheret-Dujour (2011b : 5) et Scherer (2003, *ibid.*) en cinq états émotionnels de base : joie, colère, tristesse, dégoût, peur. Sauf exception (Scherer 1996 et 2001) et alors que la plupart des théoriciens des émotions et la plupart des individus revendiquent un éventail d'émotions bien plus large, les expérimentations, empiriques, se sont jusqu'à la fin des années 90 appuyées sur ces états émotionnels de base (Keltner & Ekman 2003 : 412).

Le travail théorique nécessaire pour approcher un consensus terminologique n'a pour le moment pas abouti (Lacheret-Dujour 2011b : 5, Fónagy 2003 : 15), bien que les travaux se poursuivent, notamment dans le champ de la lexicologie (voir par exemple Novakova & Tutin: 2009).

ii) Alors que Lacheret-Dujour (2011b : 8) et d'autres (Keltner & Ekman 2003 : 412; Audibert, Aubergé & Rilliard 2008), affirment que "l'expression des émotions s'inscrit de fait dans un processus de communication multimodale : corporelle, gestuelle, vocale et verbale," l'analyse de l'expression des émotions ne s'est pourtant concentrée que sur les traits du visage, les études ne portant que rarement (voir 1.1.3 à 1.1.5) sur l'expression vocale ou linguistique des émotions (Keltner & Ekman 2003 : 411).

### 1.1.2.3 Modèles théoriques

Scherer, Johnstone & Klasmeyer (2003 : 434-436) identifient trois modèles théoriques concernant les émotions.

**a)** Le premier, "*Arousal Theories*" ou "théories de l'activation des émotions"<sup>18</sup>, est tombé en disgrâce à la fin des années 60. Ce modèle se concentrait principalement sur la prédiction de manifestations physiologiques suscitées par telle ou telle émotion, sans s'intéresser toutefois à son expression vocale ou faciale. Selon cette théorie, l'ensemble des émotions serait réparti en émotions actives et passives, et s'étendrait sur un continuum d'activation empathique.

---

<sup>18</sup> Ma traduction.

b) Au début des années 70, le modèle des émotions "discrètes" de Tomkins, popularisé par Ekman et Izard, voit le jour. Dans la lignée de Darwin, ces chercheurs ont postulé l'existence de patrons de réponses spécifiques pour les émotions de base, principalement dans le domaine des mimiques faciales, laissant de côté les expressions vocales qui leur correspondaient.

c) Enfin, dans la lignée des travaux précédents, la théorie la plus récente, (voir Scherer, Schorr, Johnstone, cités en exemple par Scherer *et al.* (2003 : 435), concerne la perception des émotions (*appraisal*), et postule que les émotions i) se différencient selon leur mode d'évaluation d'un événement, établi à partir de critères perceptifs; ii) sont un phénomène adaptable, dans le sens où l'être humain peut réévaluer la situation dans laquelle est apparue l'émotion (Ellsworth & Scherer 2003 : 572). Suite à quoi, Scherer (1986, cité par Scherer *et al.* 2003 : 435) a proposé une liste assez complète des changements physiologiques attendus, y compris dans le domaine de l'expression vocale et de ses paramètres.

#### *1.1.2.4 Les affects : un sujet d'actualité*

Le thème de l'affectivité - émotions, sentiments, attitudes, affects... - est actuellement au cœur de projets de recherche variés, qui s'étendent sur plusieurs champs : en ne se cantonnant qu'à l'Université de Grenoble, on en relèvera en lexicologie (par exemple le projet *Emolex* du laboratoire des Sciences du Langage de Grenoble 3), en sciences de l'information et de la communication (thèse de Céline Battaïa pour le laboratoire du GRESEC<sup>19</sup> de Grenoble 3), en prosodie des attitudes dans une approche interculturelle (thèse en cours de Yan Lu pour le Gipsa-lab<sup>20</sup> de Grenoble), et plus largement en prosodie de l'expression affective (voir 1.1.5). Ces recherches intéressent de très près les spécialistes du traitement automatique du langage ainsi que les cliniciens, car

*"des hypothèses fortes de la psychologie cognitive et de la neuropsychologie donnent aux émotions un statut central : traiter les émotions dans l'interaction verbale homme-machine (...) peut permettre de ne pas perturber l'efficacité même de la communication,*

---

<sup>19</sup> Groupe de Recherche sur les Enjeux de la Communication.

<sup>20</sup> Grenoble Images Parole Signal Automatique.

### 1.1.3 Prosodie

#### 1.1.3.1 Présentation

En français, la prosodie est mise en œuvre à travers les variations de la fréquence fondamentale<sup>22</sup> (dorénavant F0), la durée et l'intensité. On peut distinguer trois niveaux d'analyse : i) phonétique (acoustique), ii) phonologique (sur les unités distinctives telles que les groupes intonatifs), iii) linguistique (au niveau fonctionnel). Deux approches sont possibles : l'approche expérimentale, qui utilise les appareils de mesure disponibles, et l'approche perceptive, qui, faisant appel au traitement de l'information auditive par l'auditeur, met en jeu l'oreille comme le cerveau. **La démarche méthodique de mon travail utilisera les deux approches : d'abord une catégorisation perceptive par des auditeurs naïfs, puis une analyse acoustique, phonétique et phonologique, à l'aide d'un logiciel spécialisé.**

Deux théories se font concurrence dans le domaine de la prosodie : soit on considère que "prosodie" et "intonation" sont synonymes, les phénomènes d'accentuation et l'organisation tonale étant alors regroupés sous cette double étiquette, soit on fait la différence entre "intonation" et "prosodie", la première désignant simplement la variation de la F0, tandis que la seconde ne s'intéresse qu'à l'accentuation des mots et des groupes intonatifs (Lacheret-Dujour & Beaugendre 1999 : 12-13). **C'est cette dernière position que j'adopterai dans mon travail, et lorsque je parlerai d'"intonation", il faudra entendre "variation de la fréquence fondamentale".**

#### 1.1.3.2 Difficultés

Les spécialistes de la parole vocale quotidienne soulignent souvent la complexité de leur objet d'étude (Bachorowski 1999 : 53; Scherer *et al.* 2003 : 437-438; Devillers & Vasilescu 2004 : 1; Lacheret 2011a : 87-88 etc.). Cette complexité tient d'une part à "la dimension continue du langage, qui rend difficile le repérage et l'extraction d'unités discrètes", et à "ses fonctions linguistiques plurielles", c'est-à-dire phonologique, syntaxique et

<sup>21</sup> Aubergé (2002 : 265) s'appuie sur l'ouvrage de Damasio A.R. (*Descartes' error. Emotion, reason, and the human brain*, 1994, New York : Putnam Books), pour qui "La capacité d'exprimer et ressentir les émotions est indispensable à la mise en œuvre des comportements rationnels."

<sup>22</sup> La "fréquence fondamentale" est le nombre d'accolements entre les cordes vocales réalisé par seconde. Elle est mesurée en hertz. Plus les accolements sont nombreux, plus le son est aigu. La "variation" de la fréquence fondamentale désigne la mélodie de la voix, elle est représentée sous forme de courbe.

pragmatique (Lacheret 2011a : 87). Elle tient, d'autre part, à ce que la prosodie a longtemps été négligée par la linguistique, peut-être précisément à cause des difficultés qui viennent d'être évoquées (*Ibid.*).

On considère traditionnellement que la production vocale s'appuie sur la respiration, la phonation et l'articulation<sup>23</sup>. Parmi les paramètres articulatoires, on distingue ceux qui sont nécessaires à la production d'un énoncé intelligible (points d'articulation), et ceux qui rendent compte de facteurs extralinguistiques tels que les mimiques faciales (Scherer *et al.* 2003 : 437-438). Mais les composantes phonostylistiques propres à chaque locuteur viennent complexifier l'interprétation des paramètres articulatoires (*Ibid.*).

Afin d'extraire pour analyse les caractéristiques des variations de paramètres acoustiques, comme la variation de la F0, il est pratique de les convertir en données plus aisément lisibles, de trouver des modes de représentation de la prosodie. Toutefois, les linguistes ne sont pas parvenus à s'accorder<sup>24</sup> sur la forme que doit adopter la stylisation de ces données, ni sur les distinctions à faire entre données significatives et données secondaires (*Ibid.*); tout cela peut en outre varier en fonction de ce dont on souhaite rendre compte et du type de parole produite (Lacheret 2011a : 89).

### 1.1.3.3 L'intonation dite "neutre"

Fónagy dit (2003 : 12) que "l'auditeur [et le locuteur] ne peu[ven]t s'empêcher d'attribuer un sens aux écarts des mouvements mélodiques par rapport à une ligne mélodique **idéale (abstraite)**<sup>25</sup> parfaitement neutre." Rittaud-Hutinet (1995 : 17) affirme de son côté pour la langue orale, que, si "Notre sentiment de locuteur natif refuse d'admettre l'absence d'une langue neutre", que l'on pourrait appeler "le degré zéro de l'expressivité", l'énonciation neutre n'en est pas moins

*un mythe, une reconstruction de notre esprit qui veut croire que la parole neutre existe parce que, des mille et une expériences d'interactions langagières que nous avons vécues, nous pouvons extraire une langue abstraite, résultat de*

<sup>23</sup> Mais comme il est particulièrement difficile de différencier les effets acoustiques de la respiration et de la phonation, on les envisage la plupart du temps comme une seule et même composante de la production vocale.

<sup>24</sup> Certains utilisent des "portées" à 4 lignes sur lesquelles ils schématisent les courbes de F0, d'autres mettent les données chiffrées en tableaux, d'autres utilisent des symboles - H, H\*, HH, etc...

<sup>25</sup> Souligné par mes soins.

*l'élimination de toutes les variations, de tout ce qui n'est que particulier. (...)  
Comparer LA langue, objet unique, à la richesse des variations, n'est donc qu'une  
illusion, puisque, tout fait de parole étant soumis aux contraintes d'une situation, la  
forme du message qu'il contient est indissociable (entre autres) de la présence de  
son énonciateur dans cette situation.*

Rittaud-Hutinet (1995 : 19)

On rejoint ici le postulat de la communication de Jakobson, en ce que le "destinateur" est présent dans son message, dont il organise la "poétique", dans un "contexte" donné. Il faut conclure de cet extrait que la prosodie fait partie intégrante de la fonction poétique du langage, en particulier de ses aspects illocutoires et perlocutoires.

**Pour me situer par rapport à cette intonation "neutre"**, mon expérimentation m'a amenée à conclure qu'une telle intonation n'existe effectivement pas, du moins dans les situations d'interactions quotidiennes. Dans les corpus que j'ai pu rassembler, les rares fois où l'on pourrait éventuellement conclure à une intonation neutre, celle-ci est tout de même porteuse de sens, que l'on peut interpréter comme un désengagement du locuteur, de renoncement, voire d'abatement.

**Tout énoncé est donc émotif, ou tout au moins attitudinal.**

#### *1.1.3.4 Les fonctions de l'intonation*

Morel & Danon-Boileau (1998 : 9) octroient deux fonctions à l'intonation : fonction démarcative et fonction iconique (d'expressivité). Fónagy (2003 : 2) constate que la plupart des auteurs lui en donnent trois, - mais rarement les mêmes! - et s'essaie à en proposer une synthèse; pour lui, on peut dénombrer quatorze fonctions de l'intonation différentes : fonctions d'emphase, grammaticale, syntaxique, modale, imitative, d'appel, logique, prédictive, allusive, identificatrice, de caractérologie vocale, esthétique, expressive, exploratoire et préparatoire. Il semble toutefois possible de regrouper toutes ces fonctions dans l'une des deux catégories fonctionnelles proposées par Lacheret-Dujour (2011a : 91)<sup>26</sup> :

i) fonction de structuration du message (segmentation, intégration, hiérarchisation,

---

<sup>26</sup> Les fonctions segmentatrice et démarcative, grammaticale, syntaxique, logique et prédictive seraient à ranger dans la catégorie "structuration du message", tandis que les fonctions d'emphase, modale, imitative, d'appel, allusive, identificatrice, de caractérologie vocale, esthétique, expressive, et exploratoire seraient à classer dans la "fonction pragmatique".

groupement, équilibre), bien plus large que la fonction démarcative évoquée précédemment, et ii) fonction pragmatique (marquage de la force illocutoire d'un énoncé, expression des émotions).

Je m'arrêterai sur les fonctions d'emphase, allusive, expressive, grammaticale et syntaxique, soit parce qu'elles sont de première importance dans le champ de la prosodie, soit parce qu'elles concernent directement la recherche menée dans ce mémoire.

#### a) Fonction d'emphase

L'intonation, grâce à l'accent d'insistance visible dans les valeurs de hauteur, de durée et d'intensité, dirige l'attention de l'interlocuteur sur tel mot ou tel passage de l'énoncé. A l'inverse, une absence d'emphase relègue à l'arrière plan les informations secondaires. Cette fonction a son importance pour mon travail puisqu'elle concerne directement l'analyse conversationnelle; significative d'une manière générale en pragmatique, on verra que l'emphase peut venir appuyer l'impolitesse, voire la générer.

#### b) Fonction allusive

Là encore, nous sommes en lien direct avec la pragmatique : l'intonation, "vaguement distinctive", "permet de suggérer des opinions sans s'engager formellement ou d'exprimer des idées en germe (...). On peut dire que la mélodie des énonciations est toujours plus ou moins allusive" (Fónagy 2003 : 12). Et de donner l'ironie en exemple. On voit que cette fonction renvoie à l'implicite, (pensons aussi aux implicatures de Grice), dont sont empreints nombre d'énoncés quotidiens et que l'interactionniste a pour objectif de révéler.

c) Si certaines des fonctions de l'intonation ont pu être discutées, il n'en est rien de sa fonction expressive. "Les discussions étaient et restent centrées sur le nombre des émotions et des attitudes émotives que les moyens prosodiques permettent de distinguer." (Fónagy 2003 : 15)<sup>27</sup>. **Il me semble que la fonction a) d'emphase peut être englobée dans la fonction c) d'expressivité : je ne les distinguerai pas dans mon travail.**

d) Pour ce qui est de la fonction grammaticale, deux écoles semblent s'affronter. L'une (Pike 1945, Crystal 1969, cités par Fónagy 2003 : 4) refuse toute portée grammaticale à

---

<sup>27</sup> Voir en 1.1.5. Prosodie des émotions et attitudes.

l'intonation : Crystal va même jusqu'à parler d'"absurdité théorique" à ce sujet (*Ibid.*). Pour eux, l'émotion et les attitudes d'énonciation dominent l'intonation (Bolinger *Ibid.*). Lacheret-Dujour en revanche (2011b : 1) pose des arguments "qui militent en faveur du statut grammatical de l'expressivité vocale" :

*l'expression prosodique des émotions fait l'objet d'un apprentissage par l'enfant au même titre que la syntaxe ou le lexique et, de ce fait, relève de la grammaire à part entière. En un mot, la dimension expressive d'une grammaire ne peut plus être mise en débat aujourd'hui.*

*Ibid.*

Par ailleurs (*Ibid.* : 8), elle montre qu'

*un profil prosodique erroné greffé sur une construction verbale censée convoyer une émotion et une seule entrainera une incongruité sémantique (...) et une non reconnaissance ou tout au mieux une reconnaissance très perturbée de l'émotion sous-jacente.*

Cette dernière affirmation laisse entendre qu'une émotion ou attitude donnée peut éventuellement recevoir un certain nombre de patrons mélodiques, mais pas un nombre très étendu. Il y aurait bien un "code" de l'intonation, grammatical en ce que précisément une structure intonative ne peut être employée à la place d'une autre (ou du moins de n'importe quelle autre). **On peut à partir de là supposer que ces patrons sont identifiables, analysables et peuvent être comparés entre les différentes langues.**

e) La fonction syntaxique de la prosodie, ou intono-syntaxe, occupe quant à elle le devant de la scène, puisque la majorité des travaux en prosodie lui sont jusqu'à ce jour consacrés (voir notamment, entre autres, les travaux de Couper-Kuhlen, Di Cristo, Hirst, Martin, Mertens, Morel, Rossi, Selting). Plusieurs courants et modèles coexistent, qui décrivent et cherchent à théoriser le fonctionnement de l'intonation en ce qu'elle est régie ou non par la syntaxe de l'énoncé, en même temps qu'elle renforce, ou non, la structure syntaxique de l'énoncé.

Cette branche de la prosodie se donne pour objectif d'appréhender par exemple les modalités syntaxiques des énoncés (affirmation, interrogation, ordre etc.), ou la hiérarchie de leurs constituants (thèmes et rhèmes) et leurs démarcations. A ce rôle purement syntaxique

vient s'ajouter une fonction rythmique de la prosodie, qui s'organise autour de deux principes : i) celui d'eurythmie (espacement des accents et longueurs des groupes accentuels), et ii) de progression (augmentation progressive de la taille des constituants de l'énoncé). Ces travaux ont démontré que l'intonation est en partie commandée par la structure syntaxique de l'énoncé<sup>28</sup> (entre autres Rossi 1999 : chap. II et III; Lacheret-Dujour & Beaugendre 1999 : chap. 3; Mertens : 2008 etc.).

**Dans le cadre de ce mémoire, ce dernier point pourra poser des problèmes d'interprétation des courbes de la F0.** En effet, les énoncés que j'ai recueillis sont de structures syntaxiques très variées : l'intonation qui en ressortira sera donc à la fois conformée par la structure syntaxique et par l'expression des affects. A partir de là, on peut se demander quelle est la part de chacune dans la conformation de la courbe de F0. La syntaxe prévaut-elle sur l'expression émotive ou est-ce l'inverse? Peut-on trouver pour les différentes structures syntaxiques un (ou quelques) contour(s) intonatif(s) qui corresponde(nt) à un seul affect, à l'expression d'affects composites ou à plusieurs affects? La tâche est immense, puisque dans une langue, les structures d'énoncés sont particulièrement nombreuses - sinon infinies. On consultera à ce sujet avec profit Caelen-Haumont (2008a : 25-43).

C'est à partir de ces travaux en syntaxe de l'intonation qu'un certain nombre d'intono-syntacticiens se sont intéressés au rôle de l'intonation dans la distribution des tours de parole, rejoignant ainsi la pragmatique<sup>29</sup>.

#### 1.1.3.5 Où en est-on aujourd'hui?

Ce n'est que très récemment que la prosodie trouve ses lettres de noblesse. Ainsi, pour Lacheret-Dujour (2011a : 87) par exemple,

*contrairement au rôle qu'on a voulu lui attribuer pendant longtemps, la prosodie n'est pas une simple musique vocale qui accompagne librement les modulations de la pensée et qui, de ce fait, doit rester cantonnée dans la sphère du paralinguistique, elle s'inscrit d'emblée au cœur de la communication langagière.*

---

<sup>28</sup> Toutefois pour Bertrand, Porte & Sabio (2007 : 59) et sur un plan plus pragmatique, "le choix des patrons intonatifs met en jeu de manière cruciale la relation avec l'interlocuteur." C'est-à-dire que le contour mélodique choisi dépend, du moins pour partie, du contexte social et hiérarchique.

<sup>29</sup> Voir 1.1.4 Prosodie et pragmatique.

Pour preuve, la prosodie est initiatrice principale du langage, puisque c'est à travers elle que les nouveau-nés développent leurs premières relations intersubjectives (*Ibid.*); elle se trouve, selon une expression qui est chère à Lacheret-Dujour (2011b), "au cœur du verbal".

C'est notamment grâce aux nouveaux outils informatiques, (logiciels de mesures et d'analyse acoustiques, systèmes de synthèse vocale), que la recherche en prosodie a pu se développer. Ses applications intéressent particulièrement les neurosciences, les pathologies de la parole et la modélisation informatique ou Traitement Automatique du Langage (TAL) (Lacheret-Dujour 2011a : 87 et 93), dont Scherer *et al.* (1998) ou Aubergé (2002) par exemple pensent que les échecs ou la stagnation sont dus à l'absence de prise en compte de la part émotive dans la parole de synthèse<sup>30</sup>. Les travaux dans le champ de la prosodie appliquée se sont donc largement développés durant la dernière décennie, de même que l'importance que l'on octroie à la prosodie dans le processus de la communication parlée, ce dont atteste le nombre croissant de laboratoires qui y consacrent une partie de leur travail de recherche<sup>31</sup>. Caelen-Haumont (2008a : 38) note que dans le domaine de la prosodie, c'est parallèlement à l'évolution qui se produit dans le domaine du traitement de la parole que les prosodistes s'intéressent de plus en plus aux "facteurs sémantiques et pragmatiques" de la prosodie.

#### **1.1.4 Prosodie et pragmatique**

##### *1.1.4.1 Organisation du discours et tours de parole*

D'ailleurs, les interactionnistes eux aussi considèrent aujourd'hui qu'il faut prendre en compte la prosodie dans l'analyse des interactions, car elle influe sur le déroulement du discours. Bertrand, Portes & Sabio (2007 : 61) affirment notamment que

*selon qu'il utilise un contour montant ou descendant, le locuteur signale à son interlocuteur qu'il anticipe ou non une discussion de son propos. Dans le cas d'un montant simple ou d'un montant descendant, il signale en outre qu'il n'est pas prêt à céder sur sa position. Il l'envisage explicitement au contraire s'il utilise un contour descendant depuis la pénultième<sup>32</sup>.*

---

<sup>30</sup> Voir aussi 1.1.5.3 Evolution récente.

<sup>31</sup> Par exemple le laboratoire LaTTiCe de Paris 3 (CNRS), l'ICAR de Lyon 2, le FracTAL de Besançon, le GIPSA-Lab de Grenoble, le LPL d'Aix en Provence, etc.

<sup>32</sup> Pour plus de détails sur la question, voir Morel & Danon-Boileau (1998).

L'intonation influe aussi sur l'alternance des tours de parole (Bertrand & Portes 2012 : 19). Sacks, Schegloff & Jefferson (1974, cités par Bertrand & Portes 2012 : 26), ont défini le tour par des "unités de construction de tour" (*turn constructional unit*, TUC). Quand ils sont complets, ils laissent la place à un changement potentiel de locuteur (*transition relevant place*, TRP). Ford & Thompson (1996, cités par Bertrand & Portes 2012 : 26-27) ont montré l'importance de l'intonation dans les TRPs, puisqu'elle indique ou non la complétude de l'énoncé et donc la possibilité pour l'interlocuteur de prendre son tour de parole.

#### 1.1.4.2 Retombée dans le champ de la prosodie

La perspective interactionnelle a renouvelé l'approche des prosodistes, qui se cantonnaient auparavant à l'analyse de parole de laboratoire, très contrôlée et non interactionnelle<sup>33</sup>. Les études prennent désormais plus souvent pour support des situations de communication, des locuteurs et des langues variés, en s'inscrivant dans une approche de plus en plus comparative (Aubergé 2002; Audibert *et al.* 2008; Caelen-Haumont 2008 a et b; Bertrand & Portes 2012 : 27).

Toutefois, pour tempérer un peu cet optimisme, on constate que "l'univers sonore de l'interaction (...) reste(nt) le parent pauvre", et que "les dimensions vocales sont introduites avec plus de retenue dans des analyses [pragmatiques] qui sont par ailleurs pluridimensionnelles" (Traverso 2012 : 8; voir aussi Caelen-Haumont 2008a : 38).

#### 1.1.4.3 En analyse des conversations

En revanche, on ne voit nulle part l'analyste des conversations utiliser des mesures acoustiques qui l'aideraient à appréhender physiquement la prosodie afin de mieux cerner les actes de langage<sup>34</sup>. On a pourtant vu, et la question fait l'unanimité, que les énoncés en conversation véhiculent toujours une émotion ou une attitude, "très étroitement associées aux notions classiques en linguistique de 'force illocutoire' et de 'modalité épistémique'" (Lacheret-

---

<sup>33</sup> Voir aussi 1.2.4.

<sup>34</sup> En ce qui me concerne, je suis terriblement gênée par la quasi absence d'indications prosodiques dans les corpus transcrits - l'idéal serait de pouvoir "entendre" - : la plupart des transcriptions, à ce jour, n'indiquent que les pauses, plus ou moins longues, certaines montées et descentes de la voix en fin de segments, les chevauchements, une voix particulièrement forte par l'usage des majuscules, et quelques éléments paralinguistiques (rire, toux, etc.); or, la prosodie est encore beaucoup plus riche et nuancée que cela. D'autres protocoles ont cependant été proposés (par exemple dans l'article de Rittaud-Hutinet : Transcription de corpus oraux, *TRANEL 11*, 197-222, ou dans *Music of Everyday Speech*, Wennerstrom A. (2001), mais on se heurte à la complexité voire la lourdeur du codage, dues précisément à la richesse du matériau.

Dujour 2011 : 5, note 8). Si parfois l'on trouve des commentaires sur le contour ("prosodie montante", Moïse 2007 : 9) ou sur le rythme ("accélération du débit en fin de segment" *ibid.*), la plupart du temps, l'intonation n'est pas du tout prise en compte (Kerbrat-Orecchioni 2005), ou signalée très superficiellement et avec perplexité : "le ton (notion aussi importante que difficile à définir)" (Kerbrat-Orecchioni 2005 : 207); le "fameux 'ton', dont tout le monde s'accorde à reconnaître le rôle décisif (...) mais dont personne ne sait encore à vrai dire en quoi il consiste exactement : le mystère reste à peu près entier sur ce qu'est un "ton poli" (ou impoli)"; l'"accompagnement prosodique" (Kerbrat-Orecchioni 2002 : 11; 2010 : 57); l'énoncé "renforcé[e] par l'intonation" (Kerbrat-Orecchioni 2007 : 11; 2010 : 47); "la valeur insultante [d'un énoncé] dépend à la fois de son contenu sémantico-pragmatique, de sa réalisation linguistique, prosodique et mimo-gestuelle..." (Kerbrat-Orecchioni 2010 : 55). Tout cela montre bien que les interactionnistes ont bien sûr saisi la place centrale de l'intonation dans l'analyse des interactions, mais sans aller plus avant. Il est vrai que sur ce point, la collaboration entre les deux champs reste à réaliser<sup>35</sup>...

#### *1.1.4.4 La phonopragmatique (Rittaud-Hutinet 1995, 2005) : un cas à part*

Il existe une approche phonologique de la pragmatique, celle de Chantal Rittaud-Hutinet, qui a ouvert un nouveau champ de recherche, la "phonopragmatique". Il s'agit d'une approche nouvelle des interactions, en ce que précisément elle tâche de rendre compte de l'ensemble du processus interactif à travers les signes vocaux (Svoc), c'est-à-dire à travers la "couche vocale des énoncés". Rittaud-Hutinet définit ainsi la phonopragmatique : "Le domaine de la phonopragmatique est l'implicite de l'énonciation" (2005), en ce sens que la couche verbale ne fait pas toujours écho à la couche vocale, qui contient au travers des signes vocaux "une forme cryptée" de l'expression, et qui possède son sens propre indépendamment de la partie verbale. La phonopragmatique se donne pour objectif

*d'établir la régularité de la relation signifiant vocal/signifié pragmatique; de dégager les traits acoustiques que contiennent ces faits sonores signifiants; de déterminer leurs lieux d'apparition (...), les rapports qu'ils entretiennent avec le sens du support lexical (Ibid.).*

Elle note plus loin que les signifiés pragmatiques peuvent être exprimés par divers signes vocaux, dont certains sont synonymes. Le signe vocal révèle ou renforce la valeur des

---

<sup>35</sup> Voir aussi 1.1.4.5. Evolution récente.

actes illocutoires, appréciatifs, informatifs, et interpersonnels, il intervient au niveau de la morphosyntaxe, des rapports de place des interlocuteurs, de la structuration conversationnelle et du travail de formulation.

Cette approche, intéressante au plus haut point pour mon propos et, me semble-t-il, la pragmatique en général, n'a pas suscité d'émules... pour au moins trois raisons : une certaine avance sur son temps<sup>36</sup>, une formulation des contenus théoriques aussi bien qu'expérimentaux hautement complexifiée, ainsi qu'une certaine étanchéité entre les champs de la recherche, assurément regrettable. Qu'en est-il aujourd'hui?

#### *1.1.4.5 Evolution récente*

Depuis quelques temps déjà, les prosodistes anglo-saxons se sont lancés dans la théorisation et l'analyse de la conversation spontanée et de ses contenus illocutoires (voir par exemple Bolinger (1972), O'Connor & Arnold (1973), Halliday (1994) et Wichmann (2002), cités par Caelen-Haumont 2008b : 146, ou Wennerstrom 2001 : 131 *sq.*). En France en revanche, le phénomène est beaucoup plus récent. Ainsi Caelen-Haumont (2008a : 38) note que "l'intérêt [de la prosodie] se déporte de plus en plus de l'étude de la structuration syntaxique vers la prise en compte des facteurs sémantiques et pragmatiques". Elle consacre deux volumes (2008 a et b) à la prosodie appliquée à l'analyse conversationnelle, avec notamment la notion de "mélisme"<sup>37</sup> qu'elle préfère au terme de "focalisation"<sup>38</sup> auquel elle reproche son inexactitude et son imprécision (Caelen-Haumont 2002 : 14). Le mélisme, qui utilise les registres aigu et supra-aigu, exerce "indépendamment du linguistique, une fonction perlocutoire par l'action sur autrui, volontaire, consciente ou pas", forme sous laquelle prennent corps "les indices prosodiques spécifiques portant la marque de [l'] investissement" du locuteur (2008b : 173).

---

<sup>36</sup> Voir 1.1.4.5.

<sup>37</sup> Le mélisme est une "figure mélodique sur l'étendue du mot, telle que le nombre de notes (ou de tons) perçues est supérieur au nombre de syllabes du mot. Appliqué au domaine de la langue, ce concept désigne dans notre étude les excursions mélodiques propres à la parole".

<sup>38</sup> "Le focus mélodique s'inscrit dans la forme du mélisme. Pour notre part, nous n'employons pas le terme de « focalisation » pour 2 raisons : d'une part ce terme est inadéquat, car il réfère à un processus binaire (focal / non focal) alors que le processus est en réalité scalaire (Caelen-Haumont, 1991 ; Terken, 1991), d'autre part le terme est flou car il confond les plans acoustique, sémantique et pragmatique. S'agissant de la forme, nous emploierons donc le terme de mélisme, s'agissant du contenu, nous parlerons de valeur sémantique, pragmatique, s'agissant des croyances attachées à ces valeurs, et de leur effet sur autrui (Austin, 1991), nous parlerons de « visées perlocutoires »" (Caelen-Haumont 2002 : 14).

Lacheret-Dujour (2011a : 92) rappelle quant à elle que le locuteur communique en fonction d'un but, et que dans la mise en forme de l'information, il exprime un point de vue. Pour Victorri et Fuchs (1996, repris par Lacheret-Dujour *ibid.*), un énoncé de base serait constitué de "la description d'une scène", et de la "donnée d'un point de vue sur cette scène" qui permet un partage de la "vision" avec l'interlocuteur. On est ici au cœur de la construction de l'interaction, de la co-énonciation; le marquage du point de vue dans l'énoncé explique alors vraisemblablement nombre d'accents d'insistance ou de contours terminaux à valeur illocutoire (doute, évidence, etc.) (Lacheret-Dujour *ibid.* : 93). Morel & Danon-Boileau (1998; voir aussi Bertrand, Portes & Sabio 2007 : 61) affirment d'ailleurs que les locuteurs, selon qu'ils utilisent telle ou telle intonation, anticipent la réaction, consensuelle ou non, de leur interlocuteur. Nous ne sommes pas loin dès lors de l'expression des émotions et des attitudes, et nous allons voir maintenant dans quelle mesure les valeurs illocutoires des énoncés peuvent recouper les notions d'émotion et d'attitude.

### **1.1.5 Prosodie des émotions et des attitudes**

Les prosodistes se sont, jusqu'à très récemment, peu intéressés à l'expression vocale des émotions : "Phénomène langagier majeur qui sous-tend la communication humaine au quotidien, l'expressivité vocale est pourtant restée en marge des préoccupations dominantes dans les recherches sur le langage." (Lacheret-Dujour 2011b : 1).

Ceux qui se sont principalement emparés de la tâche affirment : "de nombreuses études ont démontré que les expressions vocales permettent d'identifier l'état émotionnel d'un individu en l'absence d'autres indices." (Bänziger & Scherer 2001 : 119), et Fónagy (2003 : 15-16) : la " transmission vocale des émotions telles que la colère, la haine, la peur, la joie, la surprise, ou des attitudes telles que le désir, la menace, la supplication, la persuasion ou l'ironie, peuvent souvent se passer de mots." Même, des expériences linguistiques (Uldall 1964, cité par Morel & Bänziger 2004 : 89), psychologiques (Papousek 1994, *ibid.*) et neuropsychologiques (Ross 1981, van Lanker & Sidtis 1992, Heilman, Bowers, Speedie & Coslett 1984, *ibid.*) laissent penser que l'intonation contribue à la communication émotionnelle indépendamment du contenu linguistique des énoncés.

### *1.1.5.1 Introduction*<sup>39</sup>

La production et la perception des propriétés acoustiques des énoncés émotifs, prises indépendamment du contenu lexico-syntaxique, sont les deux principales orientations de la recherche actuelle dans ce domaine.

La plus grande partie des études sur la question postule qu'à une émotion particulière correspondent des paramètres acoustiques donnés. Bachorowski souligne toutefois à ce sujet une contradiction entre les difficultés éprouvées par les acousticiens à caractériser les différences acoustiques entre les émotions, alors que les testeurs perçoivent eux relativement facilement ces mêmes émotions lors des écoutes perceptives : "les caractéristiques qui joueraient ce rôle ne sont pas actuellement bien identifiées" (Morel & Bänziger 2004 : 89). Le champ est donc largement ouvert.

Globalement, au cours des études qui ont été menées, les émotions fortement activées (colère, peur, angoisse, joie) sont caractérisées par une augmentation de la moyenne de la F0, une amplitude plus grande dans la variation de la F0 et une intensité plus importante, par opposition aux émotions faiblement activées (comme la tristesse), où les mêmes paramètres sont plus faibles. Fónagy (1971, 1983, 1986) puis Léon (1993) ont pu mettre en évidence quelques stéréotypes prosodiques émotifs et attitudeaux, c'est-à-dire des caractéristiques acoustiques attachées à quelques émotions et/ou attitudes.

### *1.1.5.2 Deux approches qui ont marqué la recherche en prosodie de l'expressivité des émotions et attitudes*

Le champ a jusqu'à il y a quelques années principalement été occupé par les équipes d'Ivan Fónagy et de Klaus Scherer.

#### **a) Ivan Fónagy**

Ivan Fónagy est le grand précurseur de la phonétique acoustique expressive. Bien qu'évidemment ces questions l'aient interrogé dès le début de sa recherche à la fin des années 30, (Guimbretière 2000b), c'est à partir de 1969 que Fónagy entre définitivement dans l'étude acoustique de l'expression prosodique des émotions. L'ouvrage qui l'a fait connaître du grand

---

<sup>39</sup> Bachorowski 1999, sauf indication.

public, *La Vive Voix* (1983), s'attache notamment à décrire les caractéristiques prosodiques de certaines émotions comme la colère - dont il identifie 4 types -, la tristesse, et la joie, ou d'attitudes comme la tendresse, l'ironie ou la menace. Il en décrit les accents, le débit, l'intensité, la hauteur et la variation de la F0, les mimiques buccales (c'est-à-dire les mouvements de l'appareil phonatoire), et laryngées (le comportement des cordes vocales, plus ou moins rapprochées ou comprimées, des muscles expiratoires et glottiques, plus ou moins contractés). Fónagy, Bérard & Fónagy (1983) mettent en évidence l'existence de "clichés mélodiques", du type "Eh ben dis donc!" ou bien "Et moi alors?", dont ils observent que ce sont précisément leur intonation très stéréotypée qui "trahissent" les étrangers auprès des locuteurs natifs français. S'appuyant sur la psychanalyse, il soutient que les choix prosodiques du locuteur traduisent, au moyen de l'appareil phonatoire, des émotions ou des pulsions plus ou moins inconscientes; ainsi dans l'expression de la colère, la tension des muscles correspondrait à la tension de l'esprit, la répétition de l'accent serait la marque de coups symboliquement portés à l'adversaire, etc. (Fónagy 1983 : 136 *sq.* et 152 *sq.*).

#### b) Scherer

Klaus Scherer a quant à lui abordé l'expression émotive sous l'angle de la psychologie; c'est pourquoi ses travaux, dès la fin des années 80, s'appuient d'abord sur la perception (*appraisal*) des manifestations vocales et faciales des émotions (Scherer 1988, Scherer *et al.* 1998, Bänziger & Scherer 2001, etc.), pour ensuite en venir à leur analyse acoustique. Avec Banse (1996), il est le premier<sup>40</sup> à outrepasser le cadre des six émotions de base<sup>41</sup> pour faire l'analyse de quatorze<sup>42</sup> émotions ou attitudes dont il a pu pour certaines identifier des profils acoustiques très reconnaissables<sup>43</sup>. Douze des quatorze affects ont pu être identifiés par les testeurs avec un taux élevé de précision. D'une manière générale, les expérimentations perceptives de Scherer ont pu montrer que **i) les jugements d'auditeurs inexperts sont fiables en ce qui concerne la reconnaissance d'émotions et/ou d'attitudes, ii) "les indices vocaux convoient l'affect de façon directe et indépendante du contexte" (1998).**

---

<sup>40</sup> Fónagy (1978) avait mené une expérimentation perceptive sur dix émotions/attitudes jouées par une actrice, dans l'objectif de tester un nouvel appareil de filtrage des données verbales, le Vocoder (laryngographe).

<sup>41</sup> Voir 1.1.2.

<sup>42</sup> *Hot anger, cold anger, panic/fear, anxiety, despair, sadness, elation, happiness, interest, boredom, shame, pride, disgust, contempt.*

<sup>43</sup> Mais pour lesquelles il fournit seulement une liste de chiffres sous forme de tableau, sans donner de représentations (courbes, spectres, etc.).

### 1.1.5.3 Evolution récente

Dans la suite de ce qui a été dit plus haut (1.1.3.5) à propos du développement des travaux en prosodie, les études plus spécifiques sur la parole expressive ont connu elles aussi un très large essor au cours de la dernière décennie (Aubergé 2002 : 271; Audibert, Aubergé & Rilliard 2008 : 1), et pour les mêmes raisons.

On considère aujourd'hui que "la prosodie est le vecteur privilégié des émotions de la parole" ou encore qu'"elle est le siège de l'expression directe des émotions, du codage des attitudes et des stratégies expressives pour un même matériel acoustique." (Aubergé 2002 : 264), et que "l'attitude est la fonction directe par laquelle le locuteur informe de son point de vue" (*Ibid.* : 269). De même, l'expression langagière des émotions, et donc sa partie vocale, intéresse au plus haut point la recherche clinique, notamment dans le domaine des pathologies de la parole (Lacheret-Dujour 2011a : 93), ainsi que la recherche en modélisation du langage (Aubergé 2002 : 271)<sup>44</sup> : "Pour les systèmes de synthèse de la parole, la prosodie est un enjeu important car elle représente une grande part de ce qui est humain dans la parole." (Morel & Bänziger 2004 : 91). La modélisation de la parole n'a donc plus seulement pour enjeu de proposer une reproduction acceptable du langage humain; elle doit, pour y réussir, prendre en compte les compétences interactionnelles, émotionnelles et affectives des sujets parlants; c'est-à-dire que les performances de la modélisation (de 3<sup>ème</sup> génération) ne pourront être améliorée que par une optimisation de la prise en compte de la prosodie expressive "de nature pragmatique" (Lacheret-Dujour & Morel 2011).

### 1.1.5.4 Liens entre expression des affects et impolitesse

Si, comme nous l'avons vu en 1.1.1.2, l'impolitesse consiste en la mise en œuvre de stratégies dont le but est de provoquer une rupture dans le contrat conversationnel ou de mettre en péril les faces de l'interlocuteur, alors elle doit être mise en relation avec l'expression d'émotions ou d'attitudes que l'on peut d'une manière très généraliste appeler "négatives". Ces affects sont ceux qui vont nuire d'une façon ou d'une autre à l'interlocuteur : colère bien sûr, mais aussi mépris, dégoût, menace, agacement, irritation, incrédulité etc. L'expression de l'impolitesse et d'affects menaçants pour les faces de l'interlocuteur se recoupe donc ici. A l'appui de cette hypothèse, je citerai Lacheret-Dujour (2011b : 5 et 5 note 8) qui souligne que "les attitudes contrôlées et codées linguistiquement au plan de

---

<sup>44</sup> Voir aussi 1.1.2.4 Les affects : un sujet d'actualité, et 1.1.3.5 Où en est-on aujourd'hui?

l'expression (ex. évidence, ordre, doute, etc.)" sont "très étroitement associées aux notions classiques en linguistique de **force illocutoire** et de **modalité épistémique**<sup>45</sup>." Aubergé (2002 : 265) va plus loin encore : en s'appuyant sur les travaux de Damasio (1994 : 9), elle affirme que "le traitement cognitif rationnel qu'est la production d'énoncés langagiers ne serait pas seulement porteur des émotions du locuteur qui motivent ses énoncés, il serait décidé inévitablement en fonction de celles-ci."

L'expression, plus ou moins consciente, de l'(im)politesse, peut emprunter plusieurs canaux : le canal verbal, le canal mimo-gestuel et proxémique, voire le canal physiologique (respiration, changement du teint, etc.). Au sein du canal verbal, la prosodie<sup>46</sup> est capitale dans l'expression des émotions et attitudes. **Si la pragmatique a pour objet d'analyser le discours en interaction, et que le discours en interaction ne peut être que l'expression d'une émotion, d'une attitude, d'une intention ou d'un point de vue, la pragmatique a aussi pour objet cette même expression. Si la prosodie peut rendre compte de cette expression sur le plan acoustique, la pragmatique doit en tirer bénéfice.**

Toutefois, la caractérisation de profils intonatifs (im)polis ne peut pas s'envisager seulement à partir de l'angle bien trop large de l'(im)politesse : elle doit passer par le crible de l'analyse d'émotions/attitudes impolies variées qui restent à définir précisément.

#### *1.1.5.5 Liens entre prosodie des affects et impolitesse :*

En France et à ce jour, en dehors des quelques approches présentées en 1.1.4, l'analyse acoustique appliquée à la pragmatique n'en est encore qu'à ses débuts. La description prosodique des affects quant à elle semble tournée uniquement vers les neurosciences et le traitement automatique du langage; si bien que ses liens avec la pragmatique sont encore extrêmement ténus. Les actes de parole commencent à être pris en compte mais de manière encore généraliste : si des jalons théoriques ont été posés, loin s'en faut que les modèles aient été mis à l'épreuve d'une analyse acoustique précise (Caelen-Haumont 2008 a et b; De Abreu *et al.* 2009 : 64). Quels sont par exemple, pour le français, les indices acoustiques de l'évidence, du doute, de l'autorité, du jugement et autres affects? **Loin s'en faut donc que l'(im)politesse ait été abordée sous cet angle : le champ reste entièrement à explorer.**

---

<sup>45</sup> Mis en gras par moi.

<sup>46</sup> Voir en 1.1.5.3.

On ne peut que souhaiter que, dans les années à venir, l'étude de la prosodie des affects trouve sa place dans l'analyse des interactions grâce à un rapprochement des deux champs.

### **1.1.6 Prosodie, politesse et interculturel**

#### *1.1.6.1 L'apprentissage/enseignement de la prosodie*

Si les travaux théoriques ne manquent pas sur le sujet (par exemple Brazil *et al.* 1980; Chun 2002; Santiago & Delais-Roussarie 2012), l'apprentissage/enseignement de l'intonation mériterait lui d'occuper une place plus importante dans le curriculum des langues étrangères, domaine dans lequel il est relativement peu pris en compte, par manque de temps tout d'abord : ce n'est qu'une fois que les contenus lexicaux, syntaxiques, orthographiques, culturels ont été dispensés, que l'on accorde parfois un peu de temps à la phonétique ... dont la prosodie n'est qu'une partie ... D'autre part, les enseignants perçoivent généralement la phonétique comme un domaine "inaccessible, scientifiquement réservé à une élite" (Abry & Veldeman-Abry 2007 : 7). D'ailleurs, rares sont encore les méthodes de langue qui proposent un travail de phonétique articulatoire ou prosodique :

*Si les auteurs des certifications l'ont compris, les concepteurs de manuels restent encore un peu en retrait puisqu'ils n'intègrent une rubrique 'phonétique' que pour les niveaux 1 et 2, c'est-à-dire pour les niveaux A1 et A2 du référentiel pour les langues du Cadre européen. La rubrique 'phonétique' disparaît en effet des livres 3 et 4, c'est-à-dire ceux des niveaux B1 et au-delà, ce qui est pour le moins étonnant dans la mesure où les difficultés intonatives, consonantiques et vocaliques des apprenants persistent et qu'ils seront notés sur leur prononciation.*

(Abry & Veldeman-Abry 2007 : 8)

De là à dire que l'intonation n'est jamais travaillée par l'apprenant en classe de langue, il s'en faut de peu. Quelques rares méthodes spécifiques voient cependant très tôt le jour, par exemple *Rythme et intonation* (Léon 1965), ou *Introduction à la méthode verbo-tonale de correction phonétique* (Renard 1971). *Plaisir des sons* (Kaneman-Pougatch & Pedoya-Guimbretière 1991) fait travailler implicitement l'intonation à partir de nombreuses activités de répétition. Plus récemment, *La Phonétique : audition, prononciation, correction* (Abry & Veldeman-Abry 2007) consacre seize pages à l'intonation, mais s'adresse à tous les niveaux du CECRL; *Les 500 exercices de phonétique, niveau A1/A2* (Abry & Chalaron 2010) n'accorde que neuf pages à la prosodie, mais le travail est ensuite systématiquement repris

dans les sous-parties. Un second tome est paru en 2011 pour les niveaux B1/B2, conçu selon la même maquette.

### *1.1.6.2 De la nécessité d'enseigner l'intonation pour faciliter la communication interculturelle*

Gumperz, envisageant dès 1982 (1989a : 18) les rapports intergroupes (qu'ils soient ethniques, linguistiques ou sociaux), affirme que "l'accomplissement de (ce)s tâches communicatives<sup>47</sup> sous forme de pratiques sociales est variable d'un point de vue culturel", variation qui inclut "les différentes façons de parler : l'utilisation d'un ensemble distinct de conventions linguistiques inconscientes (telles que le ton de la voix) pour accentuer, signaler les connexions logiques et indiquer la portée de ce qui est en train de se dire (...)" (*Ibid.* : 19). Plus loin (*Ibid.* : 24), il remarque que "la plupart des individus interprètent la façon de parler d'autrui en se fondant sur leurs propres conventions." Là où ils auraient dû pouvoir résoudre des problèmes "normaux" de la vie quotidienne, la communication échoue et ce sont des facteurs linguistiques qui provoquent ruptures de contrat conversationnel et malentendus. Chun (2002 : 86) cite un certain nombre de travaux qui ont très tôt établi le transfert par les apprenants de l'intonation de la langue première (L1) vers la L2.

*One factor in L2 acquisition that results in a so-called 'foreign accent' in L2 is the transfer of patterns of pronunciation from L1. Studies of both segmentals, such as Bieritz (1974), Fledge (1981), Fledge and Port (1981) and Blitz and Weiss (1969), and prosodic phenomena, e.g., Pürschel (1975) and White (1981), have shown that some recurrent phonetic patterns in learners' productions can be accounted for by "interference" or "negative transfer" from L1.*

Shochi *et al* (2009 : 48) ont ainsi relevé des cas de "prosodic false friends".

Suite à quoi, comme tout message véhicule un affect, une intention, un point de vue perceptibles par la prosodie, et si tout échange verbal constitue une menace pour les faces des interlocuteurs, un usage erroné de la prosodie, calquée à partir de la langue première, peut dans le cas de la communication en langue étrangère très rapidement conduire au conflit<sup>48</sup>. De même qu'"une prononciation erronée peut gêner la compréhension d'un message car souvent il

---

<sup>47</sup> Raconter, expliquer, argumenter, accentuer, instruire, diriger (*Ibid.* : 18).

<sup>48</sup> A noter que le processus est pertinent pour tous les rapports intergroupes, y compris pour les sous-groupes sociaux appartenant à un même hyper-groupe (par exemple les sous-groupes de l'hyper-groupe "société française").

n'y a qu'un son qui porte la distinction" (Abry & Veldeman-Abry 2007 : 8), de même une intonation erronée produira confusion et incompréhension (Lacheret-Dujour 2011b : 8) : que l'on me parle sur un ton qui ne convient pas, ne correspond pas à mes codes (socio)culturels ou linguistiques, j'aurai tôt fait de me vexer, de me sentir offensé(e) etc., et de conclure que mon interlocuteur est un malotru infréquentable! La porte est ouverte à l'incompréhension de l'autre, au mépris, au rejet, à l'exclusion, à la haine (Gumperz 1989a : 29). Comment ne pas imaginer en effet que certaines intonations dans une langue donnée soient acceptables voire recommandées, alors que dans une autre l'effet produit mette l'interlocuteur dans la gêne, le vexe ou le mette en colère?

Enfin, si la prosodie fait l'objet d'un apprentissage par le jeune enfant au même titre que le lexique et la syntaxe (Guzman 1973 : 14; Lacheret-Dujour *et al.* 2000), a fortiori si on lui octroie un statut grammatical<sup>49</sup>, la nécessité de son apprentissage par les apprenants de langue étrangère s'en trouve renforcée.

Suite à quoi les spécialistes s'accordent pour dire qu'on ne peut faire l'économie de l'enseignement de la prosodie en langue seconde (Wennerstrom 2001 chap. 9; Chun 2002; Shochi *et al.* 2006 et 2009; Lu *et al.* 2012), car hormis des capacités d'imitation naturelles favorisées par une oreille fine, l'intonation ne va pas de soi pour le locuteur de langue étrangère (Abry & Veldeman-Abry 2007 : 8).

### 1.1.6.3 Exemples<sup>50</sup>

Donnons quelques exemples de conventions prosodiques divergentes.

Bien que dans *La Vive Voix* (1983 : 136 par exemple<sup>51</sup>), Fónagy ait eu la tentation de rechercher des universaux des caractérisations acoustiques des émotions ou attitudes, leur observation dans différentes cultures a montré qu'elles sont loin d'être partout codées de la même manière (Léon 1993 : 82). Par exemple en langue quiché du Guatemala, la labialisation indique la flatterie, en zuni (sud ouest des Etats-Unis) le reproche; en tzeltal (Mexique), on

---

<sup>49</sup> Voir 1.1.3.4 d.

<sup>50</sup> J'ai choisi ici quelques exemples significatifs, qu'il aurait été possible de multiplier, notamment à partir d'études en prosodie des affects : Guzman 1973 : 6, 38, 59 *sq.*; Scherer *et al.* 2001; Shochi *et al.* 2009; Santiago & Delais-Roussarie 2012 etc.

<sup>51</sup> Mais la tendance universaliste sous-tend une grande partie de l'ouvrage, étant donnée notamment l'interprétation psychanalytique de la prosodie faite par Fónagy, censée s'appliquer à l'être humain dans son ensemble, quelle que soit sa culture.

utilisera une voix de fausset pour un discours honorifique, alors que la voix cassée indiquera la commisération (*Ibid.* : 83).

Moins loin de nous, et en contradiction avec ce qui a été relevé ci-dessus, Fónagy (1983 : 226) rapporte qu'une stagiaire soviétique, transposant l'intonation russe de la question sur ses phrases interrogatives hongroises, s'étonnait de ne pas recevoir de réponse : c'est que son intonation, pour ses interlocuteurs hongrois, était celle de l'affirmation...

Ailleurs, Fónagy (Guimbretière 2000b : 152) montre que c'est souvent une intonation erronée qui permet à un francophone natif de reconnaître un locuteur non natif, notamment dans ce qu'il appelle les "clichés mélodiques", énoncés extrêmement fréquents, repris selon la même variation de la F0 par tous les natifs (voir 1.1.5.2. a).

Gumperz (1989b : 120 *sq.*) fait une description comparative d'une intonation anglaise occidentale et d'une intonation anglo-indienne relevées sur des énoncés identiques. Il observe que l'intonation anglo-indienne est nettement influencée par la prosodie de langues du nord de l'Inde : il y a transfert prosodique de la langue première à la langue seconde, ce qui a pour résultat soit de déconcerter l'interlocuteur anglais, soit de l'amener à mésinterpréter ce qu'on lui dit. Ce transfert peut expliquer "pourquoi les rencontres entre locuteurs en anglais asiatique et en anglais occidental témoignent de difficultés de communication." (*Ibid.* : 130).

Enfin, tout récemment, Boula de Mareüil, Rilliard, Mairano & Lai (2012) ont montré que la prosodie corse des questions totales était transposée par les Corses telle quelle sur les énoncés en langue française du même type, s'écartant ainsi du français standard. Il est vraisemblable que si l'on généralisait ce type d'étude aux différentes prosodies régionales, on relèverait une influence certaine des langues qui y furent parlées peu ou prou jusqu'au début du 20<sup>e</sup> siècle, réduisant sérieusement l'étendue dudit "français standard". De même, on s'est intéressé à la description de l'accent dit "des cités" ou des "banlieues", pour tâcher de déterminer dans quelle mesure les langues premières de leurs habitants influencent la prosodie du français qui y est parlé (voir par exemple Fagyal 2003 ou Lehka-Lemarchand 2007).

#### 1.1.6.4 Conclusion

Il semble que, d'une manière générale dans la production orale en langue étrangère, les intonations de la L1 sont spontanément transférées par les locuteurs "non avertis" en L2<sup>52</sup>. Cela ne peut que favoriser l'apparition de malentendus interculturels.

### 1.1.7 Synthèse

#### 1.1.7.1 Incertitudes et questionnement

1. L'analyse acoustique de corpus écologiques enregistrés pose un problème méthodologique : comme nous l'apprend l'intono-syntaxe, "certaines constructions syntaxiques sont en elles-mêmes porteuses de schémas expressifs, et de ce point de vue contraignent les constructions prosodiques" (Lacheret-Dujour & Morel 2011). On peut aussi penser aux clichés mélodiques de Fónagy *et al.* (1996). De plus, le discours spontané produit des énoncés syntaxiquement très divers : l'intonation syntaxique peut venir brouiller l'analyse de la variation expressive de la F0. Comment alors faire la part de ce qui, dans la prosodie, revient à la syntaxe, et de ce qui ressortit à l'impolitesse?<sup>53</sup> Je tâcherai tout de même, en tenant compte des réserves que je viens d'émettre, d'identifier **si possible** des énoncés déclarés "impolis" aux profils intonatifs approchants, en tenant compte des similitudes, divergences ou convergences présentées par ces contours.

2. La notion d'impolitesse est trop large pour être abordée en tant que telle sous l'angle de la prosodie : elle est connexe à l'expression d'un certain nombre d'affects "négatifs" - émotions/attitudes menaçantes pour les faces - qui restent à définir précisément et dont la plupart n'ont pas été décrites sur le plan acoustique. Les corpus de parole spontanée que j'ai recueillis peuvent potentiellement fournir un grand nombre de ces émotions/attitudes. Il s'agira, **si possible**, de les classer sous une dénomination commune. La validation de ce classement par un panel de locuteurs natifs restera à réaliser.

#### 1.1.7.2 Hypothèses et principes théoriques

Voici les **hypothèses** et **principales positions théoriques** sur lesquelles s'appuie ce mémoire.

---

<sup>52</sup> Pour un avis plus nuancé, voir Santiago & Delais-Roussarie 2012.

<sup>53</sup> Voir aussi 1.1.3.4 e.

1. Par "intonation", j'entends "variation de la fréquence fondamentale".
2. Il est possible dans une langue donnée, ici le français, d'identifier, analyser, comparer des groupes intonatifs aux caractéristiques intonatives et/ou prosodiques proches.
3. La prosodie participe de l'expression des affects, l'intonation neutre n'existe pas car tout énoncé est émotif ou tout au moins attitudinal. A ce titre, l'analyse prosodique a un rôle à jouer en analyse interactionnelle, notamment dans le domaine de l'(im)politesse, comprise comme l'analyse du "travail des faces".
4. La prosodie d'une langue s'acquiert, au même titre que le lexique, la morphologie ou la syntaxe : elle doit être travaillée par les apprenants de langue(s) étrangère(s) afin notamment d'éviter les malentendus interculturels.

Ce mémoire se donne pour **objectifs** d'essayer

- i) d'établir des caractéristiques prosodiques d'énoncés français produits par des locuteurs natifs et déterminées par des auditeurs naïfs français comme impolies;
- ii) de cerner ce qui dans la prosodie, et notamment dans les variations de la fréquence fondamentale, peut contribuer à identifier la violence verbale et participer de "la montée en tension verbale", afin de
- iii) contribuer à identifier l'origine de malentendus interculturels dus à l'intonation, et à améliorer l'enseignement/apprentissage de l'intonation du français langue seconde.

### *1.1.7.3 Retombées possibles*

Les orientations récentes de la recherche en prosodie, et plus particulièrement en prosodie de l'expression affective, montrent un fort consensus de la part de chercheurs sur l'intérêt d'une telle recherche qui porte au premier chef sur : i) la neurophysiologie et la neuropsychologie, ii) la modélisation du langage et le traitement automatique des langues.

Il faut cependant ajouter deux autres champs d'application : i) celui de la pragmatique avec l'(im)politesse, ii) celui de la didactique des langues étrangères, ici du FLE. Sont directement concernés l'enseignement/apprentissage de la phonétique et de l'interculturel.

## 1.2 Choix méthodologiques

J'ai été amenée à opérer certains choix méthodologiques avant même de débiter la mise en place de l'expérimentation dont le déroulement s'est rapidement imposé à mon esprit (voir 1.2.8). Fallait-il filtrer les données du corpus? Quels éléments micro ou macro contextuels prendre ou ne pas prendre en compte pour l'analyse? Quels paramètres acoustiques choisir? Pourquoi travailler sur du matériau écologique, et de quel point de vue : celui de l'émetteur ou celui du destinataire? Comment comparer de manière pertinente les variations d'intonations de voix aux phonostyles bien différents?

### 1.2.1 Filtrage sémantico-lexical des énoncés

J'avais pensé dans un premier temps filtrer les unités sonores pour supprimer totalement l'influence sémantique dans l'appréciation de la qualité (im)polie des énoncés, tout en conservant la structure prosodique. Fónagy (1981, cité par Fónagy 2003 : 15-16) ou encore Scherer *et al.* (1998) affirment en effet que le contenu affectif des énoncés est conservé même si on les rend lexicalement incompréhensibles. Cependant, bien qu'il eût été très important<sup>54</sup> que seule la prosodie et, dans l'analyse prosodique, particulièrement l'intonation, fussent prise en compte par les testeurs, **j'ai finalement choisi de ne pas filtrer la partie syntaxico-sémantique des énoncés**, car en l'état actuel de mes connaissances du logiciel Praat et de ses paramètres de filtrage, le résultat sonore aurait par trop déconcerté les testeurs, qui se seraient retrouvés dans l'incapacité de juger correctement de l'(im)politesse des énoncés. Un réglage du logiciel était bien entendu possible, mais le temps imparti à la recherche dans le cadre d'un Master 2 aurait été ici insuffisant - à mon très grand regret.

### 1.2.2 Contexte

Les pragmaticiens s'accordent à dire que, si l'on veut obtenir une analyse pertinente d'une interaction, on se doit de "prendre en compte (...) tous les éléments pertinents du contexte"<sup>55</sup> (Kerbrat-Orecchioni 2007 : 3), car un énoncé dont le contenu sémantique et prosodique est perçu comme poli hors contexte, peut en contexte devenir tout à fait impoli

---

<sup>54</sup> Voir 2.2.1.2 Passation des tests - Consigne non écrite.

<sup>55</sup> Par "contexte", on entend les informations linguistiques - c'est le contexte discursif, endogène, séquentiel ou "cotexte" et qui englobe tout ce qui, verbal et non verbal, fournit des indications sur le contexte linguistique des énoncés. On entend aussi par "contexte" le contexte externe, ou exogène, qui s'étend de la situation immédiate de l'énoncé - cadre spatio-temporel, objectif de la conversation etc. - jusqu'au cadre socioculturel et sociétal - relations de place entre les interlocuteurs, classe sociale, culture partagée, etc.

(Kerbrat-Orecchioni 2005 : 208), au point que pour certains (Fraser 1990, Watts 2003, cités par Kerbrat-Orecchioni 2005 : 208), la notion d'(im)politesse serait tout entière contenue dans l'(in)adaptation de l'énoncé aux normes voulues par le contexte, en dehors duquel aucune séquence ne peut être qualifiée d'(im)polie.

Toutefois, les recherches menées sur la manifestation prosodique des émotions, Scherer, Johnstone & Sangsue (1998) par exemple, soutiennent que les indices vocaux non verbaux convoient l'affect de façon directe et indépendante du contexte. Ils se détachent donc complètement de ce dernier, pour ne se préoccuper que des phénomènes acoustiques d'énoncés "émotifs".

**C'est dans cette lignée-là que je me place :** comme je l'ai dit plus haut, il s'agit de déterminer dans quelle mesure une intonation, et elle seule, est intuitivement perçue comme (im)polie par le locuteur natif, et ce en dehors de tout contexte. De même que Kerbrat-Orecchioni (2005 : 208) affirme que certains énoncés sont par nature (im)polis, j'aimerais pouvoir montrer que certaines intonations sont elles aussi par nature (im)polies.

### **1.2.3 Mimique faciale et gestuelle**

**Pour des raisons identiques, je ne m'appuierai pas sur le mimo-gestuel,** données paralinguistiques qui, dans la mesure où elles précisent le contenu sémantico-syntaxique et prosodique, sont d'une grande utilité pour le décryptage des interactions. Fónagy (1983 : 52-53) a d'ailleurs pu prouver que plusieurs mimiques faciales, voire parfois physiques, sont "audibles" par les testeurs, qui sont capables de les reproduire à la simple écoute des énoncés. Les gestes et les expressions faciales soutiennent sans aucun doute l'intonation, mais ce sont précisément des caractéristiques de l'intonation que je cherche à saisir, indépendamment de toute autre manifestation expressive.

### **1.2.4 Nature des données**

Longtemps, qu'ils aient travaillé sur l'expression émotive vocale ou en phonologie des interactions, les acousticiens ont pris pour support des données orales non spontanées, "with isolated utterances in carefully controlled laboratory conditions, and consequently the extent to which the results will be applicable to other speech situations and conditions remain to be seen." (Di Cristo 1998 : 218).

La plupart du temps, ce sont des acteurs qui imitent telle émotion, telle attitude (Fónagy 1978 : 35, 1983, 2003 : 16 etc.; Banse & Scherer 1996; Scherer & Johnstone 2003 : 436; Bänziger & Scherer 2001: 119). Scherer & Johnstone (2003 : 436) soulignent d'ailleurs les inconvénients des corpus écologiques : le nombre de locuteurs serait la plupart du temps insuffisant, les énoncés produits le seraient dans de mauvaises conditions d'enregistrement, et ils seraient trop courts; de plus, le locuteur ne serait pas toujours capable d'identifier l'émotion qu'il a eu l'intention de produire. En conséquence de cette position théorique, Scherer *et al.* (1998) ou Bänziger, Grandjean, Bernard, Klasmeyer & Scherer (2001 : 29 *sq.*) vont jusqu'à utiliser des techniques de synthèse digitale de la parole pour mieux évaluer l'influence d'indices prosodiques spécifiques et de la qualité de la voix, dont ils font varier les paramètres, sur l'impression de l'auditeur.

Ito (2002) pourtant, dont le travail se situe à la fois dans le champ de la pragmatique et de la phonétique, émet la remarque suivante à propos de travaux récents<sup>56</sup> : "Since they collected polite/non-polite utterances with acting, (...) naturalness of the utterances was not guaranteed." La question de l'authenticité est en effet la principale réserve que l'on peut émettre au sujet de ces données artificielles : si "naturel" que soit le jeu, on ne peut pas être certain que l'acteur ne surjoue pas l'émotion, que ne s'immisce pas dans le jeu une forme d'exagération, ou d'approximation (Audibert, Aubergé & Rilliard 2004 : 2). Audibert *et al.* (2008) se demandent aussi dans quelle mesure des auditeurs naïfs sont à même de discriminer des énoncés "actés" (joués par des acteurs), d'énoncés spontanés. Voici leur conclusion :

*Les résultats obtenus suggèrent une capacité globale des auditeurs naïfs à discriminer des expressions émotionnelles multimodales actées vs spontanées, sans effet de l'émotion exprimée mais avec un fort effet inter-juge qui pourrait être lié à la notion d'intelligence affective,*

ce qui est confirmé par les travaux de Damasio (Aubergé 2002 : 271). L'auditeur moyen serait bien capable de différencier énoncés écologiques et énoncés joués, d'où toute l'importance de travailler à partir de corpus naturels si l'on veut tenter d'analyser de manière fiable ce qui se passe dans les interactions vécues au quotidien. Enfin, "la réalité langagière

---

<sup>56</sup> Voir par exemple Hirose, K., Kawamani, H. & Ihara, N., (1997). Analysis of Intonation in Emotional Speech. *ESCA Workshop on Intonation : Theory, Models and Application*, Athens, 185-188, ou Ofuka, E., MC Keown, J.D., Waterman, M.G. & Roach, P.J., (2000). Prosodic cue for rated politeness in Japanese speech. *Speech Communication*, 32, 199-217.

est beaucoup plus complexe" (Devillers & Vasilescu 2004) que ne peut l'être la parole non spontanée.

La recherche en prosodie semble avoir amorcé une évolution depuis une quinzaine d'années, et certains prosodistes basent maintenant leurs travaux sur des corpus authentiques (Rittaud-Hutinet 1995, 2005; Devillers & Vasilescu 2004; Lacheret-Dujour & François 2004; Caelen-Haumont 2008 a et b; Lacheret-Dujour, Kahane, Pietrandrea, Avanzi & Victorri 2011 par exemple). Mais il ne s'agit pas encore d'une généralité et beaucoup continuent de travailler à partir de productions non spontanées (pour une discussion à ce sujet, voir Campbell 2000 : 34-38, cité par Audibert *et al.* 2008).

En analyse des interactions en revanche, la plupart des chercheurs fondent depuis longtemps leurs analyses sur i) des données écologiques : ce sont des énoncés qui n'ont pas été prononcés *pour* la recherche mais qui ont leur propre existence indépendamment de l'exploitation que l'on pourrait en faire; ou ii) "élicitées" c'est-à-dire semi spontanées : le chercheur "provoque" des données lors d'un entretien. Dans les deux cas, la meilleure façon de sauvegarder ces données est l'enregistrement audio, ou video selon les objectifs de la recherche (Kerbrat-Orecchioni 2005 : 24-25).

Bien que les conditions d'enregistrement n'aient pas été optimales à cause de voix ou bruits parasites, **j'ai choisi de fonder mon enquête sur des données de type i) écologiques.**

### **1.2.5 Paramètre d'analyse pris en compte dans la recherche : les variations de la fréquence fondamentale**

Parmi l'ensemble des données paralinguistiques, **je ne m'attacherai donc qu'à la partie prosodique, elle-même restreinte aux variations de la F0 de la voix** ou "mélodie", car "la fréquence fondamentale se présente comme [la variable acoustique] qui a de loin le plus puissant effet sur le juge" (Scherer *et al.* 1998), ou, pour reprendre Lacheret-Dujour & Beaugendre (1999 : 12), "la fréquence fondamentale est reconnue universellement comme la plus significative et la plus facile à percevoir". D'ailleurs (*Ibid.*; Lacheret-Dujour 2011a : 89) la recherche s'est principalement concentrée sur la modélisation de la fréquence. Grichkovtsova, Lacheret & Morel (2007) ont tenté de définir lequel des paramètres de l'intonation ou de la qualité vocale était prédominant dans la reconnaissance perceptive de six

états affectifs (colère, tristesse, joie, évidence, doute et ironie), sans parvenir à montrer, selon leur postulat, que la qualité vocale jouait un rôle plus important que la variation de la F0, qui reste la plupart du temps déterminante.

Les autres traits acoustiques seront globalement laissés de côté, tels le timbre, l'étude de la (dé)labialisation, du (dé)voisement, du spectre, des formants etc. Toutefois, **lorsque cela sera nécessaire à l'interprétation des données, je m'appuierai, en complément de l'analyse de la F0, sur l'accentuation, le débit et l'intensité**, corrélats acoustiques dont la mesure sera effectuée automatiquement par le logiciel Praat<sup>57</sup>.

### 1.2.6 Comment comparer des voix différentes?

Chaque locuteur possède un phonostyle différent, qui dépend i) de caractéristiques physiologiques : genre, âge, forme, épaisseur et/ou musculature des organes de résonance et du souffle; ii) de caractéristiques contextuelles : milieu socioculturel, éducation, origine géographique, langue première ou seconde, circonstances (in)formelles, caractéristiques psychologiques de l'individu, etc. Certains locuteurs ont un phonostyle au sens propre relativement monotone, c'est-à-dire que les variations de F0 seront chez eux en général peu importantes et peu nombreuses, alors que chez d'autres, la voix, que l'on qualifie alors souvent de "mélodieuse" ou "chantante", opérera des écarts plus nombreux et plus importants. **Il faudra donc prendre en compte le phonostyle de chaque locuteur au moment de l'interprétation des indices acoustiques, après en avoir décrit les principales caractéristiques** - notamment la moyenne et la variabilité de la F0 (Bänziger *et al.* 2001 : 24). Suite à quoi **il faudra convertir les fréquences hertziennes en tons de la gamme musicale occidentale standard, et voir dans quelle mesure les potentielles variations tonales sont porteuses de sens.**

### 1.2.7 Emission ou réception des expressions émotives?

**L'enquête porte délibérément sur la réception des énoncés**, dont l'intonation sera évaluée par des "auditeurs" ou "testeurs"<sup>58</sup>, et non sur les intentions de l'émetteur<sup>59</sup>, étant

---

<sup>57</sup> Praat est un logiciel créé par Paul Boersma et David Weenink de l'Université d'Amsterdam, téléchargeable gratuitement à l'adresse : <http://www.fon.hum.uva.nl/praat/>

<sup>58</sup> Voir 2.2.1.

<sup>59</sup> Dans un premier temps, je m'étais proposé de mettre en regard des intonations françaises déclarées impolies par des tests perceptifs opérés par des locuteurs natifs du français, et des intonations allophones transposées au

donné que la question posée est de savoir si certains patterns intonatifs sont majoritairement perçus par des locuteurs natifs du français comme "impolis". Ce mémoire ne s'inscrit donc pas dans la lignée des modèles de l'"activation des émotions" (*arousal theories*), dirigés vers les origines des émotions, mais plutôt de ceux de la "perception des émotions" (*appraisal theories*), (Scherer & Johnstone 2003 : 434).

### **1.2.8 Démarche de recherche**

La démarche de recherche sera la suivante :

1. Dans un premier temps, je collecterai un corpus d'énoncés en français, que j'aurai perçus sur la base de mes compétences de locutrice native comme "polis" ou "impolis", à partir d'enregistrements de conversations quotidiennes en milieu écologique.

2. Je procéderai ensuite à une expérimentation de perception auditive à propos de l'(im)politesse des énoncés français par des locuteurs natifs français inexperts dans le domaine de l'évaluation de la voix, de manière à valider ou invalider ma propre perception.

3. Puis je procéderai à l'analyse acoustique de ces énoncés, en mettant en évidence leur mélodie à partir de l'extraction automatique de la F0 à l'aide du logiciel Praat. J'isolerais des profils proches, ou "patterns", de F0, de manière à dégager des "types" d'intonations impolies distincts. Ne seront retenus pour l'analyse acoustique que les énoncés largement discriminés comme impolis, c'est-à-dire qui recueilleront un consensus de 75% ou plus d'impolitesse.

4. Dès lors, après avoir envisagé les retombées possibles dans le champ de la pragmatique, de la phonétique et de l'interculturel, ma recherche s'orientera vers la didactique du FLE par la proposition de pistes d'activités destinées à réduire les malentendus communicationnels en contexte interculturel.

---

français, qui auraient été déclarées polies par les locuteurs natifs de la langue autre, et impolies par les locuteurs natifs du français. Je posais l'hypothèse que certaines intonations allophones tout à fait polies pouvaient en réception être reçues comme impolies par le locuteur natif français, causant ainsi certains malentendus interculturels et/ou communicationnels dus à une production prosodique erronée.





## **PARTIE 2**

### **TESTS PERCEPTIFS**

*Lorsque Monsieur Bark parle, Monsieur Linh l'écoute très attentivement et le regarde, comme s'il comprenait tout et ne voulait rien perdre du sens de ses mots. Ce que sent le vieil homme, c'est que le ton de la voix de monsieur Bark indique la tristesse, une mélancolie profonde, une sorte de blessure que la voix souligne, qu'elle accompagne au-delà des mots et du langage, quelque chose qui la traverse comme la sève traverse l'arbre sans qu'on la voie.*

*La Petite Fille de Monsieur Linh, Philippe Claudel (2005), p. 51-52 (Ed. Stock, coll. Le Livre de Poche)*

## **2.1. Présentation des corpus**

Afin d'obtenir les documents sonores les plus pertinents au regard de l'objectif que je me suis fixé, j'ai choisi d'enregistrer des conversations authentiques et spontanées, dans le cadre de moments partagés en famille ou de rencontres de socialisation. J'ai ainsi constitué des corpus dits de "spécialité", c'est-à-dire qui visent à regrouper "des données linguistiques particulières pour représenter une dimension particulière : un domaine (français médical, etc.), un thème, *une situation de communication*<sup>60</sup>, etc."(Delais-Roussarie & Durand 2003 : 94). En l'occurrence, les données que j'ai isolées ont pour la plupart été collectées dans un milieu sociologique et deux contextes interactionnels homogènes : rencontres entre pairs et amis, et moments partagés en famille); il s'agit donc de deux situations de communication particulières.

Les énoncés sont assez courts, à la modalité déclarative, interrogative ou oratoire<sup>61</sup>.

### **2.1.1 Enregistrements : cadre, prise de son**

Tous les enregistrements ont été faits entre le mois de décembre 2012 et le mois d'avril 2013, à l'aide d'un appareil à enregistrer bidirectionnel de marque "Zoom" et de type "Handy Recorder H 2". Le court laps de temps sur lequel se sont étendus les enregistrements et l'unité situationnelle des interactions ne permettent donc pas d'envisager de variation dans le phonostyle des locuteurs.

Le corpus est constitué de six enregistrements, entièrement effectués lors de rencontres informelles non planifiées entre amis ou en milieu familial. A ce titre, un certain nombre de bruits parasites apparaissent dans les enregistrements, bruits de couverts, chevauchements de voix etc., pris en compte au moment de l'analyse acoustique. Le microphone a été placé au mieux par rapport aux sources, approximativement à la hauteur de la bouche des locuteurs, à proximité du lieu de discussion en évitant le centre afin d'atténuer les chocs de vaisselle (Delais-Roussarie & Durand 2003 : 204 ).

Les sujets de conversation n'ont en aucune manière été prévus ou guidés, les interactions se sont déroulées de manière très spontanée, car une fois les premières minutes

---

<sup>60</sup> Mis en italiques par moi.

<sup>61</sup> On entend ici par "modalité oratoire" ce que les grammairiens appellent "question rhétorique" ou "interrogation oratoire", à savoir une question pour laquelle aucune réponse n'est attendue. Pierre Fontanier, grammairien du XVIII<sup>e</sup> s. spécialiste des figures de style, l'appelle «interrogation figurée» car il y voit une manœuvre du locuteur à son allocutaire qui le place dans l'impossibilité de pouvoir ni nier ni répondre.

passées, l'appareil à enregistrer a été oublié. A noter que tous les enregistrements à l'exception d'un seul, effectué par une amie à son domicile, ont eu lieu en ma présence et avec ma participation, si bien que je suis au fait du contexte interactionnel de chacun des énoncés.

	Nombre de Q°	Nombre de D°	Durée d'enregistrement totale	Durée des énoncés (fourchette)
Corpus	47	111	5 h 26 mn et 27 s	0.62 s à 7.51 s.

Tableau 1: Tableau synthétique de présentation du corpus.

### 2.1.2 Identification et durée des corpus

L'ensemble du corpus<sup>62</sup> est constitué de six enregistrements, dont les durées respectives s'étalent entre 45 et 105 minutes, pour un total de 5 heures 26 minutes et 27 secondes. En ont été extraits :

- un corpus de 47 énoncés interrogatifs (Q°) regroupant à la fois de vraies questions, ouvertes et fermées, et des questions dites "oratoires" ou "fausses questions. La durée des énoncés est comprise entre 0.62 seconde (s) et 6.83 s;

- un corpus de 111 énoncés déclaratifs affirmatifs (D°) dont la durée est comprise entre 1.1 s et 7.51 s.

Les corpus déclaratifs négatifs et injonctifs ont été constitués mais écartés<sup>63</sup> des analyses.

### 2.1.3 Procédure d'organisation des données

J'ai nommé les énoncés au fil du découpage des fichiers sonores, selon des termes qui, *de mon point de vue*, caractérisaient leur contenu sémantique. Les fichiers se sont donc automatiquement organisés par ordre alphabétique, aléatoire dans la mesure où un autre analyste aurait probablement choisi d'autres termes d'identification. J'ai ensuite anonymé ces fichiers en donnant le code "D°1" au premier énoncé, "D°2" au deuxième et ainsi de suite. Il en a été de même avec les questions (Q°1, Q°2 etc.). Les testeurs ont ainsi reçu les énoncés dans un ordre aléatoire, pour ce qui est des caractéristiques prosodiques...

<sup>62</sup> Pour la transcription du corpus français, voir en annexe B.

<sup>63</sup> Voir infra 2.1.5 Sélection des énoncés, choix des modalités.

... Mais pas en ce qui concerne les modalités : afin de ne pas ajouter pour le testeur une difficulté d'écoute supplémentaire, les énoncés ont été présentés regroupés selon leur modalité déclarative affirmative ou interrogative.

## **2.1.4 Les locuteurs et leur phonostyle**

### *2.1.4.1 Identification des locuteurs*

Le corpus fait appel à 11 locuteurs, dont

- 6 adultes (4 femmes et 2 hommes), âgés de 40 à 53 ans;
- 3 adolescents (2 filles et 1 garçon), âgés de 14, 19 et 17 ans;
- 2 enfants, (1 fille et 1 garçon), de 9 et 7 ans,

habitant tous la région Rhône-Alpes. Les phonostyles régionaux ne sont pas marqués, hormis pour deux locutrices adultes, originaires l'une de la région PACA, l'autre du Nord-Pas de Calais. Pour tous, le français est la langue première.

Les énoncés produits par les adultes représentent 92.4 % du corpus, ceux des adolescents 6.96 %, et ceux des enfants (0.63%). Les locuteurs adultes appartiennent au milieu enseignant, les adolescents sont scolarisés en filière générale de lycée.

### *2.1.4.2 Phonostyles des locuteurs*

Selon Vaissière (2006, citée par Pépiot 2013 : 25), les valeurs moyennes de F0 féminines se situent autour de 240 Hz, ou de 220 Hz pour Arnold (2012 : 5)<sup>64</sup>. Certains auteurs présentent une fourchette de variation de la F0 féminine qui s'étend de 140 à 240 Hz (*Calliope*, 1989 : 133, cité par Pépiot 2013 : 25), sachant que la F0 "peut présenter des variations considérables chez un même locuteur selon le type de phrase prononcée, son attitude et état émotif lors de l'acte de parole, etc." (*Calliope ibid.*).

Pour les voix masculines, les valeurs indiquées par Vaissière, Arnold et les auteurs de la revue *Calliope (ibid.)* sont respectivement de 120 Hz, 120 Hz et de 100 à 140 Hz.

Ne seront ici présentés que les phonostyles des locuteurs dont la voix a effectivement été analysée sur le plan acoustique, c'est-à-dire ceux dont les tests perceptifs ont déterminé des énoncés comme étant nettement impolis (plus de 75% de consensus), ou nettement polis

---

<sup>64</sup> Il semble toutefois que la valeur moyenne de F0 attendue varie sensiblement d'une communauté linguistique à l'autre, ce qui indique que des facteurs sociaux entrent en jeu dans la construction de la hauteur moyenne de la voix, masculine ou féminine (Arnold 2012, Pépiot 2013).

(plus de 80% de consensus). Je me concentrerai sur les valeurs des fréquences fondamentales (F0) et de leurs variations. Les locuteurs sont identifiés par les lettres FF ou MF selon qu'il s'agit d'hommes ou de femmes français, et par AFF ou AMF s'il s'agit de locuteurs adolescents, féminins ou masculins français.

### - Voix féminines<sup>65</sup>

La voix de FF1, 40 ans, est présente sur 18 des énoncés retenus pour l'analyse; 9 sont des énoncés interrogatifs, 9 des énoncés déclaratifs<sup>66</sup>. 16 ont été jugés "impolis" par les testeurs. La locutrice possède une voix claire aiguë, puisque la valeur moyenne de la fréquence fondamentale des énoncés qu'elle prononce se situe autour des 302 hertz (Hz), ce qui correspond à un Ré#3-<sup>67</sup>, bien au dessus des valeurs généralement proposées. La valeur du fondamental usuel (ou de confort) est d'environ 220 Hz (La2). L'amplitude de ton sur l'ensemble des énoncés est importante : la F0 minimum se situe sur le Fa#1 (131.8 Hz dans l'énoncé D°66), et la F0 maximum est proche du Do#4 (757 Hz dans l'énoncé D°105), ce qui représente 2 octaves et 2.5 tons d'amplitude. Si l'on mesure l'écart tonal maximal<sup>68</sup> pour chaque énoncé, on observe de même une grande amplitude qui va de 114.5 Hz pour l'écart le plus faible (Q°11), jusqu'à 620.2 Hz pour le plus important (D°105). Sur le plan perceptif, l'ensemble donne une ligne mélodique relativement distendue.

La voix de FF2, 46 ans, est présente dans 15 des énoncés retenus pour l'analyse; 6 sont des questions, 9 des déclarations. 7 énoncés ont été déclarés "impolis" par les testeurs. Cette locutrice fait entendre un accent méridional et parle d'une voix claire médium : la moyenne des valeurs de sa F0 est de 227.45 Hz, soit un La2+, le fondamental usuel autour de 180 Hz (Fa#2). Le ton le plus bas étant le Do2 (130.5 Hz dans l'énoncé D°22) et le plus élevé un Ré4+ (598.4 Hz en Q°22), la voix couvre 1 octave et 7 tons. A l'intérieur de chaque énoncé, l'écart tonal maximal le plus faible est de 55 Hz (énoncé Q°7), le plus élevé est de 368.3 Hz (Q°22). Les variations de F0 présentent souvent des valeurs significatives, la ligne mélodique est donc dans l'ensemble très souple.

<sup>65</sup> Pour le détail des valeurs mesurées par énoncé et par locutrice, voir en annexe D.

<sup>66</sup> Pour l'analyse et l'interprétation des valeurs selon la modalité, voir en 3.2.2 et 3.2.3.

<sup>67</sup> Les hertz sont convertis en tons musicaux, qui sont les notes de musique de la gamme dite "tempérée" : l'octave est divisée en douze intervalles égaux : do, do#, ré, ré#, mi, fa, fa#, sol, sol#, la, la#, si. Selon l'octave à laquelle la note appartient, on adjoint le numéro -2, -1, 0, etc. jusqu'à 8, en couvrant ainsi les 10 octaves du registre du piano. Voir en annexe C le tableau des valeurs fréquentielles (exprimées en hertz) converties en tons.

<sup>68</sup> C'est-à-dire la différence entre la F0 la plus basse et la F0 la plus haute dans un même énoncé.

La locutrice FF3, 44 ans, prononce 9 énoncés retenus pour l'analyse acoustique; l'un est un énoncé interrogatif, 8 sont déclaratifs. 7 ont été déclarés "impolis" lors des tests perceptifs. La voix, légèrement enrouée et d'une hauteur médium de 251 Hz proche d'un Si2+, présente par moments un accent régional du Nord-Pas de Calais. La valeur du fondamental usuel se situe autour de 200 Hz (Sol2+). L'amplitude de ton sur l'ensemble des énoncés est importante : la F0 minimum se situe sur le Do2 (129.8 Hz dans l'énoncé D°94), et la F0 maximum est proche du Ré4 (596.7 Hz en D°60), ce qui représente près de 1 octave et 7 tons. A l'intérieur de chaque énoncé, l'écart tonal maximal peut être très faible - 59.5 Hz (énoncé D°4) pour la valeur minimale, ou très ample - 420.9 Hz (D°60) pour la valeur maximale. Les variations de F0 présentant souvent des valeurs élevées, la ligne mélodique fluctue donc de même dans des proportions importantes.

La locutrice FF4, 45 ans, prononce 6 des énoncés retenus pour l'analyse. Elle possède une voix claire et relativement aiguë, puisque la moyenne de la F0 sur l'ensemble des énoncés est de 300.98 Hz, proche d'un Ré3, la valeur minimale étant de 143.3 Hz en Q°29 (Ré2) et la valeur maximale de 570.7 en Q°23 (Ré4-) : sa voix couvre donc 2 octaves. Il n'a pas été possible de repérer la hauteur de fondamental usuel. L'écart tonal maximal le plus bas est de 63.7 Hz en Q°31, et le plus élevé de 409.9 Hz en Q°29\_A. La particularité de ces énoncés est qu'ils sont tous interrogatifs, et qu'ils ont tous été classés "polis" par les testeurs.

AFF, 14 ans, ne prononce que 2 énoncés retenus pour l'analyse, 2 déclarations, dont l'une a été déclarée "impolie". Vu le peu de valeurs relevées ici, je n'ai pas calculé avec précision les moyennes de F0 comme pour les autres locutrices. On peut toutefois émettre l'hypothèse que la valeur moyenne de sa F0 avoisine les 200 Hz, soit un Sol2/Sol#2. La F0 minimum relevée est de 106.6, proche d'un Sol#1, et la F0 maximum est de 281.3 Hz, proche d'un Do#3; la voix couvre donc 1 octaves et 2 tons. Les écarts maximaux sont de 80.1 Hz pour l'énoncé 34 et de 37.5 Hz pour le 43.

**- Tableau synthétique des valeurs relevées sur l'ensemble des énoncés retenus pour l'analyse acoustique pour les voix féminines, et leur conversion en tons musicaux**

	MO F0	Eq. tons	F0 U	Eq. tons	VMi F0	VMa F0	ET max	Ampl. oct / tons
<b>FF1</b>	302.78	Ré#3-	220	La2	131.8	757.2	620.2	2 / 2.5

<b>FF2</b>	227.45	La2+	180	Fa#2-	130.5	598.4	368.3	1 / 7
<b>FF3</b>	251.21	Si2+	200	Sol2+	129.8	596.7	329.4	1 / 7
<b>FF4</b>	300.98	Ré3+	?	?	143.3	570.7	349.5	2 / 0

**Tableau 2 : Tableau synthétique des valeurs relevées pour les voix féminines et leur conversion en tons musicaux.**

MO F0 = valeur moyenne de la F0; Eq. tons = équivalence en tons musicaux par valeur approchée; F0 U : valeur approximative de la F0 usuelle ou de confort; VMi = valeur minimum de la F0; VMa = valeur maximum de la F0; ET max = valeur de l'écart tonal maximal entre la VMi F0 et la VMa F0 sur l'ensemble des énoncés de la locutrice; Ampl. oct / tons = amplitude de la voix donnée en octaves et en tons.

### - Voix masculines<sup>69</sup>

HF1, 49 ans, prononce 7 des énoncés retenus pour l'étude. 2 sont des questions, 5 des déclarations. 5 énoncés ont été déclarés "polis". Ce locuteur possède une voix légèrement enrouée et haute : la valeur moyenne de F0 est de 203.33 Hz, soit un Sol#2-, bien au dessus des valeurs standards proposées ci-dessus pour les hommes. La fréquence du fondamental usuel se situe autour de 170 Hz (Fa2-). La note la plus basse est un Fa#1 (93 Hz) dans l'énoncé Q°15, la plus haute est au-dessus d'un Mi3+ (336.9 Hz en Q°33) : la voix couvre 1 octave et 4.5 tons. L'écart tonal maximum le plus bas est de 74.4 Hz en D°62, et le plus élevé de 175.4 Hz pour D°53.

La voix de HF2, 53 ans, est présente sur 13 des énoncés retenus pour l'étude, tous des déclarations. 9 ont été déclarés "impolis". Si la valeur moyenne de sa F0 se situe au dessus des valeurs standards proposées ci-dessus pour les hommes avec 174.53 Hz (Fa2), la voix, dont la fréquence fondamentale de confort semble à l'écoute plutôt basse, possède une amplitude de 229.6 Hz soit 1 octave et 5 tons comprise entre 88.9 Hz en D°81 (Fa1) pour la valeur minimale et 323.4 Hz (Mi3-) en D°30 pour la valeur maximale. L'écart tonal maximum le plus bas est de 71.6 Hz en D°90, et le plus élevé de 189.5 Hz pour D°30.

AMF, 17 ans, produit seulement 3 des énoncés retenus pour l'étude, 3 déclarations, dont deux ont été déclarées "impolis". La valeur moyenne de la F0 semble se situer autour des 120 Hz (Si1); les écarts tonaux vont de 3 tons à 1 octave, la F0 minimum enregistrée se situant sur un Sol1 (97.5 Hz), la F0 maximum étant proche d'un La#2 (225.8 Hz). La voix est perçue comme une voix de poitrine, grave, qui varie peu en hauteur.

<sup>69</sup> Pour le détail des valeurs mesurées par énoncé et par locuteur, voir annexe D.

**- Tableau synthétique des valeurs relevées pour les voix masculines et leur conversion en tons musicaux**

	MO F0	Eq. tons	F0 U	Eq. tons	VMi F0	VMa F0	ET max	Ampl. oct / tons
<b>HF1</b>	203.33	Sol#2-	170	Fa2-	93	363.6	175.4	1 / 4.5
<b>HF2</b>	174.53	Fa2	?	?	88.9	311.4	189.5	1 / 5

**Tableau 3 : Tableau synthétique des valeurs relevées pour les voix masculines et leur conversion en tons musicaux.**

MO F0= valeur moyenne de la F0; Eq. tons = équivalence en tons musicaux par valeur approchée; F0 U : valeur approximative de la F0 usuelle ou de confort; VMi = valeur minimum de la F0; VMa = valeur maximum de la F0; ET max = valeur de l'écart tonal maximal entre la VMi F0 et la VMa F0 sur l'ensemble des énoncés du locuteur; Ampl. oct / tons = amplitude de la voix donnée en octaves et en tons.

### 2.1.5 Sélection des énoncés, choix des modalités

#### 2.1.5.1 Impolitesse directement ou indirectement adressée

Il convient de souligner ici que pour la sélection des énoncés, je n'ai pas tenu compte du fait qu'ils étaient ou non directement adressés. Pour les interactionnistes (Goffman 1973 : 164, Kerbrat-Orecchioni 2005 : 203 par exemple), l'impolitesse adressée au tiers absent, (ou "délocuté"), n'est pas de l'impolitesse au sens propre : on parle plutôt dans cette situation là de médisance : un locuteur (L1) dit à L2 du mal de L3, comme dans l'exemple ci dessous :

*/I pass'ses journées sur internet ↗ alors qu'es(t)-ce tu veux ↗/*

(énoncé Q°1)

Le concept est considéré comme "glissant" (Mougin 2006 : 14 sq.) : on ne sait pas encore si la médisance donne lieu à une structure énonciative particulière - dont il ne faut pas ignorer l'aspect sonore<sup>70</sup> - et si c'est un acte de langage à part entière. Kerbrat-Orecchioni (citée par Mougin 2006 : 15) souligne une particularité avérée de l'acte de médire, qui "n'est pas d'atteindre l'allocataire mais une tierce personne située en dehors de l'espace conversationnel."

Les passages de discours rapporté direct (DRD) m'ont aussi fourni plusieurs énoncés "impolis", car, correctement découpés, ils pouvaient passer pour du discours direct. La situation est globalement la même que pour la délocution médisante : un locuteur (L1) cite

<sup>70</sup> Ajouté par moi.

pour son interlocuteur (L2) les paroles d'un tiers absent (L3, qui peut être lui même), paroles que, par convention et sauf précision, il prétend reproduire fidèlement y compris dans leurs éléments paralinguistiques (Torck 2006). Mais comme le montrent certaines études (Klewitz & Couper-Kuhlen 1999; Couper-Kuhlen 1999), le DRD est marqué prosodiquement, et le locuteur a à sa disposition plusieurs signaux sonores : les plus courants sont i) un changement de voix, qui s'apparente à une imitation plus ou moins fidèle du phonostyle de la personne dont on reprend les paroles, ii) une intonation ou une structure rythmique différentes de celles utilisées normalement par le locuteur. Bertrand (2003 : 271) quant à elle remet en cause de manière convaincante la fidélité du DRD dont elle développe ensuite dans une optique pragmatique les fonctions, notamment en termes de places et de faces. La prosodie joue le tout premier rôle dans la distorsion sémantique du DRD et dans les actes de langage. Bertrand & Espesser (2002) ont pu mettre par exemple en évidence une augmentation significative de la F0 pour les passages de DRD de leur corpus.

On ne peut donc pas, **en contexte**, considérer ces énoncés comme impolis. Pourquoi ai-je alors fait le choix de les retenir tout de même pour l'étude?

La première raison est que le sujet de mon mémoire se limite à l'expression de l'impolitesse du seul point de vue prosodique, c'est-à-dire **hors contexte**<sup>71</sup>. Ainsi, peu importe quelles aient été réellement les intentions du locuteur, la situation de communication, ou à quel point le DRD est fidèle. Seule doit être prise en compte la réception de l'intonation par un locuteur natif du français.

La deuxième raison, qui rejoint la première, c'est que très tôt dans ma réflexion méthodologique, j'avais envisagé de filtrer sur le plan sémantique les données enregistrées : les testeurs ne pourraient dans ce cas ni reconstruire un contexte, ni s'appuyer sur le sens, ni savoir de qui parlaient ces énoncés. Malheureusement le filtrage n'a pas été possible<sup>72</sup>, c'est pourquoi j'ai expressément demandé aux testeurs de se mettre en position de destinataire, non pas de l'énoncé, mais de son intonation.

En outre, le filtrage sémantique me facilitait la tâche de collecte du corpus, car trouver des énoncés impolis directement adressés lors de conversations amicales est très difficile, de par la nature même de ces interactions : "a particular problem is that impoliteness formulae

---

<sup>71</sup> Voir 1.2.2.

<sup>72</sup> Voir 1.2.1.

are much less frequent than politeness formulae" (Culperer 2010 : 3238), et Leech (1983 : 105, cité par Culperer 2010) : "conflictive illocutions tend, thankfully, to be rather marginal to human linguistic behaviour in normal circumstances". Au point que Culperer, pour mener à bien une de ses expérimentations sur les mots impolis utilisés en certain contexte, a dû faire appel à mille étudiants pour rassembler seulement cent occurrences<sup>73</sup>. C'est que les propos impolis directement adressés se rencontrent en situation :

i) de conflits ouverts avec éclats émotifs, aussi appelés "émotions brutes" ou "émotions chocs" (Léon 1993 : 113; voir 1.1.2.1). Etymologiquement, ce type d'atteinte à la face est un ac-cident<sup>74</sup>, c'est-à-dire qu'il nous "tombe dessus sans prévenir" : le dispositif d'enregistrement est alors trop long à mettre en place au regard de la fulgurance de la manifestation des émotions - ou bien il faudrait procéder à des enregistrements d'une longueur extrêmement importante, ce qu'a dû faire par exemple Culperer dans l'expérimentation qu'il retrace à la page 3241 du n° 42 du *Journal of Pragmatics* (2010), et qui a nécessité plus de soixante heures d'enregistrements<sup>75</sup>.

ii) de conflits larvés, de rivalités souterraines, d'un je ne sais quoi qui ressortit à l'ensemble des attitudes mimo-gestuelles, praxématiques et vocales à un moment "t". Ces atteintes là peuvent être saisies au cours d'enregistrements d'une certaine durée - le corpus en comprend d'ailleurs quelques unes - mais dépendent pour l'essentiel des rapports de force entre les interactants, qui, pour diverses raisons psychosociologiques, feront preuve d'une "volonté de puissance" plus ou moins développée.

Pour pallier le manque d'énoncés impolis directement adressés en situation écologique, j'ai donc aussi sélectionné des énoncés impolis en l'absence du tiers ou des énoncés de discours rapporté direct.

L'analyse exhaustive et détaillée des tests perceptifs sera faite en partie 2.3. Mais j'anticiperai ici pour dire que le recensement des énoncés désignés comme impolis à plus de 75 % par l'ensemble des testeurs m'a réservé deux grandes surprises, la première étant que pour la très grande majorité de ces énoncés, l'impolitesse est adressée effectivement à

---

<sup>73</sup> Pour plus de détails sur l'expérimentation, voir Culperer 2010 : 3238.

<sup>74</sup> Voir note 15.

<sup>75</sup> Ce que font aussi par ailleurs les émissions de télé réalité (*Koh-Lanta*, *Secret Story* etc.) afin de pouvoir procéder ensuite à un montage satisfaisant à leurs objectifs. (voir à ce sujet l'article "Télé-réalité et marketing de l'urgence" de L. Tabary-Bolka, pages 220-232 du n°29 de la revue *Communication et Organisation* (2006), DOI 10.4000/communicationorganisation.3415, consulté le 10/03/2014).

l'interlocuteur - ou peut passer pour l'être - alors que j'avais en ce qui me concerne anticipé que nombre d'énoncés adressés au tiers absent seraient déclarés "impolis" par les testeurs <sup>76</sup>...

#### *2.1.5.2 Sélection des énoncés*

Compte tenu de ce qui précède, la presque totalité des énoncés exploitables dans les enregistrements a été retenue. L'audibilité des énoncés a été un critère majeur, jugée :

- en fonction de la qualité sonore, absence de chevauchement ou de bruit parasite autant que possible;
- en fonction de leur complétude : les énoncés n'ont pas été coupés dans leur développement.

#### *2.1.5.3 Modalités*

J'ai choisi d'exploiter les modalités déclaratives affirmatives parce qu'elles sont très largement représentées dans le corpus : 111 énoncés sur les 158 retenus (69.2 %). J'ai aussi retenu le mode interrogatif, suffisamment représenté dans les interactions (47 occurrences conservées soit 29.8 % du corpus retenu, et qui m'a paru particulièrement intéressant sur le plan prosodique, dans la mesure où deux types prosodiques se dégagent nettement selon que la question est une requête - c'est-à-dire une question véritable - ou oratoire.

Les énoncés déclaratifs négatifs et injonctifs, bien qu'eux aussi très significatifs prosodiquement, ont été écartés car ils étaient très peu représentés sur l'ensemble du corpus<sup>77</sup>, notamment les injonctions : toujours formulées par la même personne, une mère, en direction de jeunes enfants, il semble que le mode injonctif ne soit guère utilisé en interaction de socialisation.

---

<sup>76</sup> Pour la deuxième surprise, voir en 2.3.2. Pour la mise en regard des résultats attendus et des réponses proposées, voir en 2.3.3.2.

<sup>77</sup> Respectivement 23 et 25 occurrences.

## **2.2 La méthode d'expérimentation**

### **2.2.1 Les tests de perception auditive par des locuteurs français natifs**

#### *2.2.1.1 Les auditeurs-testeurs natifs français*

Les auditeurs-testeurs natifs français sont au nombre de 34 et sont identifiés aléatoirement par des numéros. La fourchette d'âge s'étend de 25 à 73 ans. Ils sont de milieu socioprofessionnel relativement uniforme : on y trouve majoritairement des étudiants, des doctorants, et des enseignants du secondaire (dont une d'éducation musicale) et du supérieur; le reste est constitué de deux travailleurs sociaux, deux cadres B de la fonction publique, de deux cadres A d'entreprise, et d'une orthophoniste.

#### *2.2.1.2 Passation des tests*

Les tests de perception auditive se sont déroulés en milieu clos dans un environnement calme, toujours en ma présence, ce qui a souvent permis de répondre à des questions ou de mettre le testeur en confiance. Un temps spécifique a été dédié à la passation du test lors de rendez-vous.

Les auditeurs ont systématiquement écouté les fichiers sons à l'aide d'un casque de marque Bose, de manière à optimiser la qualité du son ainsi que leur concentration.

J'ai dans un premier temps expliqué au testeur le sujet de mon mémoire et mon hypothèse de recherche, ainsi que la procédure matérielle. Puis une fiche support lui a été fournie, dont il a pu lire en toute tranquillité les consignes et recommandations<sup>78</sup>.

J'ai dû très souvent commenter les items 2 et 3 des consignes, à la demande des testeurs, pour éclaircir le sens des mots "intensité" et "débit". Au fil des passations, je me suis par ailleurs très vite rendue compte que la formulation du 3 n'était pas pertinente : on ne parle bien évidemment pas d'intensité "rapide"! Pour le 2, c'est un manque de perspicacité : je n'avais pas anticipé qu'un débit lent, voire très lent, peut être tout aussi impoli voire encore davantage qu'un débit rapide.

---

<sup>78</sup> Voir les consignes extraites de la fiche support page 63. Pour prendre connaissance de la fiche support dans son intégralité, voir annexe A.

- Il n'y a pas de réponses justes ou fausses.
- Ne perdez pas de vue la définition de l'impolitesse donnée ci-dessous.
- Autant que possible, ne tenez pas compte du sens des énoncés.
- Vous pouvez réécouter les enregistrements autant de fois que vous le souhaitez.

Vous allez entendre des enregistrements de paroles en conversation spontanée.

1) Face à chaque enregistrement, inscrivez la lettre **P** si vous estimez que l'intonation est **polie**, ou la lettre **I** si vous estimez qu'elle est **impolie** (**voir définition donnée ci-dessous**), un ? si vous n'avez pas d'avis.

2) Mettez une croix dans la colonne "Débit", si vous estimez que le débit rapide du locuteur vous aide à déterminer l'impolitesse.

3) Même chose pour la colonne "Intensité".

**Définition de l'intonation impolie** : est ici considérée comme **intonation impolie**, tout ce qui dans la façon de parler du locuteur pourrait sembler pour autrui menaçant, embarrassant, humiliant, vexant, méprisant etc.: ton, même léger, de colère, mépris, doute, moquerie, provocation, indignation etc.

L'item 1 n'a en lui-même pas posé de problème de compréhension, si ce n'est qu'un assez grand nombre de testeurs a éprouvé le besoin de me demander confirmation au sujet de la subjectivité de la perception auditive, qui n'était dans la consigne que sous-entendue : fallait-il vraiment donner *son propre* ressenti, ou bien tâcher d'évaluer ce que percevrait la moyenne des locuteurs français? J'ai alors confirmé que le test était bien subjectif : à chacun de faire part de son propre ressenti.

Je suis systématiquement revenue de mon propre chef sur :

i) la "définition de l'intonation impolie", de manière à bien souligner la différence entre les acceptions courante et linguistique des termes de "politesse" et d'"impolitesse" - qui ont parfois déstabilisé les testeurs.

ii) la nécessité présentée comme capitale au testeur d'abstraire le sens des énoncés : en l'absence de filtrage sémantico-lexical, je lui ai demandé de se concentrer sur la musique, la mélodie des énoncés, en tâchant de ne pas tenir compte du sens. Pour encore clarifier la consigne, j'ai procédé avec chaque testeur à un entraînement sur Q°1, le premier énoncé de la

liste<sup>79</sup> : en guise d'exemple, j'en ai chanté les notes en essayant de reproduire au mieux l'intonation que j'entendais, et j'ai précisé qu'on pouvait siffler, procéder mentalement ou à haute voix etc. Cela a bien éclairé les testeurs ... mais les a laissés face à leur capacité personnelle à reproduire ou seulement percevoir une mélodie<sup>80</sup>.

### 2.2.1.3 Entretiens après les tests

Après chaque test, un entretien a été effectué, pour évaluer la difficulté ressentie du test, pour préciser les éléments de difficulté et recueillir les éventuels commentaires des testeurs. Ces moments d'échange se sont allongés au fil des tests et m'ont apporté des éléments de réflexion très intéressants sur le plan méthodologique.

#### - Difficulté ressentie

Chaque testeur a évalué la difficulté du test sur une échelle allant de 1 (facile) à 5 (très difficile), au départ sans proposer d'utiliser les demi points. 5 testeurs ont toutefois spontanément souhaité exprimer leur difficulté de manière plus précise, et ont proposé l'évaluation 3.5, que l'on peut traduire a posteriori par "assez difficile".

Voici la synthèse de leurs réponses.

	1 facile	2 assez facile	3 pas facile	3.5 assez difficile	4 difficile	5 très difficile
Nombres de réponses / 34 testeurs	2	14	9	5	2	2
Valeurs en pourcentages	5.9	41.2	26.5	14.7	5.9	5.9

**Tableau 4 : Difficulté ressentie par les testeurs locuteurs natifs du français dans la réalisation des consignes du 1er test (corpus français).**

La majorité des testeurs se partage donc entre les valeurs "assez facile" et "pas facile", et a déclaré que la difficulté ressentie ne les avait pas empêchés de mener le test de manière satisfaisante. Ces résultats confirment mon ressenti pendant la passation des tests, à savoir que

<sup>79</sup> Sans toutefois donner mon sentiment sur la qualité (im)poli de l'énoncé.

<sup>80</sup> Voir 2.2.1.3 Entretiens après les tests - Difficulté ressentie : tableau de synthèse de l'évaluation de la facilité/difficulté du test.

selon la capacité d'écoute propre à chaque testeur, le test lui a paru plus ou moins facile/difficile.

*- Difficultés exprimées par les testeurs*

J'avais demandé aux testeurs, dans leur appréciation, de ne pas tenir compte autant que possible du sens des énoncés, pour se faire les destinataires de la seule intonation. Pour ceux qui ont noté la difficulté du test de 1 à 2, ils ont affirmé avoir globalement réussi à "chanter" les mélodies et donc à se détacher du sens. Parmi ceux-là, plusieurs ont fait la remarque pertinente que l'intonation était parfois en contradiction avec le contenu, ce qui les a étonnés mais pas gênés. Dans l'ensemble, une grande concentration a été nécessaire. Pour la majorité, le sens est resté gênant, d'autant plus lorsque les énoncés appartenaient au registre vulgaire, la tendance étant alors à la catégorisation "impoli".

De nombreux testeurs ont été gênés par le sens linguistique du terme "(im)politesse". Malgré mes explications initiales et la définition proposée sur la fiche de test, ils semblaient ne pas pouvoir faire l'adéquation entre le sens qu'ils attribuaient au mot (im)politesse, qu'ils concevaient principalement comme une manifestation d'agressivité, et les énoncés qu'ils entendaient. Ce n'est que lorsque je transposais la notion en termes de sentiments/attitudes perceptibles dans la voix, positifs ou négatifs, que la discrimination s'est trouvée facilitée. Ce malentendu sémantique pose un problème de fond en ce qui concerne la validité des réponses fournies par certains testeurs, qui ne se sont ouverts de cette difficulté qu'après le test et n'ont donc pas perçu comme "impoli" le ton d'un énoncé, simplement parce qu'il n'était pas ouvertement agressif<sup>81</sup>.

Plusieurs testeurs m'ont fait part d'une évolution de leur jugement au cours du test : apprécier la qualité (im)polie des énoncés devenait de plus en plus facile/difficile. Ils en concluaient que s'ils avaient pu reprendre le test, leur jugement sur les premiers<sup>82</sup> énoncés aurait certainement été différent.

Enfin, plusieurs testeurs se sont montrés sensibles aux phonostyles. Il leur a semblé que certains locuteurs s'exprimaient toujours sur un ton impoli, d'autres toujours sur un ton poli : ils se sont alors posé la question de savoir dans quelle mesure cela pouvait encore être considéré comme de l'(im)politesse. Ils ont par ailleurs repéré une locutrice à l'accent

---

<sup>81</sup> Voir aussi en 2.3.3.2.

<sup>82</sup> Vraisemblablement, le nombre de ces "premiers" énoncés serait variable selon les auditeurs.

méridional, ce qui a aussi posé problème car, de leur point de vue, l'accent "chantant" changeait, pour certains positivement pour d'autres négativement, les critères de perception de l'(im)politesse.

*- Remarques autres des testeurs*

Dans l'ensemble, les testeurs ont trouvé un grand intérêt à la démarche qui les a amusés. Ils ont pris conscience qu'un grand nombre d'intonations étaient impolies, que l'intonation en conversation familière était d'une extrême richesse, et que les énoncés dits "neutres" en étaient pour ainsi dire absents.

Ils ont presque toujours souligné la subjectivité d'un tel test, et ont été parallèlement amenés, lors d'une prise de conscience réflexive, à s'interroger sur leur propre perception de l'intonation et sur leurs propres schémas intonatifs, parfois à les remettre en question : comme finalement, le test les a conduits à noter "impoli" un nombre important d'énoncés, ils se sont demandé s'ils n'étaient pas trop sensibles, trop réceptifs, à l'intonation. A l'inverse, leurs propres productions étaient-elles "acceptables" pour l'interlocuteur? C'est pour pallier le caractère éminemment subjectif de ce type de test que j'ai tâché de le faire passer à un nombre de personnes suffisant, et que l'analyse ne retiendra que les énoncés pour lesquels un consensus de 75% aura été atteint.

Deux autres testeurs ont fait la remarque que certains énoncés s'adressaient visiblement à des enfants. Ils ont pu le déterminer parce qu'on entend effectivement des voix d'enfants dans l'enregistrement, parce que le type de propos ne pouvait que s'adresser à des enfants - *t'as pas fini ton riz t'as pas d'dessert c'est tout* -, mais aussi parce que le ton de réprimande était à ce moment là particulièrement perceptible<sup>83</sup>. L'écoute de ces énoncés les a remis dans la position inconfortable de relation d'autorité au parent, en réveillant chez eux des souvenirs désagréables. Ces énoncés, s'ils avaient été adressés à des adultes, auraient alors paru extrêmement impolis aux deux testeurs et auraient été pris au démarrage ou au cours d'une dispute. Si, à rapport de places égal, on rencontre très peu d'énoncés impolis entre adultes, ils sont en revanche beaucoup plus fréquents lorsque sont en présence des interactants en relation de places dissymétrique, de hiérarchie et/ou d'autorité : en contexte militaire, policier ou judiciaire par exemple (Bousfield 2008) ... ou encore scolaire ou familial... Je reviendrai sur ce point en 2.3.2.2.

---

<sup>83</sup> A noter que d'autres énoncés, non impolis, se trouvaient adressés à des enfants dans le corpus, mais ce n'était décelable ni par le contenu ni par les voix.

Enfin, de manière très consciente chez certains testeurs, a émergé l'intuition que dans un énoncé, l'intonation i) est perçue avant le sens lexical et prime dans la perception de l'impolitesse; ii) elle conditionne un nombre non négligeable de malentendus interculturels (entre gens de langues premières différentes), socioculturels (entre gens de classes sociales ou de régions différentes au sein d'un même pays), ainsi que, tout simplement, entre personnes de sensibilités différentes.

### **2.2.2 Retour réflexif**

Dès après la passation de ce premier test et les entretiens consécutifs, un certain nombre de conclusions s'imposent au sujet de la méthodologie.

#### *2.2.2.1 Enregistrement*

Tout d'abord, en ce qui concerne les paramètres d'enregistrement, j'évitais les moments de repas car les bruits, de couverts notamment, ont gêné la perception de certains testeurs et ont provoqué un certain nombre de bugs lors de l'analyse acoustique automatique opérée par le logiciel. Je positionnerais aussi le micro systématiquement plus près de la source, car pour éviter justement les bruits d'assiettes et de couverts, j'ai été amenée à l'excentrer de la conversation, ce qui a fait que l'on entendait la voix de certains locuteurs plus que celle d'autres. Je veillerais enfin à maintenir un niveau de gain uniforme : l'appareil a été dérégulé sans que je m'en rende compte et j'ai commis l'erreur de ne pas le vérifier. De tout cela, il a résulté que la perception de l'intensité des voix a été faussée. Je ne tiendrai en conséquence pas compte de la perception des testeurs en ce qui concerne l'intensité.

#### *2.2.2.2 Formulation des consignes*

Pour ce qui est du débit, la présentation actuelle du fichier de test ne convient pas. Pour les intonations jugées impolies, il faudrait sans doute proposer une échelle d'évaluation de la lenteur/rapidité de la production orale, en envisageant les variations possibles de rythme et en demandant dans quelle mesure le débit ainsi décrit renforce l'impression d'impolitesse. Les consignes étant imprécises, elles ont pu induire les testeurs en erreur et pour cela, je ne prendrai pas non plus en compte les réponses apportées sur ce point.

En ce qui concerne les termes de "politesse" et d'"impolitesse" utilisés dans la fiche de consignes, il serait peut-être préférable de les laisser de côté au profit de la seule formulation

"expression d'un sentiment ou d'une émotion positive/négative", beaucoup plus significative pour la moyenne des testeurs. C'est ce qu'ont fait par exemple Devillers & Vasilescu (2004) qui, après avoir demandé aux testeurs de reconnaître, dans des corpus écologiques de conversations téléphoniques, des émotions précises et fléchées, leur ont ensuite demandé de simplement les classer en émotions positives *vs* négatives. Ils ont alors noté une augmentation des performances dans la reconnaissance des émotions, qui sont passées de 70 à 85 %.

### *2.2.2.3 Filtrage des données*

Ensuite, et c'est le point à mon sens le plus important, je ne referais pas un tel test sans filtrer les données enregistrées; je paramétrerais si besoin le filtrage à partir de logiciels adéquats, tels que Praat ou Windpitch.

### *2.2.2.4 Limites d'exploitation d'un corpus sonore spontané*

Le travail sur le corpus - sélection et écoutes des énoncés, répartition personnelle des énoncés en termes d'(im)politesse, passation des tests et entretiens - m'a amenée à me poser la question de la limite d'exploitation d'un corpus entièrement naturel dans le cadre d'une étude de la prosodie.

Le matériau obtenu en situation écologique est en effet extrêmement disparate et ductile : i) les variables, d'un nombre considérable, se modifient d'un locuteur et d'une situation de conversation à l'autre : outre le caractère imprévisible de la progression de la conversation, influent sur la production sonore les contextes, le phonostyle, la mimogestuelle, le contenu sémantico-lexical, la structure syntaxique; ii) ce dernier point notamment pose question : dans la mesure où l'intonation, indépendamment de ses caractéristiques émotionnelles, structure l'énoncé tout en étant structurée par la syntaxe de l'énoncé<sup>84</sup>, dans une sorte de liaison immédiate et interdépendante, et que le nombre d'énoncés et par là-même d'intonations sont quasi infinis, est-il vraiment envisageable d'opérer des comparaisons d'ordre prosodique à partir d'un matériau aussi protéiforme, a fortiori si l'on envisage de comparer des langues entre elles?

J'envisagerais donc de guider plus fortement les conversations, en utilisant par exemple les objectifs communicationnels du CECRL<sup>85</sup> (exprimer ses goûts, son opinion,

---

<sup>84</sup> Voir en 1.1.3.4 Les fonctions de l'intonation - e) La fonction syntaxique de la prosodie, et tous les travaux des intono-syntacticiens (Bertrand, Couper-Kuhlen, Di Cristo, Hirst, Martin, Mertens, Morel, Rossi, Selting ...).

<sup>85</sup> CECRL : cadre européen de référence pour les langues.

argumenter sur un sujet donné), et/ou en proposant une tâche à réaliser, de manière à induire la possibilité de retrouver des structures syntaxiques identiques ou presque entre les différents enregistrements. On pourrait discriminer ces énoncés une fois filtrés en catégorie "poli/impoli", et l'on pourrait alors envisager de comparer la F0 de ces énoncés de longueur et de structures syntaxiques identiques ou proches : les variations de F0 ne seraient pas dues aux différences de structure syntaxique, ce dont il faudra tenir compte au moment de la description et de l'interprétation des indices acoustiques identifiés dans le corpus.

## 2.3 Résultats des tests perceptifs

### 2.3.1 Décompte

Les réponses des fiches de tests perceptifs ont été reportées dans des tableaux excel<sup>86</sup>.

Ont été retenus comme impolis les énoncés pour lesquels les taux de consensus étaient supérieurs ou égaux à 75 %, comme polis ceux pour lesquels les taux s'élevaient à 80 % d'accord. Un surlignage jaune signale dans les tableaux ci-après les taux de consensus supérieurs à 85%.

#### 2.3.1.1 Résultats d'ensemble

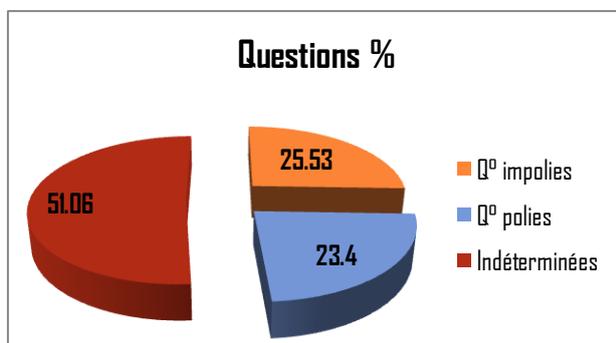


Figure I : Répartition des réponses pour l'ensemble des questions par catégorie.

Pour les déclarations, 35 sur 111 ont été jugées impolies, 18 polies, 58 restent indéterminées. Comme pour les questions, on voit qu'un peu plus de la moitié des énoncés déclaratifs ne sont pas clairement catégorisés par le test.

Le décompte a fait apparaître 12 questions sur 47 jugées impolies et 11 jugées polies. En revanche, 24 questions sur 47, soit environ la moitié du corpus de questions, n'ont pu être clairement déterminées (identifiées "indéterminées" dans les figures I et II) : les pourcentages se répartissent souvent autour de 40 % et 60 % pour l'un ou l'autre item, avec parfois une proportion non négligeable de réponses "N"<sup>87</sup>.

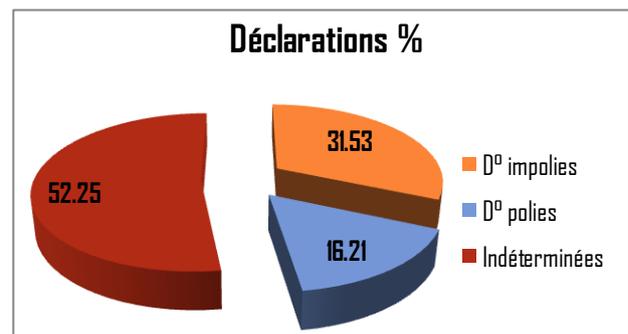


Figure II : Répartition des réponses pour l'ensemble des déclarations par catégorie.

La part importante d'énoncés indéterminés montre très nettement qu'un énoncé sur deux n'est pas compris de la même manière par ce panel de locuteurs natifs : les uns

<sup>86</sup> Voir les tableaux complets en annexe G.

<sup>87</sup> Qui a remplacé le point d'interrogation de la fiche de test dans les tableaux excel.

interprètent des intonations comme étant impolies alors que les autres non. Le caractère éminemment subjectif du test perceptif est ici bien visible, ce qui est confirmé par les proportions individuelles de réponses I, P ou N : certains testeurs perçoivent davantage une impolitesse dans l'intonation, tandis que d'autres y sont peu sensibles. On peut voir la synthèse des ces variations individuelles dans les figures III et IV ci-dessous.

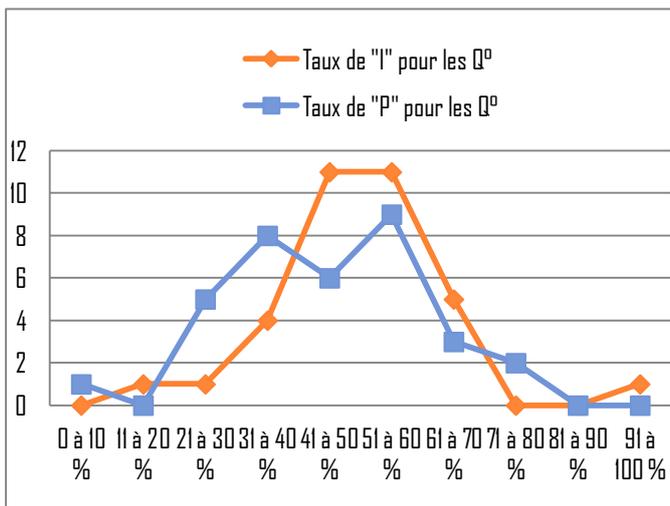


Figure III : Variations individuelles d'interprétation des intonations polies et impolies des énoncés interrogatifs. En abscisse, le nombre de testeurs.

Sans même parler des difficultés de communication entre natifs et allophones, les variations de perception individuelle mettent en lumière la fragilité de la communication entre natifs : un nombre certain de malentendus quotidiens sont assurément causés par ces variables individuelles.

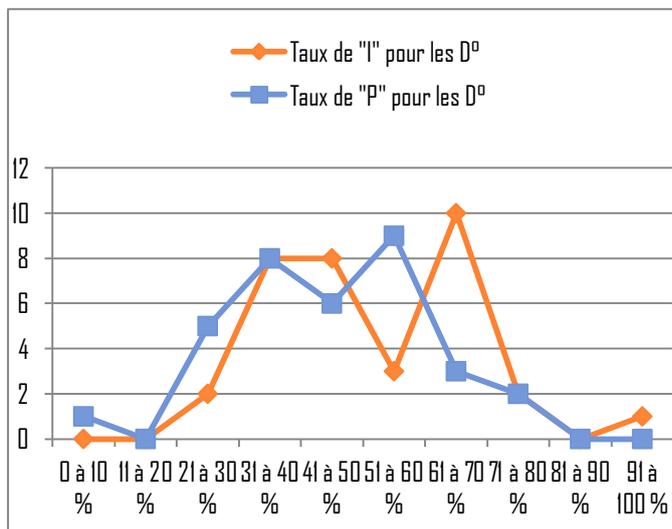


Figure IV : Variations individuelles d'interprétation des intonations polies et impolies des énoncés déclaratifs. En abscisse, le nombre de testeurs.

### 2.3.1.2 Énoncés interrogatifs

- questions jugées impolies à plus de 75 %

Q° n°	1	8	11	12	16	18	28	30	32	34	41	45
% I	97.1	76.5	94.1	97.1	79.4	97.1	91.2	94.1	76.5	94.1	97.1	76.5

Tableau 5 : Questions jugées impolies à plus de 75%.

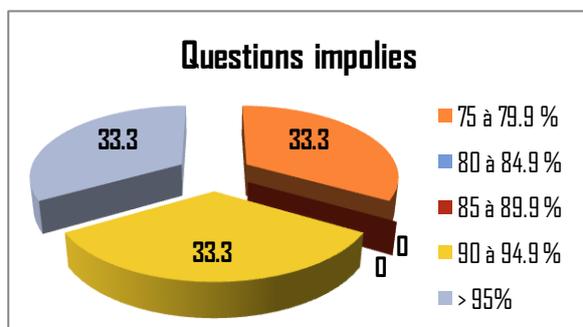


Figure V : Répartition des questions impolies par taux d'accord.

Les taux de réponse se répartissent à part égale en 3 catégories seulement : aucune réponse "intermédiaire" (entre 80 et 89.9 % d'accord) n'a été comptabilisée; au contraire, exactement les deux tiers des réponses (66.6 %) présentent un consensus de plus de 90 %, ce qui est très élevé, la dernière catégorie (taux d'accord compris

entre 70 et 79.9 %, que l'on pourra qualifier de "faible") regroupant le dernier tiers des réponses.

- questions jugées polies à plus de 80%

Q° n°	7	15	20	21	22	23	25	29	31	33	38
% P	82.4	85.3	94.1	85.3	85.3	85.3	91.2	91.2	91.2	85.3	88.2

Tableau 6 : Questions jugées polies à plus de 80%.

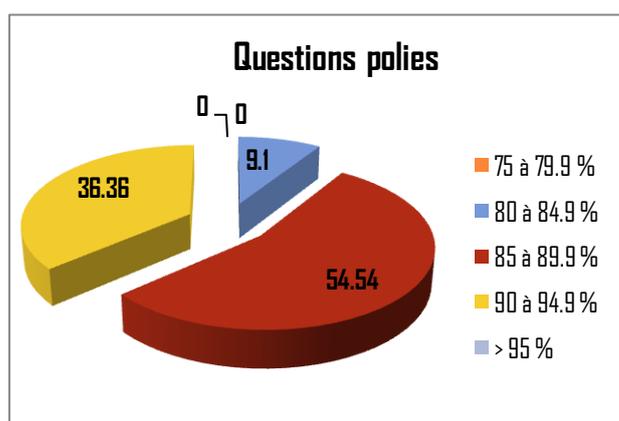


Figure VI : Répartition des questions polies par taux d'accord.

Pour les questions jugées polies, les taux d'accord sont très élevés : plus de 90 % des énoncés obtiennent un pourcentage de consensus supérieur à 85 %. Il n'y a pas de taux d'accord "faible" (75 à 79.9 %), un seul "intermédiaire".

A souligner ici : 18 des 47 "questions" sont des requêtes, c'est-à-

dire de "vraies" questions, qui attendent une réponse : 93.75% de ces requêtes ont été déclarées "polies" par les testeurs<sup>88</sup> (soit 16 sur 18) avec presque toujours de forts pourcentages d'accords. En revanche, 29 questions sur 47 sont des interrogations oratoires, dont 87.1% ont été déclarées "impolies" (soit 25 sur 29) avec le même consensus. Sur le plan pragmatique la conclusion vaut la peine d'être soulignée : les requêtes sont 9 fois sur 10 perçues comme polies alors que les fausses questions sont 9 fois sur 10 perçues comme impolies. Les actes de langage étayent cette perception : la question oratoire est toujours chargée de sentiments, émotions ou intentions, tacites pour la plupart<sup>89</sup>.

### 2.3.1.3 Énoncés déclaratifs

- **déclarations jugées impolies à plus de 75 %**

D° n°	4	8	9	15	17	19	20	23	29	30	34	37	41	51	52	53	57	58	60	66	67
% I	82.4	88.2	79.4	91.2	79.4	79.4	91.2	100	100	97.1	88.2	79.4	85.3	88.2	91.2	79.4	88.2	76.5	91.2	82.4	91.2
D° n°	68	71	74	76	77	89	90	95	96	99	101	105	109	111							
% I	79.4	94.1	79.4	94.1	94.1	76.5	91.2	79.4	76.5	91.2	94.1	91.2	97.1	88.2							

Tableau 7 : Déclarations jugées impolies à plus de 75%.

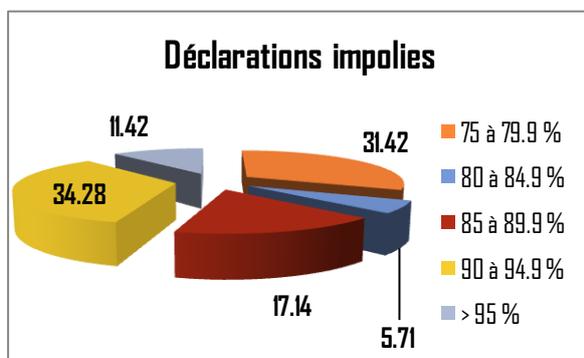


Figure VII : Répartition des déclarations impolies par taux d'accord.

"faible", alors que, en revanche, presque les deux tiers des réponses (62.84 %) se situent au-delà des 85 % d'accord.

En ce qui concerne les déclarations, les taux d'accord sont mieux répartis - ce qui s'explique par le fait que les énoncés déclaratifs sont presque deux fois et demi plus nombreux que les énoncés interrogatifs : la probabilité d'étalement des résultats augmente. Cela étant, comme pour les questions, presque un tiers (31.42 %) des réponses se situent dans le taux d'accord

<sup>88</sup> dont l'attention n'avait bien entendu pas été attirée sur le caractère véritable ou oratoire des questions.

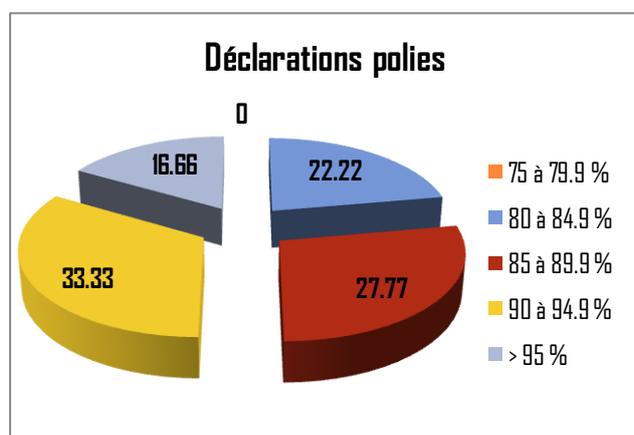
<sup>89</sup> Voir en annexe H les actes de langage des énoncés.

- **déclarations jugées polies à plus de 80%**

D° n°	2	3	16	22	26	31	32	35	43	44	54	62	63	73	75	79	81	94
%	97.1	91.2	82.4	91.2	85.3	97.1	85.3	94.1	85.3	94.1	82.4	82.4	91.2	85.3	100	91.2	82.4	88.2

**Tableau 8 : Déclarations jugées polies à plus de 80%.**

Là encore, plus des deux tiers (77.76 %) des énoncés sont jugés "polis" à plus de 85 %, 22.22 % se trouvent dans la catégorie "intermédiaire", aucun énoncé ne reçoit de taux d'accord "faible".



**Figure VIII : Répartition des déclarations polies par taux d'accord.**

### Synthèse

	75 à 79.9 %	80 à 84.9 %	85 à 89.9 %	90 à 94.9 %	> 95 %
Q° I	33.33	0	0	33.33	33.33
D° I	31.42	5.71	17.14	34.28	11.42

**Tableau 9 : répartition des réponses "énoncés impolis" par taux d'accord.**

Lorsqu'un énoncé est perçu comme impoli par les auditeurs, il semble qu'il le soit souvent (pour cette expérimentation deux fois sur trois) de manière catégorique.

	75 à 79.9 %	80 à 84.9 %	85 à 89.9 %	90 à 94.9 %	> 95 %
Q° P	0	9.1	54.54	36.36	0
D° P	0	22.22	27.77	33.33	16.66

Tableau 10 : Répartition des réponses "énoncés polis" par taux d'accord.

Plus encore que pour les énoncés impolis, les énoncés polis rassemblent des taux d'accord très élevés : la catégorisation est encore plus nette.

### 2.3.2 Impolitesse directement ou indirectement adressée

Etant donné que j'ai assisté personnellement aux conversations ou que j'en connais les locuteurs, je suis au fait des contextes d'interaction : qui parle à qui et à quel sujet. Les testeurs n'ont pas eu de leur côté connaissance de ces contextes et n'ont pas pu en tenir compte dans leur évaluation de l'(im)politesse, sauf éléments particuliers audibles dans les enregistrements. Pour cette raison, je suis à même de fournir un certain nombre d'informations qui peuvent avoir leur importance dans l'analyse.

#### 2.3.2.1 Répartition

Parmi les 12 questions retenues par les testeurs comme impolies, un énoncé seulement est pour partie indirectement adressé<sup>90</sup> : le début de l'énoncé parle d'un tiers absent, la fin s'adresse à l'interlocuteur sous forme de question phatique en registre infra-bas; un autre énoncé, que j'appelle "pseudo-adressé", est en fait indirectement adressé mais, en raison du découpage de l'énoncé, peut passer pour directement adressé<sup>91</sup>; les 10 autres sont adressés directement<sup>92</sup>. Les trois catégories représentent respectivement 8.3 %, 8.3 % et 83.3 % du corpus des questions estampillées comme "impolies" par la perception auditive.

12 déclarations<sup>93</sup>, parmi les 35 retenues par les testeurs comme impolies, sont directement adressées à l'interlocuteur, soit 34.28 %; 12<sup>94</sup> (34.28 %) sont adressées au tiers

<sup>90</sup> Q° 1.

<sup>91</sup> Q° 12.

<sup>92</sup> Q° 8, 11, 16, 18, 28, 30, 32, 34, 41 et 45.

<sup>93</sup> D° 19, 20, 41, 51, 66, 71, 76, 77, 95, 99, 101 et 105.

<sup>94</sup> D° 4, 8, 23, 29, 30, 60, 68, 74, 90, 96, 109 et 111.

absent mais peuvent passer pour être adressées à l'interlocuteur (ou pseudo-adressée); 11<sup>95</sup> (31.42 %) seulement parlent d'un absent ou ne sont pas adressées.

	QUESTIONS			DECLARATIONS		
	DA	PSEUDO DA	IA	DA	PSEUDO DA	IA
<b>%</b>	<b>83.3 %</b>	<b>8.3 %</b>	<b>8.3 %</b>	<b>34.28 %</b>	<b>34.28 %</b>	<b>31.42 %</b>
<b>Nbre d'énoncés</b>	<b>10 / 12</b>	<b>1 / 12</b>	<b>1 / 12</b>	<b>12 / 35</b>	<b>12 / 35</b>	<b>11 / 35</b>

**Tableau II : Pourcentage d'énoncés impolis directement, pseudo-directement et indirectement adressés, respectivement DA, PSEUDO DA et IA.**

### 2.3.2.2 A qui l'impolitesse s'adresse-t-elle directement?

Dans le corpus, un certain nombre d'énoncés sont issus de conversations en famille. Il est important et significatif à cet égard de souligner que sur les 10 questions impolies directement adressées, 2 seulement le sont à des pairs alors que 10 le sont à des enfants par leur(s) parent(s); de même pour 10 des 12 déclarations directement adressées et pour 9 des 12 déclarations pseudo-indirectement adressées : dans ce dernier cas, 2 des énoncés sont censés être adressés à un élève par un enseignant, un autre à un élève par un proviseur adjoint; pour les autres énoncés, un locuteur âgé de 53 ans prononce au style direct i) des paroles qu'il a lui-même adressées à un locuteur absent à un moment antérieur, ii) des paroles prononcées par autrui à son encontre, à un moment antérieur : pour i), le locuteur absent n'est autre que son propre fils âgé de 18 ans, pour ii) il reproduit les paroles de son propre père, âgé de 83 ans.

Seuls 2 énoncés déclaratifs directs et 2 énoncés pseudo-direcets sont adressés à des pairs. On trouve un cas d'énoncé pseudo-direct de la part d'un adolescent envers un adulte, dont on pourrait qualifier l'intonation d'insolente ou de désinvolte, mais cela reste à valider.

Inversement, sur les 11 énoncés impolis indirectement adressés, un seul est dirigé vers un enfant, alors que les 10 autres concernent des pairs.

<sup>95</sup> D° 9, 15, 17, 34, 37, 52, 53, 57, 67 et 89.

	QUESTIONS IMPOLIES			DECLARATIONS IMPOLIES			TOTAUX Q <sup>o</sup> + D <sup>o</sup> IMPOLIES adressées à un enfant		
	DA à un enfant	PSEUDO DA à un enfant	IA à un enfant	DA à un enfant	PSEUDO DA à un enfant	IA à un enfant	DA	PSEUDO DA	IA
<b>Nbre d'énoncés</b>	<b>10 / 10</b>	<b>0 / 1</b>	<b>0 / 1</b>	<b>10 / 12</b>	<b>9 / 12</b>	<b>1 / 11</b>	<b>20 / 22</b>	<b>9 / 13</b>	<b>1 / 12</b>
<b>%</b>	<b>100 %</b>	<b>0 %</b>	<b>0 %</b>	<b>83.3 %</b>	<b>75 %</b>	<b>9.1 %</b>	<b>90.1 %</b>	<b>69.2 %</b>	<b>4.35 %</b>

Tableau 12 : Synthèse des énoncés directement adressés (DA), pseudo directement adressés (PSEUDO DA) et indirectement adressés (IA) à des enfants.

Les énoncés impolis directement adressés le sont presque tous, à 95 %, à l'égard d'enfants par un parent, les pseudo-directs le sont à presque 70 %. Inversement, les énoncés impolis indirectement adressés le sont à 95 % à l'égard d'adultes absents.

Ces énoncés se situent pour partie dans le cadre de moments de vie de famille nucléaire, entre-soi : sont présents le(s) parents et son (leurs) enfant(s) - que la présence soit physique ou qu'elle soit recréée par le discours. Une autre partie des énoncés est censée se dérouler en milieu scolaire. On est donc là dans un contexte de relations de pouvoir, fortement hiérarchisées par la relation dissymétrique entre adultes investis socialement d'une autorité et enfants, et non dans un contexte de rencontres de socialisation. Entre parents et enfants, ce rapport de places semble pouvoir perdurer tout au long de la vie : que l'on soit un jeune enfant, un jeune adulte ou un adulte senior, le ton du parent qui réprimande sa progéniture peut être tout aussi impoli : on peut écouter et comparer à cet égard, par exemple, les énoncés D<sup>o</sup> 77 et D<sup>o</sup> 109.

La conclusion que l'on peut tirer de ces chiffres interroge : dans la vie quotidienne, les enfants - quel que soit leur âge - sont-ils les seuls à qui l'on s'adresse plus facilement et couramment sur un ton impoli, comme s'ils ne recevaient pas les mêmes égards que les pairs? Il semble en effet bien rare que l'impolitesse directement adressée soit présente dans les rencontres entre amis, au cours desquelles se (re)jouent des liens d'appartenance et d'identité de groupe : on observe bien souvent au travers des conversations une réactivation d'un "we-code" vs "they code" (Gumperz 1989b). L'impolitesse adressée au tiers absent fait alors partie des rituels conversationnels amicaux qui participent d'une mise en scène des relations sociales.

### 2.3.3 Prolongements

#### 2.3.3.1 Pour d'autres analyses

Afin d'affiner les résultats, on peut envisager d'étendre le panel et le nombre de testeurs en faisant varier le genre, l'âge, le milieu social et/ou les origines géographiques : on pourrait alors observer si la perception auditive se différencie d'un genre à l'autre, d'un âge à l'autre, d'un milieu social à l'autre, d'un phonostyle régional à l'autre. Parallèlement, on pourrait catégoriser les testeurs en fonction de la difficulté ressentie pendant les tests, de manière à voir si ce critère a un quelconque effet sur les résultats : ceux dont le ressenti de difficulté est le plus élevé pourrait-il présenter des catégorisations d'(im)politesse plus diffuses que ceux qui trouvent le test facile, une perception de l'(im)politesse plus aléatoire pour des désaccords entre testeurs plus importants? Comment les résultats se répartiraient-ils pour ceux qui jugent le test facile? Peu d'impolitesse, peu de politesse, plus d'accords sur la perception, tout autant ou moins?

#### 2.3.3.2 Ecart entre les résultats effectifs des tests et les résultats attendus

Pour l'ensemble des énoncés, j'avais passé le test moi-même avant de le soumettre aux testeurs, de manière à pouvoir situer ma perception par rapport à celle d'un groupe suffisamment large.

En ce qui concerne les énoncés interrogatifs, les résultats du test concordent entièrement avec ce que j'en avais anticipé : pour 43 des 46 questions, mes perceptions correspondent à celles des testeurs (93.5%). Pour les énoncés déclaratifs, mon ressenti et celui des testeurs s'accordent pour 95 énoncés sur 111, soit 85.6% d'accord et 14.4% de désaccord.

Malgré cette apparente convergence, une analyse plus fine amène à s'interroger une fois de plus sur la procédure : j'ai compté ci-dessus pour "accord" les taux égaux ou supérieurs à 50% du côté des testeurs. Or, parmi ces énoncés peu clairement catégorisés par les tests, mais aussi parmi ceux pour lesquels je suis en complet désaccord, il me semble que souvent, les testeurs ont écouté davantage le sens que l'intonation.

Prenons pour exemple l'énoncé D°18, jugé "poli" à 61.8% par les testeurs et "impoli" par moi même : le contenu lexical ne présente aucune impolitesse :

*"quand v'z'avez employé un mot, ça va vouloir dire une chose et pas une autre."*

Il n'en va pas tout à fait de même pour l'intonation<sup>96</sup>.

On peut segmenter l'énoncé en 3 parties :

- partie A jusqu'à 8A,
- partie B, jusqu'à 7B,
- partie C jusqu'à 3C.

N° syl	1A	2A	3A	4A	5A6A	7A	8A
Transcr <sup>o</sup>	quand	v'z'a	vez	em	ployé	un	mot
F0 Hz	194.7	177.7	160	155.5	139.1	194.4	167.8
ton	Sol2	Fa2	<Mi2	Re#2	Do#2	Sol2	Mi2>
aug. ton		-1	-1	-0.5	-1	<b>+3.5</b>	-2
Tps en ms	215	114	75	149	293	193	<b>334</b>
total aug. tons		-3.5				+3.5	-2
total Tps en ms		631				193	334

N° syl	1B	2B	3B4B	5B	6B	7B	8B
Transcr <sup>o</sup>	ça	va	vouloir	dire	une	cho	s'et
F0 Hz	185.7	175.8	179	202.8	352	214.4	190.8
ton	Fa#2	Fa2	>Fa2	>Sol2	>Fa3	>La2	>Fa#2
aug. ton	+1.5	-0.5	=	+1	<b>+6</b>	<b>-5</b>	-1.5
Tps en ms	176	188	239	<b>314</b>	<b>330</b>	<b>427</b>	293
total aug. tons	+1.5	-0.5	=	+7		-6.5	
total Tps en ms	176	188	239	644		720	

1C	2C	3C
pas	une	autre
181	152.4	∅
<Fa#2	<Re#2	
=	-2	
184	<b>276</b>	294
=	-2	

Tableau 33 : indices acoustiques de l'énoncé D°18\_P\_61.8%.

La première syllabe donne une note assez élevée pour ce locuteur (HF1), un sol2 (194,7 Hz); on relève ensuite dans l'énoncé 2 montées brutales (ou pics intonatifs) qui se font écho, sur 7A et 6B, respectivement de 3.5 tons et de 6 tons, dans les 2 cas en l'espace d'une syllabe : le locuteur insiste sur le fait qu'**un** mot correspond à **une** définition, avec la volonté de persuader son auditoire du bien fondé de son affirmation. Dans les 2 cas, la F0 redescend brutalement sur la syllabe suivante ( 8A et 7B). De 7B à 3C la valeur de F0 continue de baisser normalement par paliers.

<sup>96</sup> Pour plus de précision, il faudrait aussi mesurer l'intensité de l'énoncé, dont la variation est audible à la simple écoute.

Même en position de continuation majeure (C<sub>Ma</sub>) ou mineure (C<sub>Mi</sub>)<sup>97</sup>, 8A, 5B, 6B, 7B et 2C portent un accent de durée expressif : c'est-à-dire que même en tenant compte de cette éventuelle position, la durée de ces syllabes est au-delà de la durée standard.

Les pics intonatifs présentés par cet énoncé associés aux durées des syllabes font apparaître une forme d'impolitesse<sup>98</sup>.

Cette différence d'interprétation conduit à poser trois questions :

- les testeurs ont-ils tous bien compris la définition pragmatique de l'impolitesse?
- Dans quelle mesure ont-ils réussi à extraire ou non l'intonation comme si elle leur était personnellement adressée?
- Donc, dans quelle mesure ces tests perceptifs sont-ils fiables?

---

<sup>97</sup> Pour une explication de ces termes, voir en 3.2.1.

<sup>98</sup> Voir en 3.2.3.





# **PARTIE 3**

## **ANALYSES**

### **ACOUSTIQUES**

*Je ne lui reproche pas de ne pas être sensible à cette peinture, il n'a pas l'éducation pour, il y a tout un apprentissage qu'il n'a pas fait, parce qu'il n'a jamais voulu le faire ou parce qu'il n'avait pas de penchant particulier, peu importe, ce que je lui reproche c'est son ton, sa suffisance, son absence de tact.(...)*

*Tu as pris un petit ton suave, subalterne, qui ne te va pas du tout d'ailleurs. (...)*

*Il a un petit ton douceâtre. Un petit sourire entendu derrière chaque mot. On a l'impression qu'il s'efforce de rester aimable.(...)*

Art, Yasmina Reza (1994), p. 30, 38, 42 (Ed. Magnard, coll. *Classiques et Contemporains*)

## **3.1. Procédure**

### **3.1.1 Logiciel Praat**

Pour réaliser les analyses acoustiques, j'ai utilisé le logiciel Praat<sup>99</sup>, libre de droits, conçu à l'Institut de Sciences Phonétiques de l'Université d'Amsterdam par Paul Boersma et David Weenink. Praat permet l'analyse et la visualisation automatiques du spectre, de la hauteur (F0), de l'intensité, des formants d'un son et de la mesure de la durée des segments jusqu'au millième de seconde.

### **3.1.2 Mesures de F0 retenues**

Dans un premier temps, pour chaque énoncé jugé "impoli" à plus de 75 % et pour les énoncés jugés "polis" à plus de 80 %, j'ai systématiquement effectué les mesures de F0 suivantes : la valeur de la F0 moyenne de l'énoncé, les valeurs de F0 en début et en fin d'énoncé, les valeurs de F0 minimum et maximum. Cela m'a permis de me familiariser avec le fonctionnement du logiciel Praat tout en étudiant les phonostyles des locuteurs<sup>100</sup>.

Dans un second temps, quand il s'est agi d'effectuer véritablement les analyses acoustiques, j'ai, quand c'était possible, mesuré la valeur moyenne de la F0 de chaque syllabe. Lorsqu'une syllabe, à la vue de la courbe de F0 retranscrite par Praat ou à l'écoute, me paraissait chuter ou augmenter de manière notable, j'ai aussi effectué 2 à 3 points de mesure précis, dont les valeurs minimum et maximum de la F0 de cette syllabe.

### **3.1.3 Mesures autres**

Bien que cela n'eût pas été prévu dans le projet de recherche, j'ai été amenée à prendre en compte les deux autres paramètres acoustiques, durée et intensité.

J'ai la plupart du temps mesuré la durée des syllabes, ce qui permet de voir à quelle vitesse s'opèrent les changements de valeurs de F0 : comme on le verra ci-dessous, la rapidité et l'amplitude du changement engendrent souvent la sensation d'impolitesse. Dans le cas où ces 2 paramètres n'ont pas suffi à caractériser l'impolitesse de l'énoncé, j'ai dû avoir recours à la mesure de l'intensité moyenne des syllabes, que j'ai prises en compte dans leur valeur relative et non absolue car les modalités d'enregistrement n'ont pas permis d'obtenir des mesures fiables<sup>101</sup>.

---

<sup>99</sup> téléchargeable à l'adresse : <http://www.fon.hum.uva.nl/praat/>

<sup>100</sup> Pour consulter ces valeurs, voir en annexe D.

<sup>101</sup> Voir en 2.2.2.1.

### 3.1.4 Identification des éléments d'une fenêtre Praat

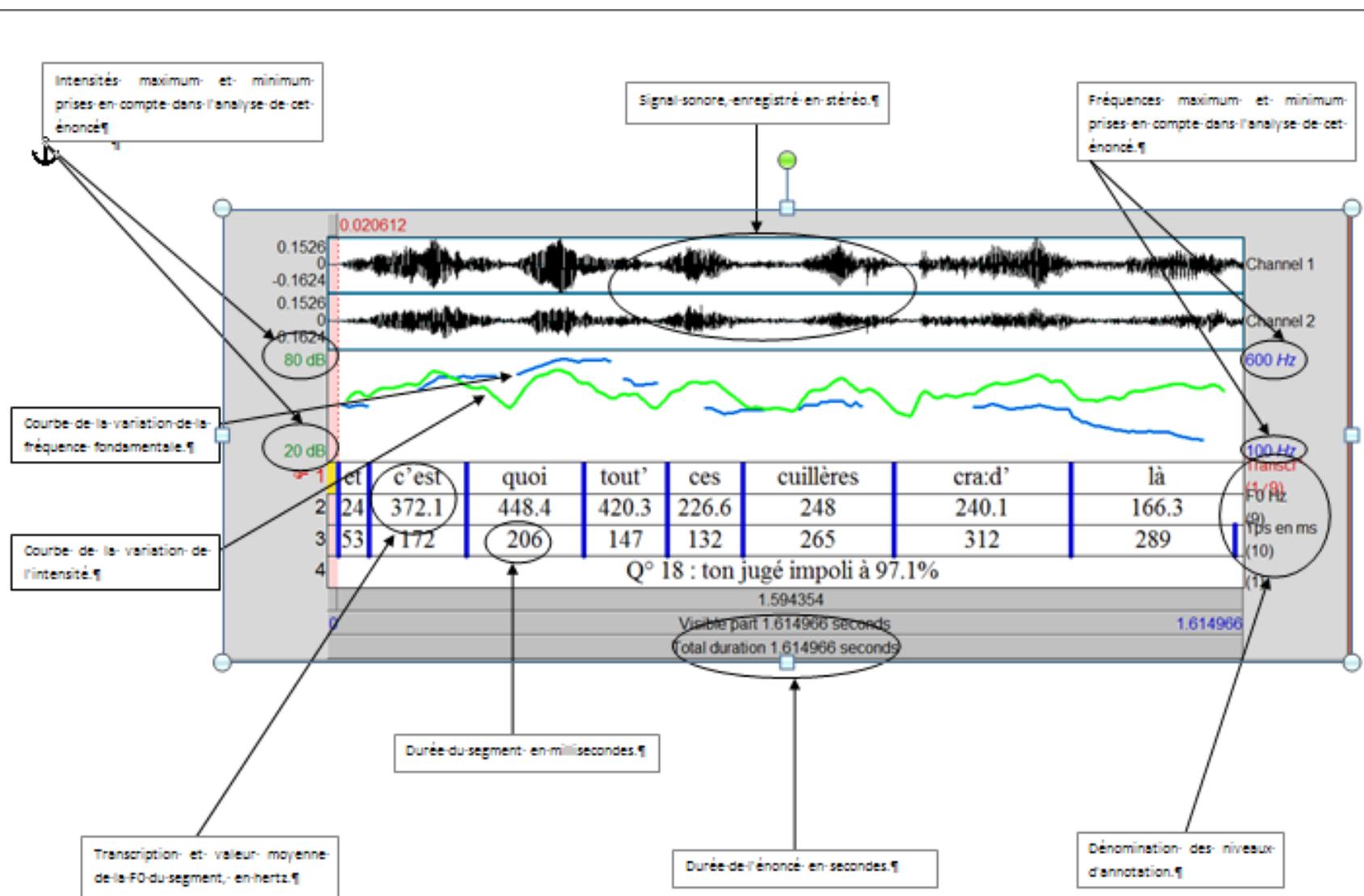


Schéma 1 : Identification des éléments d'une fenêtre Praat.

## 3.2 Résultats significatifs

### 3.2.1 Rappels

Afin de mieux prendre en compte les écarts ou absence d'écarts tonaux que présenteront les énoncés à l'étude, je souhaite procéder ici à quelques rappels.

La courbe de variation de fréquence d'un énoncé français standard, celle qui paraîtra harmonieuse et "neutre"<sup>102</sup> au natif, présente un profil fluide, dont les valeurs montantes ou descendantes sont en général progressives : les montées/descentes par paliers (ou plateaux), les variations brutales et importantes (pics et creux intonatifs), mais aussi l'absence de variations, sont le résultat de l'expression d'affects, et doivent faire l'objet d'interprétations pragmatiques (Caelen-Haumont 2008 a et b).

La syllabe en fin de groupe sémantique (ou rythmique) porte un accent, visible à travers les valeurs de hauteur et de durée - la variation de l'intensité étant moins notable. Delattre (1966) propose dix patrons mélodiques de base qu'il choisit de représenter sur une grille à 4 niveaux. Les patrons se répartissent en 4 modalités de phrase : déclarative, interrogative, parenthétique et exclamative. La modalité déclarative gouverne 3 types d'intonations : i) la continuation mineure - entre 2 groupes syntaxiques fortement liés - varie du niveau 2 à 3, soit d'après mes observations entre 0.5 et 1.5 tons d'augmentation; ses valeurs sont moins élevées qu'en position de ii) continuation majeure - entre 2 groupes syntaxiques dont les liens sont plus lâches, comme pour 2 propositions ou une proposition et un complément circonstanciel - où on passe du niveau 2 au niveau<sup>103</sup> 4, soit entre 1 et 2 tons d'augmentation, toujours d'après les observations que j'ai pu faire à partir du corpus à l'étude. Enfin, iii), la finale descend du niveau 2 au niveau 1. "La validité des contours de Delattre sera partiellement mais non fondamentalement remise en cause par Di Cristo et de nos jours ces types de contours sont communément admis et repris par la communauté scientifique" (Caelen-Haumont, 2008a : 39).

---

<sup>102</sup> On se souvient toutefois que l'intonation neutre est "un mythe" (Rittaud-Hutinet 1995); (voir 1.1.3.3).

<sup>103</sup> Il faut noter que les niveaux ne correspondent pas aux changements exacts de tons mais n'indiquent que les mouvements généraux de la voix : *Il n'est (donc) pas question ici d'établir le nombre de niveaux sur lesquels l'intonation française reposerait si les intonèmes étaient des unités discrètes, comme les phonèmes segmentaux, mais simplement de marquer des points de repère pratiques pour faciliter la représentation schématique des intonations de base. Il est d'ailleurs probable, comme le croit Dwight Bolinger, qu'aucun système de niveaux ne peut fidèlement représenter toutes les intonations d'une langue.* (Delattre 1966 : 7).

Les unités discrètes vocaliques ne présentent pour ainsi dire pas de variation de timbre (Leblanc, 2012 : 13) ou de F0, on peut dire que leur hauteur est globalement stable<sup>104</sup>.

Dans un énoncé interrogatif (Delattre 1967 : 330-332), la note la plus haute, en général 2 niveaux au dessus de la fréquence fondamentale usuelle de la voix, est portée par :

- le mot interrogatif, si la question en contient un;
- la dernière syllabe, si la question est inversée ou de forme affirmative.

Un énoncé exclamatif ou un commandement voient leur valeur de F0 descendre du niveau 4 au niveau 1, de manière plus abrupte pour le commandement (*Ibid.*).

Les caractéristiques phonologiques de la voix (ou phonostyle) varient d'un individu à l'autre, et peuvent dépendre du sexe, de l'âge, des habitudes familiales, de l'origine géographique et socioculturelle, ainsi que de traits physiologiques propres à l'individu. La recherche a cependant pu identifier des niveaux de valeurs moyennes pour les 3 paramètres acoustiques que sont i) la fréquence fondamentale, ii) la durée et iii) l'intensité d'une voix standard.

i) Selon Vaissière (2006, citée par Pépiot 2013 : 25) et Léon & Léon (1997 : 66)<sup>105</sup>, les valeurs moyennes de F0 féminines se situent autour de 240 Hz, de 220 Hz pour Arnold (2012 : 5)<sup>106</sup>. Pour ces auteurs, la valeur moyenne d'une F0 masculine se situe autour de 120 Hz. Certains auteurs présentent une fourchette de variation de 140 à 240 Hz pour la F0 féminine, et de 100 à 150 Hz pour la F0 masculine (*Calliope*, 1989 : 133, cité par Pépiot 2013 : 25).

ii) La valeur moyenne de la durée d'une syllabe non accentuée est de 70 à 80 millisecondes (ms) en conversation ordinaire, et de 150 à 200 ms pour une syllabe accentuée; "mais dans un parler expressif, on peut avoir des syllabes beaucoup plus courtes ou beaucoup plus longues" (Léon & Léon 1997 : 68). Dans le corpus à l'étude, la durée des syllabes non accentuées est comprise entre 51 et 240 ms, et pour les syllabes accentuées entre 122 et 571 ms : ces écarts montrent à quel point la parole expressive vivante est malléable, bien loin des valeurs moyennes de référence. Pour être perçue par l'oreille, l'écart entre 2 durées doit être égale ou supérieure à 20% (Léon 1993 : 95).

---

<sup>104</sup> Hormis, parfois, de légères variations sur la finale d'énoncés interrogatifs de forme grammaticale affirmative.

<sup>105</sup> Voir aussi en 2.1.4.2 Phonostyles des locuteurs, pour les caractéristiques vocales des locuteurs du corpus.

<sup>106</sup> Il semble toutefois que la valeur moyenne de F0 attendue varie sensiblement d'une communauté linguistique à l'autre, ce qui indique que des facteurs sociaux entrent en jeu dans la construction de la hauteur moyenne de la voix, masculine ou féminine (Arnold 2012, Pépiot 2013).

iii) En ce qui concerne l'intensité, la valeur moyenne d'une conversation courante se situe autour de 50 dB; pour la voix chuchotée elle est de l'ordre de 20 dB, et de 70 dB pour une dispute (Richard, Camps *et al.* 2013 : 243). Pour que l'oreille puisse percevoir une augmentation d'intensité, celle-ci doit être égale ou supérieure à 3dB (Léon 1993 : 95).

Il est bien entendu que chaque énoncé présente ses propres caractéristiques acoustiques. Mais en observant et comparant les courbes des 12 énoncés interrogatifs et des 35 énoncés déclaratifs jugés impolis, j'ai pu dégager un certain nombre de phénomènes<sup>107</sup> communs, à partir desquels j'ai constitué des groupes d'énoncés présentant une courbe de F0 proche.

Pour chaque phénomène affectant la courbe de la F0, je mettrai en regard un énoncé impoli caractéristique et un énoncé poli présentant un phénomène approchant, s'il existe, pour tâcher d'éclairer les marques acoustiques qui différencient un énoncé perçu comme poli ou impoli, au delà de la courbe elle même.

### **3.2.2 Comparaison des indices acoustiques des Q° polies et impolies**

J'ai pu repérer pour les énoncés interrogatifs "impolis"<sup>108</sup> 5 caractéristiques dont plusieurs se croisent généralement dans un même énoncé :

- 1) une augmentation brutale très importante de F0 sur la dernière syllabe de l'énoncé (Q°28);
- 2) une augmentation et/ou diminution brutale très importante de F0 en cours d'énoncé (Q°12, 16, 18, 30, 34);
- 3) des augmentations/diminutions compensatoires successives (Q°12, 30, 34 et 41);
- 4) une variation brutale de F0 dans une même syllabe (Q°12, 28, 30, 34, 41);
- 5) une chute brutale de F0 i) sur un segment de plusieurs syllabes en fin d'énoncé (Q°1, 8, 11, 34 et 45); ii) sur la dernière syllabe de l'énoncé (Q°16, 18 et 34).

---

<sup>107</sup> Pour voir l'ensemble des analyses acoustiques, aller en annexes E et F.

<sup>108</sup> L'énoncé Q°32, composé d'un seul mot ("*quoi*"), a été mis de côté car Praat semble fournir des données erronées.

3.2.2.1 Augmentation brutale très importante de F0 sur la dernière syllabe de l'énoncé (Q°28\_I\_91.2)

N° syl	1	2	3	4	5	6		7	
Transcr <sup>o</sup>	t'es	l'pen	seur	de	Ro	din		là	
MO FO Hz	235.1	233.5	276.1	259.5	284.8	372		524.8	
V Ma/Mi FO						321.5	483.5	444.5	590
ton	La#2	La#2	Do#3	Do3	Do#3+	Mi3-	So13-	La3	Ré4
ton						(Fa#3= +3)		(Do4= +3.5)	
aug. tons		=	+2	-0.5	+0.5	+1.5	+2	+1	+3
Tps en ms	171	230	255	134	142	226		368	
total Tps en ms	401		255	134	736				
total aug. tons	=		+2	-0.5	+7				
dB	58	60.5	59.5	59.4	63	62.6		63.43	
position			<b>CMi</b>			<b>Fin.</b>			

Tableau 14 : indices acoustiques de l'énoncé Q°28\_I\_91.2%.

**N° syl** : numérotation des syllabes; **Transcr<sup>o</sup>** : transcription lexicale; **MO FO** : valeur moyenne de la FO sur la syllabe; **V Ma/Mi FO** : valeurs maximales et minimales de la FO sur la même syllabe; **aug. ton** : évolution du ton par rapport à la syllabe précédente; **Tps en ms** : temps en millisecondes; **total aug. tons** : addition des augmentations/diminutions de tons sur un groupe de syllabes; **total Tps en ms** : addition de la durée du groupe de syllabes affecté d'une augmentation/diminution de tons; **dB** : valeur en décibels de l'intensité; **position** : place occupée par la syllabe dans le segment; **CMa** : continuité majeure; **CMi** : continuité mineure; **Fin.** : finale.

Q° 28\_I\_91.2 est produit par FF1, dont les valeurs moyennes de F0 se situe à 302.78 Hz et de fondamental usuel à 220 Hz. L'énoncé, interrogatif, est de structure syntaxique affirmative : à ce titre, il est normal que la F0 s'élève en 6<sup>109</sup> de 2 ou 3 tons. Sa courbe présente globalement un profil distendu, c'est-à-dire que les écarts entre les valeurs de F0 sont supérieures aux valeurs moyennes communément relevées : en 3, la continuité mineure est plus haute qu'attendu, de même que 6 en position de finale. Quant à 7, déictique menaçant à la valeur de durée triple de celle d'une syllabe standard, la F0 y poursuit sa montée en flèche de 3.5 tons supplémentaires.

<sup>109</sup> Ce chiffre renvoie au numéro de la syllabe.

On trouve le même phénomène d'augmentation brutale de la F0 sur la dernière syllabe pour des questions de forme affirmative estimées polies à plus de 80 %<sup>110</sup>. Q°23\_P\_85.3, prononcé par la locutrice FF4, présente les caractéristiques les plus intéressantes.

N° syl	1	2	3	4	5
Transcr <sup>o</sup>	i	l est	sur	face	book
MO F0 Hz	237.5	268.5	296	303.7	533
ton	La#2	Do#3	Ré3	Ré3+	Do4
dB	58.4	57.5	54.3	53.7	49.2
aug. tons		+2	+0.5	=	+6
Tps en ms	82	98	134	276	350
total aug. tons		+2.5		=	+6
total Tps en ms		232			

Tableau 15 : indices acoustiques de l'énoncé Q°23\_P\_85.3%.

Progressive de 1 à 4, la courbe se fait abrupte sur 5 où l'on a une voix de tête à 533Hz soit + 6 tons. La durée des 3 dernières syllabes augmente dans les mêmes proportions que celles de Q°28\_I\_91.2. L'énoncé n'a pourtant pas paru impoli aux testeurs. Le seul paramètre qui semble jouer est l'intensité, régulièrement décroissante en Q°23, alors qu'en Q°28 elle augmente sensiblement sur les 3 dernières syllabes.

La courbe de la F0 ne suffit donc pas dans ce cas à distinguer un énoncé poli d'un énoncé impoli.

3.2.2.2 *Augmentation et/ou diminution brutale très importante de F0 en cours d'énoncé (Q°12\_I\_97.1, Q°16\_I\_79.4, Q°18\_I\_97.1, Q°30\_I\_94.1, Q°34\_I\_94.1)*

L'énoncé Q°18 prononcé par la locutrice FF1 en donne un bon exemple.

Seg. synt.	A			B				
N° syl	1A	2A	3A	1B	2B	3B	4B	5B
Transcr <sup>o</sup>	et	c'est	quoi	tout'	ces	cuillères	crad'	là
MO F0 Hz	247.2	372.1	448.4	420.3	226.6	248	240.1	166.3

<sup>110</sup> Q°20, 23 et 29.

V Ma/MI F0								190.5	143
ton	Si2	Fa#3	La3	Sol#3	La2	Si2	Si2	Sol2-	Re2-
ton								(Mi2 = -4)	
aug. tons		+4.5	+1.5	-0.5	-6.5	+1	=	-2	-3
Tps en ms	53	172	206	147	132	265	312	289	
total aug. ton		+6		-7		+1	=	-4	
total Tps en ms		378		279		577		289	
position			CMa				Fin.	parenthèse	

Tableau 16 : indices acoustiques de l'énoncé Q°18\_I\_97.1%.

On observe une augmentation de 6 tons très rapide, en l'espace de 2 syllabes (2A et 3A), puis immédiatement après une chute de 7 tons (1B et 2B), enfin une chute supplémentaire de 4 tons sur 5B. La courbe de F0 présente donc un profil extrêmement distendu, avec des écarts aux valeurs très supérieures à celles des énoncés polis, pour lesquels le corpus ne fournit pas de caractéristiques comparables.

3.2.2.3 *Augmentations/diminutions compensatoires successives (Q°12\_I\_97.1, Q°30\_I\_94.1 et Q°41\_I\_97.1)*

L'énoncé Q°12\_I\_97.1, produit par FF3 (MO F0 251.21 Hz, F0 usuel 200 Hz), en est caractéristique :

Seg. synt.	A		B			
N° syl	1A	2A	1B	2B	3B	4B
Transcr <sup>o</sup>	c'est	quoi	cette	co	mmu	ni
MO F0 Hz	287.3	320	219.4	328.6	∅	∅
ton	Re3	Mi3-	La2	Mi3	∅	∅
aug. tons		+1	-4	+4	=?	=?
Tps en ms	153	140	231	244	111	98
total Tps en ms				453		
dB	57.3	60.1	56	61.2	59.5	56.3
position		CMa				

Seg. synt.	B					C	
N° syl	5B	6B	7B	8B	9B	1C	2C
Transcr <sup>o</sup>	ca	tion	à	la	noix	ici	euh

MO F0 Hz	301	301	299	272.3	301	268	Ø
ton	Re3+	Re3+	Re3+	Do#3	Re3+	Do#3	Ø
aug. tons	-1	=	=	-1	+1	-1	=?
Tps en ms	120	230	108	178	316	416	366
total Tps en ms	518						
dB	65.7	61.6	65.4	61.3	60.6	60	66.7
position		<b>CMi</b>			<b>CMa</b>		<b>Fin.</b>

Tableau 17 : indices acoustiques de l'énoncé Q°12\_P\_97.1%.

On observe dans cet énoncé que: a) 2 plateaux de 3 syllabes se suivent : le premier, supposé à l'écoute, en 2B, 3B et 4B, arrive immédiatement après 1B, et le second avéré en 5B, 6B et 7B; b) un écart de 4 tons sépare 2A et 1B ainsi que 1B et 2B; c) les tons perdus en 1B sont immédiatement regagnés en 2B, de même pour 8B, 9B, 1C et 2C. C'est ce phénomène que j'ai appelé augmentation/diminution compensatoire; la sensation auditive, désagréable, peut être associée à celle provoquée par une sirène<sup>111</sup>.

Un seul énoncé parmi les 11 jugés polis présente un phénomène partiellement comparable, Q°25\_P\_91.2. Produit par la locutrice FF2, (MO F0 227.45 Hz, fondamental usuel 180 Hz), Q°25 peut se scinder en 2 segments, A et B, séparés par une hésitation en 4A. Q°25 a paru poli à 91.2% des testeurs alors que a) la courbe de F0 montre 2 pics intonatifs très importants en 3A et 5B, immédiatement compensés en 4A et 6B, et que b) la F0 descend progressivement de 1B jusqu'à 4B; le mouvement de courbe de 5B à 8B serait quasi parallèle à celui de 1B à 4B, s'il n'y avait pas le pic intonatif sur 5B. La fin de la courbe est donc inverse à celle attendue pour une question totale de forme grammaticale affirmative<sup>112</sup>.

Tableau 18 : indices acoustiques de l'énoncé Q°25\_P\_91.2%.

<b>Seg. synt.</b>	<b>A</b>			
<b>N° syl</b>	<b>1A</b>	<b>2A</b>	<b>3A</b>	<b>4A</b>
<b>Transcr°</b>	et	y a	pas	des
<b>MO F0 Hz</b>	208	231	359	223
<b>ton</b>	Sol#2	La#2	Fa3+	La2
<b>aug. tons</b>		+1	<b>+4.5+</b>	<b>-5</b>
<b>Tps en ms</b>	104	122	174	316

<sup>111</sup> Une sirène est un dispositif d'alarme sonore électronique ou électromécanique, caractérisée par un son montant/descendant d'un écart en général de 2 tons.

<sup>112</sup> Pour le phénomène b), voir aussi 5) i) ci-dessous.

<b>dB</b>	54	57	56.2	62
<b>position</b>			<b>CMi</b>	<b>CMa</b>

<b>Seg. synt.</b>	<b>B</b>							
<b>N° syl</b>	<b>1B</b>	<b>2B</b>	<b>3B</b>	<b>4B</b>	<b>5B</b>	<b>6B</b>	<b>7B</b>	<b>8B</b>
<b>Transcr<sup>o</sup></b>	des	li	brai	ries	où	ils	le	font
<b>MO FO Hz</b>	193.4	184	182.6	169	324	186.5	171	156.4
<b>ton</b>	Sol2	Fa#2	Fa#2	Mi2	Mi3	Fa#2	Fa2	Ré#2
<b>aug. tons</b>	-1	-0.5	=	-1.5	<b>+7</b>	<b>-5.5</b>	-0.5	-0.5
<b>Tps en ms</b>	124	145	242	176	129	115	180	262
<b>dB</b>	58	60	60.5	57.5	55.9	58	54.5	49
<b>position</b>				<b>CMi</b>				<b>Fin.</b>

Pour les raisons a) et b), l'énoncé aurait pu être catégorisé "impoli". Il est toutefois remarquable que dans Q°25\_P\_91.2, le pic d'intensité intervient dans la courbe comme une exception : une seule syllabe du segment porte une note très élevée, mais dès la syllabe suivante la courbe retrouve une forme progressive. Ce n'est pas les cas pour Q°12, 30 et 41, qui présentent des courbes autrement perturbées<sup>113</sup> : on a parlé des plateaux intonatifs hauts pour Q°12; Q°30 voit sa dernière syllabe s'étirer jusqu'à 571 ms, ses deux dernières syllabes montrent une F0 instable avec des variations importantes au sein des syllabes<sup>114</sup>. Dans Q°41, on observe également 2 plateaux (1<sup>115</sup> - 2 et 4 - 5) et une instabilité de F0 (4). En outre, les valeurs moyennes d'intensité par syllabe se situent en deçà de 60.5 dB pour Q°25, mais au-delà pour Q°12, 30 et 41 (entre 57.5 et 71 dB).

Q°12, 30 et 41 ont reçu des taux d'impolitesse très élevés entre 94.1 et 97.1%. Il semble donc que ce type de courbe soit caractéristique des énoncés ressentis comme très impolis.

<sup>113</sup> J'ai vérifié que comme pour Q°25, les pics de Q°30 et 41 ne soient pas suivis de paliers intonatifs, et que la variation de hauteur se fasse vers l'aigu puis le grave et non du grave vers l'aigu comme dans Q°12 - ce qui aurait pu éventuellement jouer sur la perception d'impolitesse.

<sup>114</sup> Voir ci-dessous en 4).

<sup>115</sup> Ces chiffres renvoient au numéro de syllabe.

3.2.2.4 Une variation brutale de F0 dans une même syllabe (Q°12\_I\_97.1, Q°28\_I\_97.1, Q°30\_I\_94.1, Q°34\_I\_94.1, Q°41\_I\_94.1)

Q°34\_I\_94.1, produit par FF1 (MO F0 302.78 Hz; F0 usuel 220 Hz), illustre bien ces variations de F0 au sein d'une même syllabe, en 2, 7 et 8<sup>116</sup>, avec des valeurs extrêmement importantes, de l'ordre de l'octave. Or on sait que les voyelles du français parlé contemporain ne présentent pas ce caractère d'instabilité car elles sont décrites comme stables en hauteur. Il faut souligner ici que lorsqu'une syllabe porte de telles variations de valeur de F0, c'est la dernière partie de la fluctuation que l'auditeur garde en dernier ressort en mémoire; c'est pourquoi j'ai calculé l'écart de F0 de la syllabe suivante à partir de la dernière valeur fluctuante et non à partir de la valeur moyenne de F0 de la syllabe.

Tableau 19 : indices acoustiques de l'énoncé Q°34\_I\_94.1%.

N° syl	1	2	3	4	5	6	7	8		
Transcr <sup>o</sup>	t'as	vu	c'que	t'as	dans	ton	bol	toi		
MO F0 Hz	∅	288	235.1	373.5	331	292.6	266.1	219.6		
ton	∅	Re3	La#2	Fa#3	Mi3	Re3	Do3	La2		
aug. tons		?	-2.5	+6	-1.5	-1	-1	-2		
VMa/Mi F0 Hz		214	439.2				205	391.3	277.6	186.4
ton		La2-	La3				Sol#2	Sol3	Do#3	Fa#2
aug. tons			+1 oct.				-3.5	+7.5	-4	-4
Tps en ms	164	188	188	122	82	185	267	261		
position		CMa		CMi			CMa	Fin.		

Le corpus ne fournit pas de phénomène équivalent parmi les énoncés interrogatifs jugés polis, dont les variations intonatives sur une même syllabe ne dépassent pas 2.5 tons.

Ces 5 énoncés ont reçu des taux d'impolitesse très élevés entre 91.2 et 97.1% : il semble que ces variations de hauteurs sur une même voyelle soient caractéristiques des énoncés ressentis comme très impolis.

<sup>116</sup> La syllabe 8 doit être également rapprochée du paragraphe suivant 5) ii) concernant les chutes importantes de F0 sur la dernière syllabe.

### 3.2.2.5 Chute brutale de F0

i) sur un segment de plusieurs syllabes en fin d'énoncé (Q°1\_I\_97.1, Q°8\_I\_76.5, Q°11\_I\_94.1, Q°34\_I\_94.1 et Q°45\_I\_76.5)

Seg. synt.	A								
N° syl	1A	2A	3A	4A	5A	6A	7A	8A	9A
Transcr <sup>o</sup>	i'	pass'	ses	jour	nées	sur	in	ter	net
MO F0 Hz	341.8	384.8	341	315	315	338.8	347	351.5	366.2
ton	Fa3	Sol3	Fa3	Ré#3	Ré#3	Fa3-	Fa#3	Fa3	Fa#3
aug. tons		+1	-1	-1.5	=	+1.5	+0.5	-0.5	+0.5
Tps en ms	207	185	111	131	161	179	98	150	204
total aug. tons		+1	-2.5			+2		-0.5	+0.5
total Tps en ms		185	403			257		150	204
position		CMi			CMi				CMa

Seg. synt.	B			
N° syl	1B	2B	3B	4B
Transcr <sup>o</sup>	alors	qu'est-ce	tu	veux
MO F0 Hz	217.8	220.7	241	256.7
ton	La2	La2	Si2	Do3
aug. tons	-5.5	=	+1	+1
Tps en ms	221	192	126	202
total aug. tons	-5.5		+2	
total Tps en ms	413		328	
position	CMi			Fin.

Tableau 20 : indices acoustiques de l'énoncé Q°1\_I\_97.1%.

L'énoncé Q°1\_I\_97.1, produit par la locutrice FF1 (MO F0 302.78 Hz, F0 usuel 220 Hz), peut se scinder en 2 segments, A et B, un par acte de langage : un constat, et une fausse question qui demande en fait l'assentiment et/ou la compassion de l'interlocuteur. La F0 se maintient sur A à une hauteur assez élevée. On observe à l'attaque de B un décrochage de 5.5 tons; B est formulé à la hauteur du fondamental usuel, et peut être considéré comme un commentaire : le mot interrogatif ne porte pas l'accent typique de hauteur et d'intensité.

A contrario, l'énoncé Q°25 a paru poli à 91.2% des testeurs alors que la fin de la courbe est décroissante comme dans Q°1 : la F0 descend progressivement de 1B jusqu'à 4B et le mouvement de courbe de 5B à 8B serait quasi parallèle à celui de 1B à 4B s'il n'y avait pas le pic intonatif sur 5B. La fin de la courbe est donc inverse à celle attendue pour une question totale de forme grammaticale affirmative, ce que l'on peut interpréter ici comme un sentiment d'étonnement légèrement fâché ou de suspicion.

La différence majeure entre Q°1 et Q°25 tient à l'intensité : la valeur moyenne de l'intensité des 2 segments B est comparable, respectivement 57.2 dB et 59.5 dB. Les segments A en revanche présentent une valeur moyenne respectivement de 63.3 dB et 59.5 dB. Ce serait donc l'intensité élevée de A Q°1 ainsi que la différence d'intensité entre les segments A et B Q°1 qui produiraient la sensation d'impolitesse. On ne peut éviter par ailleurs de se demander si le sens n'a pas influencé le jugement des testeurs.

*ii) sur la dernière syllabe de l'énoncé (Q°16\_I\_79.4, Q°18\_I\_97.1 et Q°34\_I\_94.1)*

Les énoncés Q°16, Q°18 et Q°34, prononcés par FF1 (MO F0 302.78 Hz; F0 usuel 220 Hz), voient leur F0 décroître fortement sur la dernière syllabe, qui porte en même temps une fluctuation importante de hauteur.

N° syl	1	2	3	4	
Transcr <sup>o</sup>	en	quel	ho	nneur	
MO F0 Hz	105.4 (bug?)	240	208.5	210	
VMa/Mi F0 Hz				242.5 VMaF0	169.5 VMiF0
ton	Sol#1	Si2	Sol#2	Si2	Mi2+
ton				(Sol#2 =)	
aug. tons		+1 oct 0.5 tons	-1.5	+1.5	-4
Tps en ms	140	137	124	353	
Tps en ms				193	126

Tableau 21 : indices acoustiques de l'énoncé Q°16\_I\_79.4%.

Dans l'énoncé Q°16\_I\_79.4, 1 démarre très bas pour cette locutrice, à moins que la valeur ne soit erronée; 2, mot interrogatif, porte l'accent d'intensité et de hauteur : a) ce qui crée une différence de F0 d'une octave et 0.5 ton entre les 2 syllabes, en 137 ms. 4 remonte à 210 Hz, comme il est normal pour une question de ce type. Il faut cependant y regarder de plus près : sur cette syllabe, b) la F0 retombe en l'espace de 126 ms de 4 tons; la fin de la syllabe est donc bien moins aigüe que ce que le laisse croire la valeur moyenne de la syllabe. c) 4 est aussi bien plus longue que la durée standard, avec un étirement qui occupe à lui seul la moitié de la durée de l'énoncé. On peut imputer la sensation d'impolitesse à l'ensemble des phénomènes a, b et c.

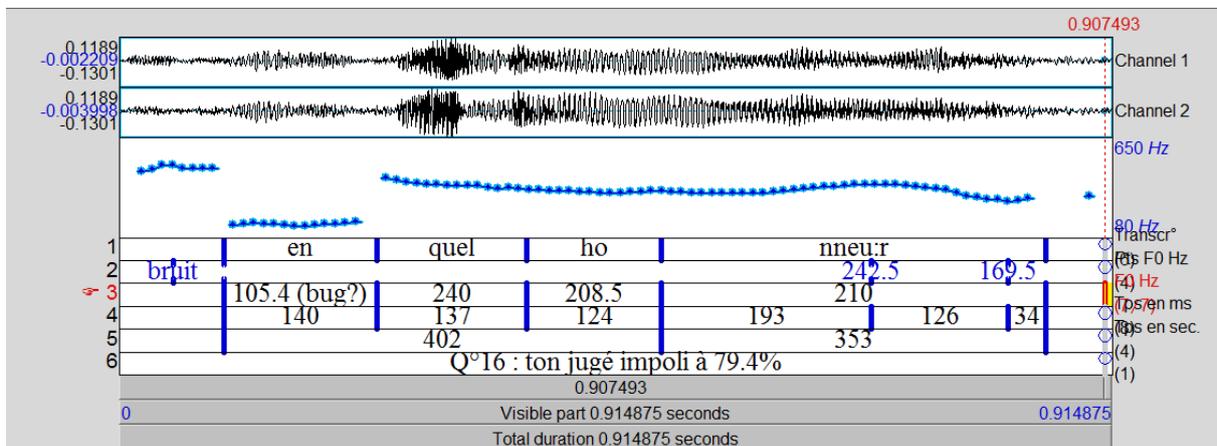


Schéma II : Capture d'écran de la fenêtre d'analyse acoustique de Praat pour la Q°16\_I\_79.4%.

Le corpus ne fournit pas de courbe équivalente parmi les énoncés interrogatifs élicités polis.

### 3.2.2.6 Conclusion de l'analyse acoustique des énoncés Q°

Arrivés à ce point, voici les observations que l'on peut faire :

A courbe de F0 comparable, c'est l'intensité qui permet de discriminer auditivement un énoncé impoli d'un énoncé poli.

Un énoncé interrogatif est perçu comme impoli lorsque :

1) la courbe de F0 présente des écarts entre les valeurs de F0 supérieures aux valeurs standard attendues (valeurs distendues);

2) la courbe de F0 montre des augmentations/diminutions de valeurs compensatoires successives;

3) des variations de F0 interviennent au sein d'une même syllabe (instabilité syllabique);

4) plusieurs syllabes à la suite ont la même valeur de F0 (plateaux intonatifs);

5) la courbe chute brutalement et de manière importante sur la dernière syllabe de l'énoncé. Il est à noter que dans le corpus, ce phénomène est toujours associé à celui de l'instabilité syllabique.

Plus les écarts de valeurs de F0 sont importants et plus les phénomènes précédemment cités se rencontrent dans le même énoncé, plus l'énoncé est perçu "impoli".

### 3.2.3 Comparaison des indices acoustiques des D° polies et impolies

Dans le cas des énoncés déclaratifs, il m'a paru préférable de présenter d'abord les caractéristiques des énoncés jugés "polis", ce qui permettra de mettre en valeur leurs caractéristiques pour les opposer ensuite à celles des énoncés jugés "impolis".

#### 3.2.3.1 Description et analyse des indices acoustiques des énoncés D° jugés "polis"

Les énoncés D° jugés "polis", au nombre de 16, ont tous au moins l'une des deux caractéristiques suivantes :

a) La courbe de F0 suit un schéma bien identifiable, formé d'un ou plusieurs sous groupes intonatifs convexes - en forme de "u" écrasé -, délimités par les positions syntaxiques de continuations mineure, majeure ou de finale.

*Exemple :*

Tableau 42 : illustration par les indices acoustiques d'un énoncé D° poli de type a) standard (D°43).

Seg. synt.	A			B					C				
N° syl	1A	2A	3A	1B	2B	3B	4B	5B	1C	2C	3C	4C	
Transcr <sup>o</sup>	et	a	près	à	qua	tre	heu	re et d'	mie	y a	Charlotte	qui	vient
MO F0 Hz	178.9	168.3	196.7	185.1	181.7	194.9	186.7	194.5	194.3	181.6	186.2	172.6	
ton	Fa2	Mi2	Sol2	Fa#2	Fa#2-	Sol2	Fa#2	Sol2	Sol2	Fa#2-	Fa#2	Fa2	
aug. tons		-/	+2	-0.5	=	+0.5	-0.5	+0.5	=	-0.5	=	-0.5	

Tps en ms	202	153	175	136	116	302	172	178	138	203	138	277
position			CMa			CMi		CMa		CMi		Fin.

**A l'intérieur de ces groupes**<sup>117</sup>, les valeurs de F0 ne dépassent pas, sauf exceptions, le ton et demi. Les exceptions sont i) D°63\_P\_91.2<sup>118</sup> où il s'agit de mettre en relief le nom d'une personne qui sert de caution au propos; ce nom, en 3 syllabes dont la dernière est en position de continuation majeure, porte les valeurs d'augmentation/diminution de F0 de +2/-4/+2.5; le reste de l'énoncé contient des segments syntaxiques standards ou plats, dont le dernier forme une courbe descendante régulière<sup>119</sup>. Le taux de politesse pour cet énoncé est de 91.2%, pour une locutrice (FF3) dont les modulations vocales sont assez importantes. ii) D°32\_P\_85.3<sup>120</sup> : "la" porte une chute de 6 tons destinée à la focalisation du mot suivant, "conséquence", dont la première syllabe montre une augmentation d'1 octave et 0.5 ton. Il s'agit ici d'attirer l'attention de l'interlocuteur sur les mots qui véhiculent pour le locuteur le cœur du message, ou focalisation. D°32 obtient un taux de politesse de 85.3%.

**En finale de groupe**<sup>121</sup>, la valeur de F0, descendante en finale et montante en position de continuation majeure, ne dépasse en général pas les 2.5 tons, et le demi ton pour les continuations mineures. Les exceptions sont i) D°32\_P\_85.3<sup>122</sup> : "vous", en position de continuation majeure, subit une augmentation de 5 tons. ii) D°31\_P\_97.1<sup>123</sup>, énoncé exclamatif, donne 2 syllabes en position de continuation majeure, "ouais" et "cool", qui portent respectivement un accent de hauteur de +3 et de +3.5 tons. iii) La dernière exception, D°81<sup>124</sup> avec un taux de politesse de 82.4%, contient une finale qui perd 6 tons. Cet énoncé peut aussi être classé dans le groupe b).

**b)** La courbe de F0 descend régulièrement et progressivement entre la première et la dernière syllabe. Les continuations mineures sont faiblement marquées (+0.5), et le plus souvent elles sont non marquées : elles portent la même valeur de F0 que la syllabe précédente, voire une valeur inférieure jusqu'à -1.5. Les continuations majeures continuent la

<sup>117</sup> C'est-à-dire non prises en compte les syllabes en position de continuation mineure ou majeure.

<sup>118</sup> Voir fenêtre Praat en annexe F'.

<sup>119</sup> Voir b) ci-dessous.

<sup>120</sup> Voir fenêtre Praat en annexe F'.

<sup>121</sup> C'est-à-dire en position de continuation mineure, majeure, ou finale.

<sup>122</sup> Voir fenêtre Praat en annexe F'.

<sup>123</sup> Voir fenêtre Praat en annexe F'.

<sup>124</sup> Voir fenêtre Praat en annexe F'.

plupart du temps d'être normalement marquées (+2 à +2.5 tons), mais elles peuvent prendre parfois des valeurs négatives (-1.5) comme en 9B de D°63\_I\_91.2.

*Exemple :*

Tableau 23 : illustration par les indices acoustiques d'un énoncé D° poli de type b) descendant (D°22).

Seg. synt.	A		B		C							
	1A	2A	1B	2B	1C	2C	3C	4C	5C	6C	7C	8C
N° syl												
Transcr <sup>o</sup>	et	ça	en	plus	ça	le	met	dans	z'une	ra	ge	noire
MO FO Hz	162.6	239.3	195.4	204.4	175.3	164.4	161.2	157.7	152.4	159	157.2	201.5
ton	Mi2	La#2	Sol2	Sol#2	Fa2	Mi2	Mi2-	Ré#2	Ré#2-	Ré#2	Ré#2	Sol2
aug. tons		+3.5	-1.5	+0.5	-1.5	-/	=	-0.5	=	=	=	+2.5
Tps en ms	121	303	149	168	92	79	83	166	383	109	144	381
position		CMa		CMa			CMi			CMi		CMa

Seg. synt.	D					
	1D	2D	3D	4D	5D	6D
N° syl						
Transcr <sup>o</sup>	parce	que	<i>silence</i>	c'est	des	lâches
MO FO Hz	∅	172.2		145.1	140.8	139.5
ton	(190?)	Fa2		Ré2	Do#2	Do#2
aug. tons	-	-1.5	<i>silence</i>	-2	-0.5	=
Tps en ms	241	400	499	302	124	411
position		CMi				Fin.

### 3.2.3.2 Description et analyse des indices acoustiques des énoncés D° jugés "impolis"

Les énoncés jugés impolis présentent tous l'une ou plusieurs des caractéristiques suivantes :

a) Pour 10 énoncés<sup>125</sup>, la courbe contient un ou plusieurs segments plats que l'on peut désigner par le terme de "**plateaux**" : la valeur de F0 se maintient à l'identique sur plusieurs syllabes consécutives au lieu de prendre la forme convexe évoquée précédemment : les

<sup>125</sup> D°15, 19, 34, 37, 57, 60, 67, 76, 90 et 99.

continuations mineures ne sont pas marquées, les continuations majeures ne s'élèvent pas au dessus d'1 ton.

Seg. synt.	A			B				C	
N° syl	1A	2A	3A	1B	2B	3B	4B	1C	2C
Transcr <sup>o</sup>	ah	là	là	ça	m'a	gon	flée	ce	truc
MO F0 Hz	142.2	141.7	144.5	144	141.3	149.5	163.4	176.1	162.8
ton	Ré2-	Ré2-	Ré2	Ré2	Ré2-	Ré2+	Mi2	Fa2	Mi2
aug. tons		=	=	=	=	=	+/-	+/-	-1
Tps en ms	79	80	108	138	110	<b>229</b>	191	91	595
dB	48.4	51.7	52.7	50.6	50	57	56	52.9	45.5
position			<b>CMa</b>				<b>Fin.</b>	<b>écho/parenthèse</b>	

Tableau 24 : indices acoustiques de D°15\_I\_91.2%.

Dans D°15\_I\_91.2, le segment A, exclamatif, ne porte aucune modulation; 1B en début de segment aurait dû être légèrement plus basse que 3A, et la F0 aurait dû remonter sur 3B en position de pré-finale pour descendre en finale sur 4B.

Il est d'ores et déjà important de souligner que cette configuration ne suffit pas à générer la perception d'impolitesse, puisque l'on trouve quelques énoncés "polis" avec segments en plateaux<sup>126</sup>, en bien moins grand nombre il est vrai. A la différence des énoncés "impolis" cependant, ils ne présentent pas d'autres phénomènes acoustiques associés, développés ci-dessous. Ainsi D°15, outre le plateau initial, montre un transfert d'accent de 3B sur 4B, ce qui affecte les 2 syllabes : 3B est plus basse et 4B plus haute qu'attendu. Y sont associés 2 accents expressifs : un accent de durée sur 3B et, contre toute attente, une baisse de F0 sur 2C (si le segment est un écho), ou un accent de hauteur sur 1C (si le segment est une parenthèse).

b) 18 énoncés<sup>127</sup> portent divers **accents expressifs** sur les CMa/Mi. On distinguera ce groupe d'énoncés impolis du groupe e) en ce que i) les valeurs, quoique plus élevées que d'ordinaire, resteront ici raisonnables : les écarts ne dépasseront pas 3 tons; ii) les accents

<sup>126</sup> D°54, 63 et 94.

<sup>127</sup> D° 15, 17, 20, 23, 29, 34, 41, 51, 60, 67, 68, 74, 89, 90, 96, 99, 109 et 111.

expressifs de hauteur sont toujours en position de CMa, de CMI ou de finale. Par exemple, dans l'énoncé D°17\_I\_79.4,

Tableau 25 : indices acoustiques de D°17\_I\_79.4%.

N° syl	1	2	3	4	5	6	7	8
Transcr <sup>o</sup>	ça	va	pas	ren	trer	dans	la	case
FD Hz	187	172	<b>235</b>	∅	286	257.5	245	202
ton	Fa#2	Fa2	La#2	∅	Re3-	Do3-	Si2	Sol#2-
aug. ton	∅	-0.5	<b>+2.5</b>	-	<b>+2.5</b>	-1	-1	<b>-1.5</b>
Tps en ms	178	153	171	<b>231</b>	148	145	123	<b>564</b>
position			<b>CMi</b>		<b>CMi</b>			<b>Fin.</b>

on voit que i) 3 et 5, les 2 CMi, portent chacune une augmentation de 2.5 tons alors que d'ordinaire elles se cantonnent à une augmentation de 0.5 à 1 ton; que ii) 4 porte un accent de durée; iii) que 8, la syllabe finale, est un peu plus bas qu'attendu et surtout s'étire bien au delà des valeurs standard.

Il faut cependant noter que 14 D° "polis"<sup>128</sup> sur 16 montrent le même type d'accents expressifs. Cela représente une proportion très importante et laisse penser que le phénomène d'accent expressif à hauteur modérée en position de CMa, CMI ou finale affecte aussi bien les énoncés polis qu'impolis.

c)<sup>129</sup> La courbe forme un **angle** relativement **aigu** : elle monte de manière significative sur le premier segment et redescend sur le second dans les mêmes proportions. Les 2 segments prennent donc des contours très différents de celui décrit ci-dessus en 1) a).

<sup>128</sup> D°2, 3, 16, 26, 31, 32, 54, 62, 63, 73, 79, 81 et 94.

<sup>129</sup> D°4, 23, 53, 66, 77, 90 et 99.

Seg. synt.	A						B		
N° syl	1A	2A	3A	4A	5A	6A	1B	2B	3B
Transcr <sup>o</sup>	et	j'vais	au	Ca	na	da	A	man	dine
FO Hz	232	269	341	404.5	448	384	302	265	219
ton	La#2	Do3+	Fa3	Sol3+	La3+	Fa#3+	Ré3+	Do3+	La2
aug. ton	∅	+1.5	+3	+1+	+1	-1.5	-2	-1	-2
total aug. tons	<b>+7+</b>					<b>-6.5</b>			
Tps en ms	131	197	148	174	170	<b>282</b>	70	<b>203</b>	<b>594</b>
total Tps en ms	<b>820</b>					<b>1149</b>			
position		<b>CMi</b>				<b>CMA</b>			<b>Fin.</b>

Tableau 26 : indices acoustiques de D°66\_I\_82.4%.

On remarque en outre que i) les durées de syllabe sont bien plus longues sur B pour une syllabe de moins, avec un fort contraste entre la première syllabe du prénom, extrêmement brève, et les 2 autres; que ii) la CMA en 6A chute au lieu de monter; enfin que iii) 1B, 2B et 3B, une apostrophe, auraient pu constituer un écho (valeurs hautes), mais ici la FO descend significativement à chaque syllabe.

Aucun énoncé jugé "poli" ne présente de courbe comparable.

d) <sup>130</sup> Une même syllabe voit sa valeur de FO varier brutalement en l'espace de quelques millisecondes<sup>131</sup>. Pour les 12 énoncés concernés, les écarts sont compris entre 3 tons et 1 octave et 3.5 tons. On trouve de 2 à 3 **voyelles instables** par énoncé "impoli".

Seg. synt.	C			
N° syl	1C	2C		
Transcr <sup>o</sup>	c'est	tout		
MO FO Hz	326	297		
points de FO		271	520	233
ton		Do#3	Do4	La#2
aug. tons		-1.5	+7	-1 oct.
ton	Mi3	Ré3		

<sup>130</sup> D°20, 29, 30, 41, 53, 57, 58, 66, 71, 105, 109 et 111.

<sup>131</sup> Voir aussi 3.2.2.4.

aug. tons	-1.5	(-)
Tps en ms	159	556
dB	57.7	63.2
position		Fin.

Tableau 27 : indices acoustiques de D°20\_I\_91.2%.

N° syl	0A	1A			2A	3A	1B	2B		3B
Transcr°	à	l'e			nne	mi	e	nne		mi
MO FO Hz	161	212			221.5	214	186	204.5		189
Points de FO		165	194	243				187	225.5	
ton	Mi2	Mi2 =	Sol2 +2	Si2 +2	La2	Sol#2+	Fa#2	Fa#2	La2	Fa#2
aug. tons		=	+2	+2				=	+1.5	
ton		Sol#2						Sol#2		-1
aug. tons			+2.5		+0.5	-0.5	-1	+1		
Tps en ms	121	339						344		291

Tableau 28 : indices acoustiques de D°41\_I\_85.3%.

Dans le cas de D°20\_I\_91.2 et D°41\_I\_85.3, l'évolution de la courbe de F0 dans la fenêtre Praat montre très nettement les variations de fréquence.

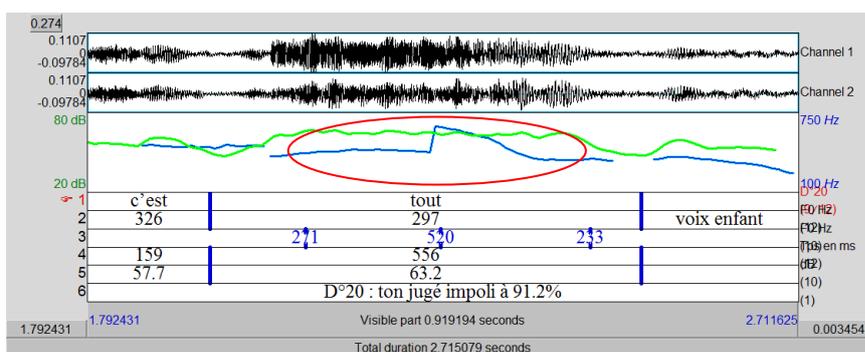
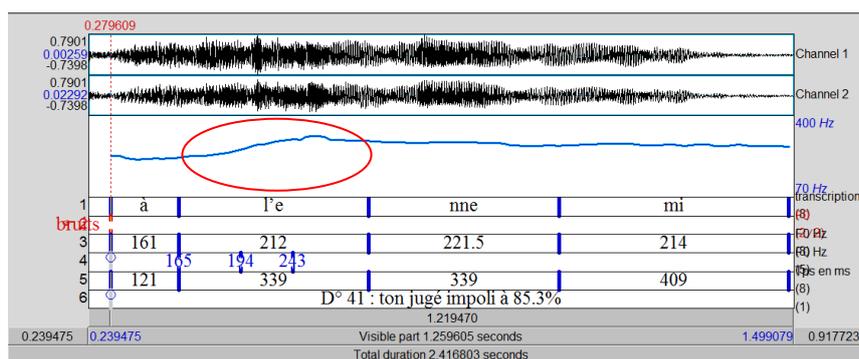


Schéma III : Capture d'écran des fenêtres d'analyse acoustique de Praat pour la D°20\_91.2% et la D°41\_85.3%.



4 énoncés classés "polis" montrent des voyelles instables, une par énoncé (3 fois : D°3\_P\_91.2, D°22\_P\_91.2 et D°31\_P\_97.1), ou deux (1 fois : D°26\_P\_85.3); les écarts sont compris entre 3 et 7.5 tons. D°3, 26 et 31 sont des énoncés exclamatifs qui ont été perçus comme positifs. Pour D°3 et 31 c'est effectivement un mot exclamatif qui porte la voyelle instable, "ah" et "ouais" en position de CMa. Pour D°22 et 26, il s'agirait davantage d'un accent d'insistance en position de CMa pour le premier et de CMi pour le second. Comme en a) et b), on ne peut se fier seulement à la présence de voyelles instables pour déterminer l'impolitesse d'un énoncé, il faut aussi regarder les autres indices qui ne manquent pas d'apparaître.

e) 11 énoncés impolis<sup>132</sup> présentent des écarts aux **valeurs de F0 très distendues**, bien au-delà de ce que montrent les énoncés "polis". 1/3 au moins des syllabes sont affectées par rapport à la syllabe précédente de variations comprises entre 3 tons et 1 octave et 3.5 tons. On remarque dans D°111 pris en exemple ci dessous, que la CMa en 3A-4A et la CMi en 3B portent les accents de hauteur attendus, au dessus des valeurs standard toutefois pour la CMa. En revanche, 4 autres syllabes sans position prosodique spécifique portent des accents de F0 : 1A, 1B, 2B, 3B et 4B. Ce phénomène est très intéressant car il affecte la plupart des énoncés du groupe e) : les accents aux valeurs distendues, ici de hauteur, ajoutent de l'impolitesse **lorsqu'ils se placent en dehors des positions de CMa, CMI, voire finale**, comme on va le voir en f).

Tableau 29 : indices acoustiques de D°111\_1\_88.2%.

Seg. synt.	A				B								
N° syl	1A	2A	3A	4A	1B	2B	3B			4B	5B	6B	
Transcr <sup>o</sup>	c'est	po	ssi	ble	puis	que	lui			il	l'a	fait	
MD FD Hz	201.4	254.6	394.5	413.4	147.6	243.5	258.1			198	217	213.5	
Points de FD	239	167.7					208.8	264.8	310				
ton	La#2	Mi2					Sol#2	Do3+	Ré#3				
ton	(Sol#2-)	Do3-	Sol3	Sol#3	Ré2	Si2	(Do3 = +1)			Sol2	La2-	La#2-	
aug. ton		-3.5	<b>+5</b>	<b>+3</b>	+0.5	<b>-1 oct. - 2.5</b>	<b>+5</b>	-1.5	+2.5	+1.5	<b>-4</b>	+1	-0.5
Tps en ms	262	197	171	180	116	126	359			99	133	404	
total aug. ton			<b>+ 1 oct. + 0.5</b>		<b>- 1 oct. - 2.5</b>	<b>+5</b>	<b>(+1)</b>			<b>-4</b>	<b>+1</b>		
total Tps en ms			548									537	

<sup>132</sup> D°20, 30, 57, 67, 71, 74, 95, 96, 105, 109 et 111.

dB	66.7	69	67	66.5	62.4	62.1	70.6	66.9	67.9	57.1	
position			CMa				CMi				Fin.

Aucun énoncé jugé "poli" ne présente de phénomène comparable.

f) Un très grand nombre (20) d'énoncés "impolis"<sup>133</sup> montrent des **finale**s à la valeur de **F0 trop élevées** par rapport à l'attente. La plupart, 18, correspondent à des exclamations d'indignation ou de colère. La forme de la courbe ne correspond pas à celle proposée par Delattre (1966), qui propose pour ce type d'énoncé un arc descendant du niveau 4 au niveau 1, c'est-à-dire aussi une forme inverse à celle des énoncés des groupes 1) et 2) en 3.2.2 et c) en 3.2.3.

Tableau 30 : indices acoustiques de D°71\_I\_94.1%.

N° syl	1		2		3				4		
Transcr <sup>o</sup>	mais		je		m'en				fous		
MO FO Hz	150		152		244.5				465		
May de FO section syll.					192		305		310	538	
Points de FO					153	221.6	288	318	301	434.4	574.5
ton	Ré2+		Ré#2-		Sol2		Ré#3-		Ré#3	Do4+	
ton					Si2 (+4.5)				La#3 (+6.5)		
aug. tons	=				+2.5		+4.5		=	+5.5+	
Tps en ms	73		204		88		90		92	256	
total Tps en ms	277				200				410		
dB	51		53		62.5				60.5		
position									Fin.		

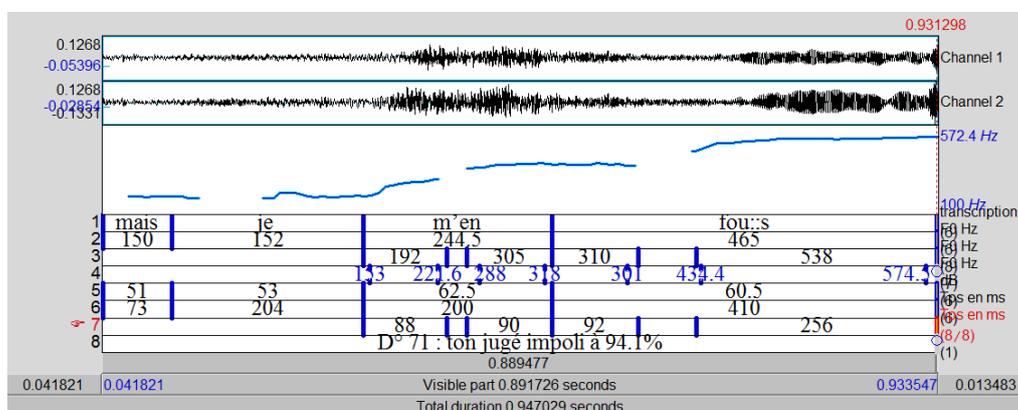


Schéma IV : Capture d'écran des fenêtres d'analyse acoustique de Praat pour la D°71\_I\_91.2% et la D°41\_I\_94.3.

<sup>133</sup> D°15, 17, 19, 20, 23, 29, 34, 37, 53, 57, 58, 60, 67, 68, 71, 74, 89, 99, 105 et 111.

En comparaison, seulement 3 énoncés polis<sup>134</sup> sur 16 présentent un fin haute, mais pour 2 d'entre eux, D°16\_P\_82.4 et D°32\_P\_85.3, la différence est immédiatement visible : alors que pour les D° "impolis" les valeurs en finale sont souvent très élevées comme en D°71, la valeur de F0, au lieu de descendre, s'élève d'1 ton seulement par rapport à la syllabe précédente pour D°32, et reste stable pour D°16. Dans ces 2 cas, l'énoncé reste comme en suspens, comme pour introduire une conclusion tacite que le locuteur laisse à l'interprétation de l'interlocuteur, ce que l'on appelle communément un sous-entendu. D°3\_P\_91.2 laisse davantage perplexe<sup>135</sup> ...

### 3.2.3.3 Conclusion de l'analyse acoustique des énoncés D°

Parmi les 6 phénomènes acoustiques a), b), c), d), e) et f) affectant les valeurs de F0 énumérés en 3.2.3.2, seuls 2 sont le fait des seuls énoncés impolis : i) les courbes présentent (c) un angle aigu, ii) (e) des valeurs distendues; on peut considérer toutefois que c) est une sous catégorie de e). Les autres phénomènes, a), b), d) et f), affectent aussi des énoncés "polis", mais bien moins souvent et avec des valeurs d'indices bien moindres - sauf pour les accents expressifs, identifiables dans la presque totalité des énoncés : n'oublions que l'énoncé dit "neutre" n'existe pas.

**Tableau 31 : Nombre d'énoncés D°I et D°P affectés par les différents types d'indices acoustiques.**

**Ac. expr.** : accents expressifs; **Voy. inst.** : voyelles instables; **Val. dist.** : valeurs distendues; **Fin. H** : finale haute.

	a) plateaux	b) Ac. expr.	c) Angle aigu	d) Voy. inst.	e) Val. dist.	f) Fin. H
<b>D°I /35</b>	<b>12</b>	<b>18</b>	<b>7</b>	<b>12</b>	<b>11</b>	<b>20</b>
<b>D°P /16</b>	3	<b>14</b>	Ø	4	Ø	3

Tx I	75-79.9	80-84.9	85-89.9	90-94.9	95-100
<b>M. ind. ac.</b>	2.9	4	4.75	4.5	5.25

**Tableau 32 : Moyenne du nombre d'indices acoustiques par taux d'impolitesse.**

**Tx I** : taux d'impolitesse; **M. ind. ac.** : moyenne des indices acoustiques par tranche.

<sup>134</sup> D°3, 16 et 32.

<sup>135</sup> Voir en 3.3.3.

On ne peut donc dire qu'un seul type d'indice acoustique assure de l'impolitesse prosodique d'un énoncé, sauf peut-être la catégorie e)<sup>136</sup>. On peut dire en revanche, que :

i) les énoncés "polis" sont proportionnellement beaucoup moins souvent touchés par les indices acoustiques mis en évidence que les énoncés "impolis";

ii) la présence simultanée de plusieurs des phénomènes évoqués précédemment caractérise l'énoncé "impoli", qui en montre toujours au moins 2<sup>137</sup>. On voit en outre que la moyenne du nombre d'indices acoustiques par énoncé augmente avec le taux d'impolitesse;

iii) plus les écarts de valeurs de F0 sont importants et en dehors des positions de CMa, CMi ou finale, plus l'énoncé est perçu "impoli".

### **3.3 Conclusions**

#### **3.3.1 Mise en regard des observations faites sur les corpus Q° et D°**

J'ai regroupé sous différentes étiquettes les caractéristiques de la F0 des énoncés interrogatifs et déclaratifs.

Certaines semblent à première vue être la marque des seuls énoncés interrogatifs comme i) la chute brutale et importante sur la dernière syllabe de l'énoncé, quand la syntaxe voudrait que la F0 monte; mais je proposerais de ranger les énoncés qui présentent ce phénomène dans une catégorie d'énoncés porteurs d'accents expressifs contraires à la syntaxe; ii) des augmentations/diminutions compensatoires successives; mais il est vraisemblable, même si le corpus n'en propose pas, que ce dernier phénomène doit pouvoir affecter aussi des énoncés déclaratifs.

Le phénomène inverse à i) ci dessus, la présence d'une finale haute là où la syntaxe demanderait une baisse de valeur de F0, n'affecte que les énoncés déclaratifs pour marquer la colère ou l'indignation la plupart du temps; je proposerais aussi de classer ces énoncés dans la catégorie des énoncés porteurs d'accents expressifs contraires à la syntaxe.

Les autres caractéristiques se retrouvent *de facto* dans les deux types d'énoncés : i) plusieurs syllabes à la suite ont la même valeur de F0 - plateaux intonatifs; ii) la courbe de F0 présente des écarts entre les valeurs de F0 supérieures aux valeurs standard attendues - valeurs

---

<sup>136</sup> A laquelle je rattache le c).

<sup>137</sup> Voir tableau des phénomènes acoustiques affectant les énoncés impolis en annexe I.

distendues; iii) des variations de F0 interviennent au sein d'une même syllabe - instabilité syllabique.

On peut donc lister comme suit, en tenant compte des précisions formulées précédemment, les marques acoustiques possibles de l'impolitesse :

- 1) accents expressifs contraires à la syntaxe;
- 2) augmentations/diminutions compensatoires successives;
- 3) plateaux intonatifs;
- 4) valeurs de F0 distendues, notamment en dehors des positions syntaxiques habituelles (C<sub>Ma</sub>, C<sub>Mi</sub>, finale);
- 5) instabilité syllabique.

Il convient de souligner ici que la place dans les groupes intonatifs des phénomènes acoustiques que l'on vient d'énumérer joue un rôle de premier ordre dans la perception de l'(im)politesse : avec des valeurs ou une ampleur modérée et s'il(s) occupe(nt) la position syntaxique standard, le(s) phénomène(s) caractérise(nt) un énoncé poli; avec des valeurs ou une ampleur plus grandes et en dehors de la position syntaxique standard, ils caractérisent un énoncé impoli.

### 3.3.2 Pour un usage de la prosodie en analyse interactionnelle

J'ai tenté de regrouper les énoncés du corpus par catégorie selon le phénomène prosodique qui m'a paru le plus saillant, au risque d'occulter momentanément d'autres caractéristiques pour les besoins du propos. Cela n'a pas toujours été facile. Car chaque énoncé mériterait une analyse prosodique particulière et *in extenso* afin de mettre en valeur ses particularités, puisque chacun est finalement unique : les configurations des indices acoustiques entre eux sont extrêmement diverses et contribuent tous à l'analyse pragmatique des énoncés. En voici un exemple, le D°109\_I\_97.1 :

Tableau 33 : indices acoustiques de D°109\_I\_97.1%.

Seg. synt.	A						
	1A	2A	3A	4A	5A	6A	7A
N° syl							
Transcr <sup>o</sup>	mais	c'est	pas	ça	qu'j'	te'd'	mande
MO F0 Hz	133.3	∅	207.5	199	283.5	152.2	138.8
ton	D <sub>o</sub> #2	∅	Sol#2	Sol2	D <sub>o</sub> #3	Ré#2	D <sub>o</sub> #2

aug. tons			<b>+4.5</b>	-0.5	<b>+3</b>	<b>-5.5</b>	-1
Tps en ms	120	100	93	110	123	136	200
total aug. tons			<b>+4.5</b>		<b>+3</b>	<b>-6.5</b>	
total Tps en ms						336	
dB	53.7	54.7	<b>61.6</b>	<b>62.5</b>	60.7	58.8	60.2
position				<b>CMi</b>			<b>Fin.</b>

Seg. synt.	B					
N° syl	1B	2B	3B	4B	5B	6B
Transcr <sup>o</sup>	ch'	d'mande	quel	cé	page	c'est
MO F0 Hz	(135?)	(135?)	167	174	166.4	161.7
	voix féminine					
Points de F0					<b>177.2</b>	<b>154.2</b>
ton					<b>Fa2+</b>	<b>Ré#2</b>
ton	Do#2	Do#2	Mi2+	Fa2	Mi2 (-2)	
aug. tons	=	=	<b>+3.5</b>	<b>+1</b>	<b>-1</b>	<b>-1.5</b>
Tps en ms	122	158	133	<b>164</b>	<b>201</b>	<b>397</b>
total aug. tons			<b>+4.5</b>		<b>-2.5+</b>	
total Tps en ms	122	158	297		598	
dB	59.4	60.8	<b>62.5</b>	<b>62.4</b>	<b>61.8</b>	58.8
position		<b>CMi</b>	<b>CMi</b>			<b>Fin.</b>

Cet énoncé est produit par le locuteur HF2, dont la MO F0 est de 174.53 Hz et la F0 usuelle doit se situer autour de 150 Hz : malgré la perturbation de la voix féminine qui intervient de 1B à 3B, la hauteur de F0 de l'ensemble de l'énoncé semble tout à fait ordinaire.

On note :

- a)
- une voyelle légèrement instable en 5B;
  - quelques valeurs de F0 élevées voire distendues en 3A, 5A, 6A et 3B, dont 3 localisées en dehors des CMa/Mi;
  - les segments A et B sont reliés par un plateau intonatif de 7A à 2B au niveau d'un Do#1, ce qui correspond à la valeur du fondamental usuel du locuteur. Le ton est donc assez bas sur ce segment alors que l'on aurait attendu une montée sur la CMi en 2B;

- le ton aurait dû redescendre en 4B après la CMi, ainsi qu'en 6B en finale; au contraire il monte.

L'énoncé standard se trouve perturbé par les hauteurs de F0 élevées, le non respect des positions habituelles des CMa/Mi, des valeurs de F0 qui les suivent, de la hauteur finale. L'ensemble de ces éléments fonde l'impolitesse acoustique de l'énoncé.

b)

- Si 3B et 4B portent un accent de hauteur, 4B porte aussi un accent de durée, comme 5B et 6B, accent de durée visible par rapport à la rapidité des syllabes précédentes. Le segment "*quel cépage*" est ainsi nettement souligné.

- 4A, 5B et 6B portent en outre un accent d'intensité très net, en écho à 3A et 4A, "*pas ça*".

Les 2 segments sont très énergiquement focalisés par les indices acoustiques, ce qui contribue à construire l'impolitesse de l'énoncé.

Ce type d'analyse vient consolider et enrichir l'analyse interactionnelle traditionnelle, dans le domaine de l'(im)politesse acoustique notamment.

### 3.3.3 Limites

Malgré l'intérêt que cela représente, on est bien obligé d'admettre qu'il reste dans l'analyse acoustique de la voix une zone d'ombre, ou, pour reprendre un terme parfois utilisé dans le champ, une zone sombre, que met en évidence l'analyse de D°3\_P\_91,2. Cet énoncé présente plusieurs caractéristiques partagées par les énoncés "impolis" : les valeurs en sont distendues, une de ses syllabes est instable, et la fin très haute.

Tableau 34 : indices acoustiques de D°3\_P\_91,2%.

N° syl	1			2	3	4	5	6	7
Transcr°	ah			ben	c'est	comme	les	pu	tois
MO F0 Hz	381.2			277.8	229.2	223.3	209.9	332.5	416.2
Points d eF0	322.9	422	443.6						
ton	Mi3	Sol#3+	La3						
aug. ton		+2.5	+0.5						
ton	Sol3			Do#3	La2	La2	Sol#2	Mi3	Sol#3

<b>aug. ton</b>		<b>-4.5</b>	<b>-2</b>	<b>=</b>	<b>-0.5</b>	<b>+4.5</b>	<b>+2.5</b>
<b>Tps en ms</b>	230	74	189	143	160	120	417
<b>total aug. ton</b>		<b>-7</b>				<b>+7</b>	
<b>total Tps en ms</b>	230	566				537	
<b>dB</b>	<i>59.7</i>	<i>58.9</i>	57.2	54.3	52.8	<i>54.9</i>	<i>56.8</i>

La durée des syllabes ne montre rien de notable par rapport à d'autres énoncés jugés "impolis" : la durée de la finale, au dessus des valeurs standard, est un accent expressif que l'on retrouve dans bien des énoncés, "polis" ou "impolis". En ce qui concerne l'intensité, seuls 1 et 2, mots exclamationnels, sont mis en valeur - ils sont les indices acoustiques de la "révélation", réponse à la question vraisemblablement posée par l'interlocuteur - ainsi qu'à un moindre degré 6 et 7, "*putois*", mot porteur de l'essentiel du message.

Qu'est-ce donc qui fait que cet énoncé a été ressenti "poli" par 91.2% des testeurs? Le sens bien entendu a dû jouer un grand rôle, et on ne peut s'empêcher de penser qu'il a souvent influencé la catégorisation. Pourtant, il faut réécouter attentivement l'énoncé et accepter de remplacer le mot "*putois*" par un autre mot de 2 syllabes à valeur axiologique, "crétins" par exemple. Ma perception auditive me laisse entendre une prosodie qui resterait "polie" : autre chose est en jeu, car la voix reste chargée de vibrations émotives positives, relativement joyeuses. Il me semble que cela concerne la pression avec laquelle l'air est expulsé à travers le larynx ainsi que la mise en tension relative des cavités de la phonation et des muscles, notamment faciaux.

Même si le cas est rare puisque l'énoncé D°3 est unique en son espèce dans le corpus, on se doit de convenir qu'à ce jour, l'analyse acoustique seule ne permet pas toujours de discriminer un énoncé poli d'un énoncé impoli.

## ***CONCLUSION(S) ET PERSPECTIVES***

- ◆ **Plusieurs questions ou hypothèses ont été posées en introduction et première partie :**

*1) Comment, dans l'analyse acoustique, faire la part des contraintes accentuelles des énoncés et la part de l'expression des affects?*

J'ai pris en compte, au moment de l'analyse, les variations de tons établies par la tradition pour les énoncés interrogatifs et pour les déclarations en fonction des principales constructions syntaxiques. Ainsi, en position de continuation mineure, on s'attend à une montée de F0 de 0.5 à 1.5 tons; pour une continuation majeure à une montée de 1 à 2.5 tons; pour une finale à une baisse de 0.5 à 1.5 tons. Cela m'a permis de i) mesurer les écarts de F0 des syllabes en position de CMa, CMi ou finale, par rapport à ces valeurs standard attendues; ii) identifier les accents de hauteur "déplacés", c'est-à-dire affectant une syllabe qui ne se trouvait pas dans l'une de ces positions syntaxiques caractéristiques.

*2) L'intonation syntaxique prévaut-elle sur l'intonation affective?*

Etant donné ce qui vient d'être dit, il semble que le locuteur ne puisse s'affranchir totalement de la couche accentuelle syntaxique, présente dans tous les énoncés. Il peut en revanche, selon sa volonté ou sa pulsion du moment, en exagérer plus ou moins les caractéristiques, et créer des perturbations de hauteur en plaçant des pics, creux ou plateaux intonatifs sur et en dehors des positions syntaxiques standard. Si bien que, si l'intonation syntaxique prévaut en ce qu'elle structure - et est structurée par - la couche profonde de l'énoncé, ce n'est pas elle que l'auditeur remarque au premier chef, mais bien l'intonation expressive.

*3) Est-il possible d'identifier des patrons intonatifs typiques de l'(im)politesse?*

Comme on vient de le voir en partie 3, il est en tout cas possible d'identifier des caractéristiques prégnantes d'énoncés perçus comme "(im)polis". Quant à les regrouper par affects spécifiques, la tâche est encore loin d'être accomplie puisqu'il conviendrait tout d'abord de différencier clairement les sentiments/émotions ou états émotionnels (agressivité, colère, fierté etc.) des attitudes, d'identifier clairement les attitudes (accusation, dégoût, doute,

indignation, provocation etc.) par des tests perceptifs, pour pouvoir enfin procéder à leur analyse acoustique. Certains travaux reflètent cette démarche : Banse & Scherer (1996), Bänziger & Scherer (2001), Devillers & Vasilescu (2004), Devillers & Vidrascu (2006), Shochi *et al.* (2008), Lu *et al.* (2013), par exemple.

#### ◆ Perspectives

Etant donné qu'il a été possible d'isoler un certain nombre de phénomènes acoustiques affectant des énoncés perçus comme "(im)polis", nous allons voir maintenant dans quelle mesure l'analyste des conversations ou le didacticien des langues pourra en tirer avantage dans sa réflexion théorique et dans sa pratique.

#### - En analyse conversationnelle ... et en prosodie des affects

L'analyste des interactions d'un côté a tout intérêt à travailler en lien avec des prosodistes car, étant donné ce qui a été dit en 1.1.3 et en 1.1.5, on sait que la prosodie, et principalement l'intonation, véhicule les affects même indépendamment du contenu lexical. Le pragmaticien pourrait ainsi utiliser des travaux qui montreraient des indices acoustiques de tel ou tel affect (im)poli, afin de confirmer, approfondir ou préciser sa propre analyse interactionnelle - dont la prosodie reste la grande absente alors que le matériau d'analyse est principalement oral.

Il est envisageable d'autre part, pour le prosodiste des affects, de poursuivre la route sur laquelle se sont engagés certains de ses confrères, en tournant résolument sa recherche vers la part émotionnelle et affective de l'analyse interactionnelle, notamment vers une "prosodie de l'(im)politesse". Il ne fait pas de doute que le décryptage acoustique de tours de parole relevés en milieu écologique pourrait aider à identifier avec plus de certitude les actes de parole implicites. Plus particulièrement, il serait possible de décrire puis analyser les marques acoustiques de la violence verbale (colère, haine, ironie, menace, mépris, provocation etc.<sup>138</sup>), pour expliquer partiellement, donc participer à déjouer, le phénomène de "montée en tension verbale".

#### - Dans la didactique des langues

---

<sup>138</sup> Pour la distinction perceptive de ces attitudes, il faudrait pouvoir mener de larges tests à vue d'identification, ce qui n'est pas une mince affaire; voir aussi 3) *supra*.

On peut émettre l'hypothèse que des intonations perçues comme "polies" ou acceptables dans une L1 puissent en réception L2 paraître "impolies" ou inadaptées. Dans ce cas, la transposition d'intonation de la L1 à la L2 va poser problème. Même si le transfert d'intonation depuis la L1<sup>139</sup> ne constitue pas la seule source d'erreur pour l'apprenant en L2 (Santiago & Delais-Roussarie 2012), il en est tout de même la source principale, largement attestée<sup>140</sup>. L'existence de "faux amis prosodiques" (Shochi 2008) a par exemple été établie entre le japonais, le français et l'américain : le locuteur de L1 croit reconnaître en L2 une prosodie familière alors qu'il n'en est rien : il s'agit d'une confusion dont l'issue pragmatique peut être à divers degrés gênante ou catastrophique, interculturellement parlant.

D'autre part, une intonation qui ne serait pas exactement transférée de la L1 mais s'avérerait fautive tout de même ne peut que déstabiliser l'auditeur natif, voire nuire à la compréhension du message. Lacheret (2011a : 93) note que les recherches interlangues sur la dimension expressive à l'oral se multiplient actuellement "pour mieux saisir les traits caractéristiques des types de langues, des types de discours et, d'une façon plus générale, la dimension culturelle".

Il serait bien sûr intéressant de multiplier les études comparatives interlangues en fonction de la langue d'origine et de la langue cible. Mais on peut peut-être plus simplement dans un premier temps, dans la pratique de l'enseignement en L2, envisager de centrer l'apprentissage sur les pièges à éviter dans la prosodie, et particulièrement dans l'intonation, de la langue cible.

Ainsi, il faudrait déterminer les phénomènes intonatifs à éviter, tels que, pour le français L2, les accents expressifs contraires à la syntaxe, les augmentations/diminutions compensatoires successives, les plateaux intonatifs, les valeurs de F0 distendues, notamment en dehors des positions syntaxiques habituelles (C<sub>Ma</sub>, C<sub>Mi</sub>, finale), l'instabilité syllabique : toute production d'apprenant qui contiendrait ces phénomènes, perçus comme "impolis" par les natifs, devrait être rapidement corrigée de manière à éviter les malentendus linguistiques ou culturels ainsi que leurs conséquences (préjugés, stéréotypes, mésentente, violence etc.).

#### ◆ Proposition d'activités

Tout d'abord, l'enseignant doit i) expliquer/mettre en évidence/faire percevoir les spécificités de l'intonation en L2 et la présence possible de faux amis prosodiques, comme on

---

<sup>139</sup> On peut émettre l'hypothèse que le locuteur natif de plusieurs langues ne rencontrera pas cette difficulté?

<sup>140</sup> Voir en 1.1.6.

le fait pour les faux amis lexicaux, ii) mettre l'apprenant en garde sur les conséquences interactionnelles des confusions ou des mésusages de l'intonation, iii) mettre en place un travail de remédiation spécifique et éventuellement personnalisé pour déjouer la production de patterns prosodiques erronés, notamment caractéristiques de l'impolitesse.

### **Voici quelques approches de remédiations possibles :**

1) Il va sans dire qu'une bonne acquisition de la prosodie passe par l'écoute et la compréhension orale de séquences de langage authentiques et variées, qui demanderont de travailler avec l'apprenant différentes situations de communication et les diverses stratégies prosodiques des locuteurs (Guimbretière 1997 : 169), et ce dès le niveau débutant du CECRL<sup>141</sup>, car "l'oralité ne se découpe pas en strates de complexité croissante" (Lhote *et al.* 2000 : 23). Il faut, nous dit Lhote (*ibid.*), dans l'optique d'une pédagogie communicationnelle, entraîner l'apprenant à "ajuster dans l'instant sa façon d'écouter et sa façon de parler à la situation et à son interlocuteur", en s'appuyant sur les compétences interactionnelles déjà développées en L1 et dont l'apprenant doit prendre conscience. Dans ce contexte là, l'enseignant fera travailler la perception de l'intonation en même temps que la compréhension orale (pré-écoute, écoute globale, écoute sélective), en demandant aux apprenants de dégager les caractéristiques prosodiques des énoncés.

La perception et l'acquisition d'intonations (im)polies trouvent tout à fait leur place dans ces exercices : on sélectionnera préférentiellement - en prenant soin d'en adapter la difficulté au niveau de l'apprenant - des extraits de films, de journaux ou entretiens télévisés, de publicités ou de pièces de théâtre qui montrent des échanges à divers degrés conflictuels<sup>142</sup>, sur l'intonation desquels on fera travailler plus spécifiquement les apprenants comme indiqué précédemment.

2) Ce travail peut être consolidé par des exercices systématiques de discrimination intonative : Chun (2002 : 92) cite en référence une étude qui a montré que l'apprenant progresse dans l'expression prosodique surtout grâce à l'amélioration de sa perception de l'intonation native, plutôt que par l'écoute de sa propre production qui ne semble guère aider l'apprenant à progresser (Bot & Mailfert, cités par Chun 2002 : 92).

Dans le cas de l'(im)politesse, il faudra créer des énoncés adéquats pour constituer les exercices audio.

---

<sup>141</sup> Cadre européen commun de référence pour les langues.

<sup>142</sup> Plutôt que des extraits d'émissions radiophoniques car l'image aide à clarifier la situation de communication.

3) Certains chercheurs proposent d'utiliser des nouveaux logiciels d'analyse acoustique, comme WinPitch<sup>143</sup> Language Teaching and Learning (LPL), pour aider l'apprenant à corriger lui-même ses erreurs prosodiques. WinPitch en effet a été conçu dans une optique pédagogique; il permet d'analyser et de visualiser en temps réel la production orale de l'apprenant, que ce dernier peut comparer autant de fois qu'il le souhaite avec une courbe et une production modèles (Léon & Martin 2000 : 145-146). Lors de ces exercices, on fera travailler à l'apprenant aussi bien les intonations polies qu'impolies, en mettant bien en évidence les caractéristiques respectives. On peut envisager de faire traduire par l'apprenant ces énoncés dans sa L1, de les lui faire enregistrer en tâchant de reproduire dans sa langue l'affect mobilisé en français L2, pour finalement en comparer visuellement leurs courbes intonatives respectives. L'exercice va bien plus loin qu'une simple imitation puisque i) l'apprenant peut voir jusqu'à quel point sa propre courbe intonative diverge ou se rapproche du modèle, ii) l'exercice n'est plus seulement auditif, il est surtout visuel : par tâtonnements et améliorations successifs, l'apprenant va pouvoir inscrire dans sa mémoire kinesthésique la meilleure position des organes de la parole pour la meilleure production possible<sup>144</sup>.

4) Car il faut bien sûr travailler la production orale des patterns (im)polis spécifiquement. Pour ce faire, une situation de classe banale risque - heureusement - de ne pas fournir les opportunités nécessaires; on a donc tout intérêt à créer un cadre communicatif dans lequel de tels affects peuvent apparaître sans susciter de tensions ou de gêne : dans une situation de communication fictive mais vraisemblable, celle du jeu théâtral.

Des activités théâtrales variées permettent de mettre en jeu la voix **et** le corps, ce qui est très important dans le cas des affects car comme on l'a vu ils s'enracinent dans le corps<sup>145</sup>. On travaillera plus facilement dans ce cadre, sans compter la réactivation ou l'acquisition de lexique ou de structures syntaxiques, tout l'environnement prosodique de la L2 : prononciation/articulation, groupes de souffle, rythme, accents, expressifs ou non.

Régine Llorca<sup>146</sup>, par exemple, propose de "lier le geste à la parole", afin d'inscrire le rythme intonatif dans la mémoire kinesthésique, par une sorte de mime intonatif. Le travail

---

<sup>143</sup> WinPitch est un logiciel gratuit créé par Philippe Martin en 1996. Il est téléchargeable à l'adresse : [http://www.winpitch.com/winpitch\\_tltl.htm](http://www.winpitch.com/winpitch_tltl.htm)

<sup>144</sup> On peut aussi bien entendu penser à l'apprentissage de la production orale chez les personnes nées sourdes.

<sup>145</sup> Voir en 1.1.2.1.

<sup>146</sup> Voir par exemple : Jeux de sons et Variations sur l'intonation, magazine *Schulfunk*, n° 9, 44-45, et n° 11 / 12, 31-32; Jeux de groupe avec la voix et le geste sur les rythmes du français parlé, *Actes du colloque «Quelles pédagogies pour l'enseignement des langues aux adultes aujourd'hui?»*, sept. 2001, Publications de l'Université de Saint-Etienne, 141-150; en collaboration avec Elisabeth LHOTE : Le geste, outil d'écoute. *Le Français dans le Monde*, numéro spécial janv. 2001, 160-164.

pourra aussi se faire à partir de mots d'une puis plusieurs syllabes - pour commencer des interjections et des interpellations - jusqu'à l'énoncé complet, associés à des gestes significatifs en L2 (gestes culturels), des postures corporelles, des déplacements. Sur ces brefs supports qui ne doivent pas mobiliser outre mesure sa mémoire, l'apprenant-acteur peut faire toutes sortes de gammes prosodiques d'expression affective (im)polie. Dans ces moments là particulièrement, l'enseignant peut proposer d'alourdir le jeu et d'accentuer les caractéristiques prosodiques. On peut ensuite développer le texte autant que l'on veut, texte littéraire, tiré d'une méthode de langue, texte écrit voire improvisé par les apprenants, que l'on fait jouer sous forme de sketches ou de saynètes, et avec lesquels les apprenants joueront pour mieux endosser la dimension théâtrale qu'implique toute production en L2, puisque la prosodie "naturelle", celle de la L1, est délaissée pour celle, d'abord artificielle, de la L2. Pris par le *ludus*<sup>147</sup>, l'apprenant s'autorise plus facilement sous le masque de l'*hypokritès*<sup>148</sup> à restituer en L2 des affects qu'il doit aussi savoir reconnaître et produire, ou non, selon la situation de communication.

A l'issue de cette réflexion, il semble donc envisageable de réduire les malentendus communicationnels et /ou interculturels par l'apprentissage en L2 d'intonations (im)polies, pourvu que l'on permette aux apprenants de prendre conscience du phénomène et qu'on leur propose des activités d'écoute et de production adaptées.

---

<sup>147</sup> Le mot *ludus* (lat.) est en général traduit par "jeu" ou "école". Il désigne en fait une activité qui consiste à "faire semblant", à faire les choses "pour de faux" avant de les faire "pour de vrai", dans le monde réel.

<sup>148</sup> Acteur du théâtre grec antique, par contraste avec le ludion, acteur du théâtre romain.

## **BIBLIOGRAPHIE**

ABRY, D. & VELDEMAN-ABRY, J. (2007). *La Phonétique : audition, prononciation, correction*. Paris : CLE International.

ABRY, D. & CHALARON, M.L. (2010). *Les 500 exercices de phonétique, niveau A1/A2*. Paris : Hachette FLE.

ABRY, D. & CHALARON, M.L. (2011). *Les 500 exercices de phonétique, niveau B1/B2*. Paris : Hachette FLE.

ARNOLD, A. (2012). Le rôle de la fréquence fondamentale et des fréquences de résonance dans la perception du genre. *TIPA - Travaux Interdisciplinaires sur la Parole et le Langage*, 28, 1-18.

AUBERGE, V. (2002). Prosodie et émotion. Dans *Actes des deuxièmes assises nationales du GDR I*, 263-273. Consulté le 07/06/2013 à l'adresse : <http://www.irit.fr/GDR-I3/fichiers/assises2002/papers/15-ProsodieEtEmotion.pdf>

AUDIBERT, N., AUBERGE, V. & RILLIARD, A. (2004). EWiz : contrôle d'émotions authentiques. Dans *Actes des XXV<sup>e</sup> Journées d'Étude sur la Parole*. Consulté le 29/05/2013 à l'adresse : [http://www.afcp-parole.org/doc/Archives\\_JEP/2004\\_XXVe\\_JEP\\_Fes/actes/jep.htm](http://www.afcp-parole.org/doc/Archives_JEP/2004_XXVe_JEP_Fes/actes/jep.htm)

AUDIBERT, N., AUBERGE, V. & RILLIARD, A. (2008). Emotions actées vs. spontanées : variabilité des compétences perceptives. Dans *Actes des XXVII<sup>e</sup> Journées d'Étude sur la Parole*. Consulté le 04/06/2013 à l'adresse : [http://www.afcp-parole.org/doc/Archives\\_JEP/2008\\_XXVIIe\\_JEP\\_Avignon/PDF/avignon2008\\_pdf/JEP/063\\_jep\\_1660.pdf](http://www.afcp-parole.org/doc/Archives_JEP/2008_XXVIIe_JEP_Avignon/PDF/avignon2008_pdf/JEP/063_jep_1660.pdf)

BACHOROWSKI, J.A. (1999). Vocal Expression and Perception of Emotion, *American Psychological Society*, 8 (2), 53-57. Consulté le 04/06/2013 à l'adresse : <http://www.utdallas.edu/~assmann/hcs6367/bachorowski99.pdf>

BANSE, R. & SCHERER, K.R. (1996). Acoustic Profiles in Vocal Emotion Expression, *Journal of Personality and Social Psychology*, 70 (3), 614-636.

BÄNZIGER, T. & SCHERER, K.R. (2001). Relations entre caractéristiques vocales perçues et émotions attribuées. Dans Aubergé, V., Lacheret-Dujour, A. & Loevenbruck, H. (Eds), (2003). *Journées Prosodies*, 119-124. Grenoble.

BÄNZIGER, T., GRANDJEAN, D., BERNARD, P.J., KLASMEYER, G. & SCHERER, K.R. (2001). Prosodie de l'émotion : étude de l'encodage et du décodage. *Cahiers de linguistique française*, 23, 11-37.

BERTRAND, R. (2003). Etre soi avec les mots d'autrui. Dans MERLE, J.M. (Ed.) (2003). *Le Sujet*, 269-280. Paris : Ophrys.

BERTRAND, R. & ESPESSER, R. (2002). Voice Diversity in Conversation : a Case Study. *Speech Prosody Aix-en-Provence France 2002*. Aix-en-Provence : Presses Universitaires de Provence.

BERTRAND, R., PORTES, C. & SABIO, F. (2007). Distribution syntaxique, discursive et interactionnelle des contours intonatifs du français dans un corpus de conversation. *Travaux neuchâtelois de linguistique*, 47, 59-77.

BERTRAND, R. & PORTES, C. (2012). Pour une approche phonologique de la prosodie dans l'interaction. *Langue française* 175, 19-36.

BOULA DE MAREÛIL, P., RILLIARD, A., MAIRANO, P. & LAIJ.P. (2012). Questions corse : peut-on mettre en évidence un transfert prosodique du corse vers le français? Dans *Actes de la conférence conjointe JEP-TALN-RECITAL 2012 Grenoble, 1*, 609-616. Consulté le 09/06/2013 à l'adresse :

<http://aclweb.org/anthology//F/F12/F12-1077.pdf>

BOUSFIELD, D. (2007a). Beginnings, Middles and Ends : Towards a biopsy of the dynamics of impoliteness. *Journal of Pragmatics*, 39, (12), 2185-2216.

BOUSFIELD, D. (2007b). Impoliteness, preference organization and conductivity. *Multilingua*, 26, (1/2), 1-33.

BOUSFIELD, D. & LOCHER, M. (2008). *Impoliteness in language*. D. Bousfield et M. Locher (Eds), Berlin : Mouton de Gruyter.

BOUSFIELD, D. (2008). *Impoliteness in Interaction*, Amsterdam/Philadelphie : John Benjamins Publishing Company.

BOUSFIELD, D. (2010). Researching impoliteness and rudeness : Issues and definitions. Dans M. Locher et S. Lambert (Eds), *Interactional Pragmatics*, 101-134. Berlin : Mouton de Gruyter.

BRAZIL, D., COULTHARD, M. & JOHNS, C. (1980). *Discourse Intonation and Language Teaching*. Harlow (Royaume Uni) : Longman.

BROWN, P. & LEVINSON, S. (1978). Universals in language usage : Politeness phenomena. Dans E. N. Goody (Ed.), *Questions and politeness: Strategies in social interaction* (56-289). New York : Cambridge University Press.

BROWN, P. & LEVINSON, S. (1987). *Politeness : some universals in language usage*. Cambridge : Cambridge University Press.

CAELEN-HAUMONT, G. (2002), Prosodie et dialogue spontané : valeurs et fonctions perlocutoires du mélisme. *Travaux Interdisciplinaires du Laboratoire Parole et Langage*, 21, 13-24.

CAELEN-HAUMONT, G. (2008a). *Prosodie et sens : une approche expérimentale*, volume 1. Paris : L'Harmattan, Marges linguistiques.

CAELEN-HAUMONT, G. (2008b). *Prosodie et sens : une approche expérimentale*, volume 2. Paris : L'Harmattan, Marges linguistiques.

CHUN D. (2002). *Discourse Intonation in L2 : From Theory and Research to Practice*. Amsterdam : John Benjamins. Consulté en ligne (extraits) à l'adresse : <https://benjamins.com/#catalog/books/llt.1/preview>

COUPER-KUHLEN, E. (1999). Coherent Voicing : On Prosody in Conversational Reported Speech. Dans BUBLITZ, W., LENK, U. & VENTOLA, E. (Eds) *Coherence in Spoken and Written Discourse* (11-34). Amsterdam/Philadelphia : John Benjamins Publishing Company.

COUPER-KUHLEN, E. & SELTING, M. (1996). *Prosody in Conversation. Interactional studies*. Cambridge : Cambridge University Press.

CULPERER, J. (1996). Towards an anatomy of impoliteness. *Journal of Pragmatics*, 25, 349-367.

CULPERER, J. (2010). Conventionalised impoliteness formulae. *Journal of Pragmatics*, 42, 3232-3245. doi : 10.1016/j.pragma.2010.05.007

CULPERER, J., BOUSFIELD, D. & WICHMANN, A. (2003). Impoliteness revisited : With special reference to dynamic and prosodic aspects. *Journal of Pragmatics*, 35, (10/11), 1545-1579.

DELAIS-ROUSSARIE, E. & DURAND, J. (Eds) (2003). *Corpus et variation en phonologie du français : méthodes et analyses*. Toulouse : Presses Universitaires du Mirail.

DAMASIO, A.R. (1994). *Descartes' error. Emotion, reason, and the human brain*. New York : Putnam Books.

DE ABREU, S., MATHON, C. & MOSCA, A. (2009). Prosodic Cues of Anger in French : A Comparative Study with Two Other Romance Languages, Italian and Portuguese. Dans HANCIL, S. (Ed.), *The Role of Prosody in Affective Speech*, 61-93. Bern : Peter Lang.

DELATTRE, P. (1966). Les Dix Intonations de base du français. *The French Review*, 40, (1), 1-14.

DELATTRE, P. (1967). La Nuance de sens par l'intonation. *The French Review*, 41 (3), 326-339. Consulté le 10/01/2013 à l'adresse :  
<http://www.jstor.org/stable/385163>

DEVILLERS, L. & VASILESCU, I. (2004). Détection des émotions à partir d'indices lexicaux, dialogiques et prosodiques dans le dialogue oral. Dans *Actes des XXV<sup>e</sup> Journées d'Étude sur la Parole*. Consulté le 29/05/2013 à l'adresse :  
[http://www.afcp-parole.org/doc/Archives\\_JEP/2004\\_XXVe\\_JEP\\_Fes/actes/jep2004/Devilleers-Vasilescu.pdf](http://www.afcp-parole.org/doc/Archives_JEP/2004_XXVe_JEP_Fes/actes/jep2004/Devilleers-Vasilescu.pdf)

DEVILLERS, L. & VIDRASCU, L. (2006). Real-life emotions detection with lexical and paralinguistic cues on Human-Human call center dialogs. 9<sup>e</sup> ICSLP, Interspeech 17-21 septembre 2006. Consulté le 29/05/2013 à l'adresse :  
<http://www.cs.columbia.edu/~julia/papers/devillers%26vidrascu06.pdf>

DI CRISTO, A. (1998). Intonation in French. Dans D. Hirst (dir.), *Intonation System : a Survey of Twenty Languages*, 195-218. Cambridge : Cambridge University Press.

DODANE, C. & GUILLEMINOT, C. (2002). Détection de la stabilité de timbre des voyelles : 24<sup>ème</sup> Journées d'étude sur la parole, Nancy, 24-25 juin, 141-144. Consulté le 28/07/2014 à l'adresse :  
[http://www.afcp-parole.org/doc/Archives\\_JEP/2002\\_XXIVe\\_JEP\\_Nancy/JEP2002/papiers/101.pdf](http://www.afcp-parole.org/doc/Archives_JEP/2002_XXIVe_JEP_Nancy/JEP2002/papiers/101.pdf)

ELLSWORTH, P.C. & SCHERER, K.R (2003). Appraisal Processes in Emotion. Dans DAVIDSON, R. J., SCHERER, K.R. & GOLDSMITH, H. H. (Eds), *Handbook of Affective Sciences*, 411-414. Oxford : Oxford University Press.

FAGYAL, Z. (2003). La prosodie du français populaire des jeunes: traits héréditaires et novateurs. *Le Français aujourd'hui*, 143, n<sup>o</sup> spécial "Français de l'école et langues des élèves : quel statut, quelles pratiques ?", 47-55.

FÓNAGY, I. (1971). La synthèse de l'ironie. *Phonetica*, 23, 42-51.

FÓNAGY, I. (1976). La mimique buccale. Aspect radiologique de la vive voix. *Phonetica* 33, 31-44.

FÓNAGY, I. (1978). A New Method of Investigating the Perception of Prosodic Features. *Language and Speech*, 21/1, 34-49.

FÓNAGY, I. (1981). Emotions, voice and music. Dans J. Sundberg (Ed.), *Research aspects on singing*, 137-152. Stockholm : The Royal Swedish Academy of Music.

FÓNAGY, I. (1983). *La Vive Voix*. Paris : Payot.

FÓNAGY, I. (2003). Des fonctions de l'intonation : essai de synthèse. *Flambeau*, 29, 1-20. Consulté le 14/04/2013 à l'adresse :

<http://ed268.univparis3.fr/lpp/pages/EQUIPE/vaissiere/fonagy/articles/fontions%20de%20l'intonation.pdf>

FÓNAGY, I., BERARD, E. & FÓNAGY, J. (1983). Clichés mélodiques, *Folia Linguistica* 17 (1-4), 153-186. DOI : 10.1515/flin.1983.17.1-4.153, July.

GOFFMAN, E. (1973). *La mise en scène de la vie quotidienne. 1. La présentation de soi*. Paris : Minuit.

GOFFMAN, E. (1973). *La mise en scène de la vie quotidienne. 2. Les relations en public*. Paris : Minuit.

GRICE, H. P. (1975). Logic and conversation. Dans P. Cole et J. Morgan (Eds), *Speech Acts*, 41-58. New York : Academic Press.

GRICHKOVTSOVA, I., LACHERET, A. & MOREL, M. (2007). The role of intonation and voice quality in the affective speech perception. *Interspeech*, 2245-2248, conférence du 30 août 2007. Antwerp, Belgique : ISCA. Consulté le 04/06/2013 à l'adresse : <http://lacheret.com/Xinha/UPLOAD/39.Emotions-Interpseech%202007.pdf>

GUIMBRETIERE, E. (1997). Analyse de faits prosodiques dans une perspective d'enseignement/apprentissage. Dans PERROT J. (Coord.) *Polyphonie pour Ivan Fónagy : Mélanges offerts à Ivan Fónagy par un groupe de disciples, collègues et admirateurs*, 163-171. Paris : L'Harmattan.

GUIMBRETIERE, E. (Dir.) (2000a). *Apprendre, enseigner, acquérir : la prosodie au cœur du débat*. Rouen : Presses Universitaires de Rouen-CNRS.

GUIMBRETIERE, E. (2000b). Entretien avec Ivan Fónagy. Dans GUIMBRETIERE, E. (Dir.) (2000). *Apprendre, enseigner, acquérir : la prosodie au cœur du débat*, . Rouen : Presses Universitaires de Rouen-CNRS.

GUMPERZ, J. (1989a). *Engager la conversation. Introduction à la sociolinguistique interactionnelle*. Paris : Editions de Minuit.

GUMPERZ, J. (1989b). *Sociolinguistique interactionnelle : une approche interprétative*. Paris : L'Harmattan.

GUZMAN, M. (1973). *Les Interférences de l'intonation interrogative dans le français parlé par des Chiliens*. (Thèse de 3<sup>o</sup> cycle, Université Sorbonne Nouvelle, Institut de Phonation et Langage, Paris).

HIRST, D. (dir.), (1998). *Intonation System : a Survey of Twenty Languages*. Cambridge : Cambridge University Press.

ITO, M. (2002). *Japanese Politeness and Suprasegmentals - a Study based on Natural Speech Materials*. Department of Theoretical and Applied Linguistics, University of Edimburgh, United Kingdom. Consulté le 11/04/2013 à l'adresse : <http://www.sk.uni-bonn.de/lehre/informationen-materialien/informationen-und-materialien-kopho/materialien-1/wagner/stimmqualitaten/ito-m.pdf>

JAKOBSON, R. (1963). *Essais de linguistique générale*. Paris : Éditions de Minuit.

KANEMAN-POUGATCH, M. & PEDOYA-GUIMBRETIERE, E. (1991). *Au Plaisir des sons*. Paris : Hatier/Didier.

KELTNER, D. & EKMAN, P. (2003). Introduction : expression of emotion. Dans DAVIDSON, R. J., SCHERER, K.R. & GOLDSMITH, H. H. (Eds), *Handbook of Affective Sciences*, 411-414. Oxford : Oxford University Press.

KERBRAT-ORECCHIONI, C. (2002). Politesse en deçà des Pyrénées, impolitesse au-delà : retour sur la question de l'universalité de la (théorie de la) politesse. *Marges linguistiques*, 3, 1-18. Saint-Chamas, France : M.L.M.S. Eds. Consulté le 10/03/2013 à l'adresse : <http://icar.univ-lyon2.fr/Equipe1/documents/MargesKerbrat.pdf>

KERBRAT-ORECCHIONI, C. (2005). *Le discours en interaction*. Paris : Armand Colin.

KERBRAT-ORECCHIONI, C. (2007). L'Analyse du discours en interaction : quelques principes méthodologiques. Université Lumière, Lyon. Consulté le 11/04/2013 à l'adresse : [http://www.fflch.usp.br/dlcv/enil/pdf/Artigo\\_Catherine\\_Kerbrat\\_Orecchioni.pdf](http://www.fflch.usp.br/dlcv/enil/pdf/Artigo_Catherine_Kerbrat_Orecchioni.pdf)

KERBRAT-ORECCHIONI, C. (2010). L'impolitesse en interaction : aperçus théoriques et étude de cas. *Lexis Special 2 "Impoliteness / Impolitesse"*. D. Jamet & M. Jobert (Eds). Consulté le 19/02/2013 à l'adresse : [http://lexis.univ-lyon3.fr/IMG/pdf/Lexis\\_special\\_2.pdf](http://lexis.univ-lyon3.fr/IMG/pdf/Lexis_special_2.pdf)

KLEWITZ, G. & COUPER-KUHLEN, E. (1999). Quote - Unquote? The Role of Prosody in the Contextualisation of Reported Speech Sequences. *Pragmatics*, 9 (4), 459-485.

LACHERET-DUJOUR, A. (2011a). La prosodie au cœur du verbal. *Rééducation orthophonique*, 246, 87-104. Consulté le 04/06/2013 à l'adresse : <http://www.lacheret.com/Xinha/UPLOAD/4.Revue-ortho.pdf>

LACHERET-DUJOUR, A. (2011b). Le corps en voix ou l'expression prosodique des émotions. *Evolutions psychomotrices*, 23, 90. Consulté le 30/05/2013 à l'adresse : <http://www.lacheret.com/Xinha/UPLOAD/Lacheret-corps-en-voix.pdf>

LACHERET-DUJOUR, A. & BEAUGENDRE, F. (1999). *La prosodie du français*. Paris : CNRS Editions.

LACHERET-DUJOUR, A., MARTEL, K. & VIVIER, J. (2000). Prosodie et acquisition du français langue maternelle : l'émergence de la prosodie dialogique chez les jeunes enfants. Dans GUIMBRETIERE, E. (Dir.) (2000). *Apprendre, enseigner, acquérir : la prosodie au cœur du débat*, 205-226. Rouen : Presses Universitaires de Rouen-CNRS.

LACHERET-DUJOUR, A. & FRANÇOIS, J. (2004). Circonstances et prédication en français parlé : contraintes sémantico-pragmatiques et filtrage prosodique. *Revue de syntaxe et sémantique : la prédication verbale et averbale dans les langues*, 6, 35-56. Lacheret-Dujour, C. & Behr, I. (Eds). Caen, PUC. Consulté le 18/04/2013 à l'adresse : <http://www.lacheret.com/Xinha/UPLOAD/28.Predication-LAC-FRA.pdf>

LACHERET-DUJOUR, A. & MOREL, M. (2011). Modéliser la prosodie pour la synthèse à partir du texte : Perspectives sémantico-pragmatiques. Dans NEVEU, F., BLUMENTAL, P. & LEQUERLER, N. (Eds), *Au commencement était le verbe : syntaxe, sémantique et cognition, Mélanges en l'honneur du Professeur Jacques François*, 299-325. Peter Lang. Consulté le 08/06/2013 à l'adresse : <http://lacheret.com/publications.php>

LACHERET-DUJOUR, A., KAHANE, S., PIETRANDREA, P., AVANZI, M. & VICTORRI, B. (2011). Oui mais elle est où la coupure, là ? Quand syntaxe et prosodie s'entraident ou se complètent. *Langue française*, 170, 61-80. Paris : Larousse. Consulté le 15/04/2013 à l'adresse : <http://www.lacheret.com/Xinha/UPLOAD/3.Langue-francaise.pdf>

LAKOFF, R. (1972). Language in Context. Dans *Language*, 48/4, 907-927, Linguistic Society of America. Consulté le 27/05/2013 à l'adresse : [https://5361invention.pbworks.com/f/Lakoff\\_72\\_fletch.pdf](https://5361invention.pbworks.com/f/Lakoff_72_fletch.pdf)

LAKOFF, R. (1973). The logic of politeness; or, minding your p's and q's. Dans *Papers from the Ninth Regional Meeting of the Chicago Linguistics Society*, 292-305. Chicago Linguistics Society. Consulté le 27/05/2013 à l'adresse : [http://www.linguistik.uni-erlangen.de/clue/fileadmin/docs/rrh/papers/presuppositions\\_and\\_quantifiers.pdf](http://www.linguistik.uni-erlangen.de/clue/fileadmin/docs/rrh/papers/presuppositions_and_quantifiers.pdf)

LEBLANC, G. (2012). *Une étude acoustique des voyelles orales susceptibles d'être diphtonguées en français québécois*. (Mémoire, Université de Laval, Département de Langues, Linguistique et Traduction, Faculté des Lettres, Laval, Québec). Consulté en ligne le 27/07/2014 à l'adresse : <file:///D:/Downloads/28362.pdf>

LEHKA-LEMARCHAND, I. (2007). *Accent de banlieue : approche phonétique et sociolinguistique de la prosodie des jeunes d'une banlieue rouennaise*. (Thèse de doctorat, Université de Rouen, département des Sciences du langage et de la Communication, Rouen). Consulté en ligne le 16/06/2013 à l'adresse : [http://www.afcp-parole.org/doc/theses/these\\_ILL07.pdf](http://www.afcp-parole.org/doc/theses/these_ILL07.pdf)

LEON, M. (1965). *Le rythme et l'intonation françaises*. Paris : Hachette-Larousse - BELC.

LEON, P. (1993). *Précis de phonostylistique, parole et expressivité*. Paris : Nathan Université.

LEON P. & LEON M. (1997). *Prononciation du français*. Paris : Nathan Université (Rééd.).

LEON, P. & MARTIN, P. (2000). Prosodie et technologie. Dans GUIMBRETERE, E. (Dir.), *Apprendre, enseigner, acquérir : la prosodie au cœur du débat*, 135-149. Rouen : Presses Universitaires de Rouen CNRS.

LHOTE, E., BARRY, A.-O. & TIVANE, A. (2000). Acquisition et apprentissage de la prosodie : une double approche vocale et discursive. Dans GUIMBRETERE, E. (Dir.), *Apprendre, enseigner, acquérir : la prosodie au cœur du débat*, 21-45. Rouen : Presses Universitaires de Rouen CNRS.

LU, Y., AUBERGE, V., RILLIARD, A. & GU, W. (2012). Prosodic Cross-linguistic Perception of Social Affects in Mandarin Chinese by Native, French and Vietnamese Listeners. Dans MELLO, H., PETTORINO, M. & RASO, T. (Eds) *Proceedings of the VII<sup>th</sup> GSCP International Conference : Speech and Corpora*. Florence : Florence University Press.

MERTENS, P. (2008). Syntaxe, prosodie et structure informationnelle : une approche prédictive pour l'analyse de l'intonation dans le discours. *Travaux de linguistique*, 56/1, 97-124. DOI : 10.3917/tl.056.0097

MOÏSE, C. (2007). Contexte et violence verbale. *La mise en œuvre des langues dans l'interaction, 4ème colloque du réseau de sociolinguistique*, 79-101. Paris : L'Harmattan.

MOÏSE, C., AUGER, N., FRACCHIOLA, B. & SCHULTZ-ROMAIN, C. (Eds) (2008). *De l'impolitesse à la violence verbale* (2 tomes). Paris : L'Harmattan.

MOREL, M.A. & DANON-BOILEAU, L. (1998). *Grammaire de l'intonation : l'exemple du français*. Paris : Ophrys.

MOREL, M. & BÄNZIGER, T. (2004). Le rôle de l'intonation dans la communication vocale des émotions : test par la synthèse. *Cahiers de l'Institut de Linguistique de Louvain*, 30 (1-3), 207-232. Consulté en ligne le 08/06/2013 à l'adresse : [http://hal.archives-ouvertes.fr/docs/00/10/03/47/PDF/article\\_proso\\_Geneve\\_CIL\\_Louvain.pdf](http://hal.archives-ouvertes.fr/docs/00/10/03/47/PDF/article_proso_Geneve_CIL_Louvain.pdf)

MOUGIN, S. (Dir.) (2006). *La Médisance*. Reims : Presses Universitaires de Reims.

NOVAKOVA, I. & TUTIN, A. (2009). *Le lexique des émotions*. Grenoble : ELLUG.

OPREA, A.G. (2012). *Le système de la politesse confronté aux défis du "talk-show" : Politesse, impolitesse et a-politesse à l'épreuve du spectacle et de la violence dans "On n'est pas couché" et "Tout le monde en parle"*. (Thèse de doctorat, Université de Babeş-Bolyai, Cluj-Napoca, Université Lumière Lyon 2, Ecole doctorale 3LA (Lettres, Langues, Linguistique et Arts), Cluj-Napoca, Roumanie).

PEPIOT, E. (2013). *Voix de femmes, voix d'hommes : à propos de l'identification du genre par la voix chez des auditeurs anglophones et francophones*. (Thèse de doctorat, Université de Paris 8, Ecole doctorale 31 Pratiques et théories du sens, Sciences du langage Spécialité Phonétique / Linguistique anglaise).

PRIEGO-VALVERDE, B. (2001). "C'est du lard ou du cochon?" : lorsque l'humour opacifie la conversation familière. *Marges linguistiques*, 2, 195-208. Saint-Chamas, France : M.L.M.S. Eds. Consulté le 10/03/2013 à l'adresse : <http://aune.lpl.univ-aix.fr/~fulltext/1266.pdf>

QUNTILIEN (90 env.). *Institution oratoire*, Paris : Les Belles Lettres.

RENARD, R. (1971). *Introduction à la méthode verbo-tonale de correction phonétique*. Paris : Didier.

RICHARD D., CAMPS J.-F. , EUGENE D., GAUTHIER M., *et al.* (2013). *Neurosciences - Tout le cours en fiches*. Paris : Dunod.

RITTAUD-HUTINET, C. (1995). *La phonopragmatique*. Berne, Suisse : Peter Lang.

RITTAUD-HUTINET, C. (2005). Les signes vocaux : au carrefour des strates d'interprétation des messages oraux. *Actes du colloque international Interface discours - prosodie d'Aix-en-Provence (septembre 2005)*. Edition sur CDRom. Consulté le 12/04/2013 à l'adresse : <http://aune.lpl.univ-aix.fr/~prodige/idp05/actes/rittaud-hutinet.pdf>

ROSSI, M. (1999). *L'intonation, le système du français*. Paris : Ophrys.

SANTIAGO, F. & DELAIS-ROUSSARIE, E. (2012). La Prosodie des énoncés interrogatifs en français L2. *Actes de la Conférence conjointe JEP-TALN-RECITAL 2012*, 4 au 8 juin 2012 Grenoble, 1, 265-272.

SCHERER, K.R. (Ed.), (1988). *Facets of Emotions : Recent Research*. Hillsdale, New Jersey.

SCHERER, K.R., JOHNSTONE, T. & SANGSUE, J. (1998). L'état émotionnel du locuteur : facteur négligé mais non négligeable pour la technologie de la parole. Carouge, Suisse : FPSE Université de Genève. Consulté le 14/03/2013 à l'adresse : <http://psyphz.psych.wisc.edu/~johnstone/jep81.pdf>

SCHERER, K.R., BANSE, R. & WALLBOTT, H.G. (2001). Emotions Interferences from Vocal Expression Correlate across Languages and Cultures. *Journal of cross-cultural Psychology*, 32 (1), 76-92. Washington : Western Washington University.

SCHERER, K.R., JOHNSTONE, T. & KLASMEYER, G. (2003). Vocal Expression of Emotion. Dans DAVIDSON, R. J., SCHERER, K.R. & GOLDSMITH, H. H. (Eds), *Handbook of Affective Sciences*, 433-456. Oxford : Oxford University Press.

SEARLE, J. (1969). *Speech Acts*. Cambridge : Cambridge University Press.

SHOCHI, T., AUBERGE, V. & RILLIARD, A. (2006). How prosodic attitudes can be false friends : Japanese vs. French social affects. Consulté le 23/07/2014 à l'adresse : [http://20.210-193-52.unknown.qala.com.sg/archive/sp2006/papers/sp06\\_249.pdf](http://20.210-193-52.unknown.qala.com.sg/archive/sp2006/papers/sp06_249.pdf)

SHOCHI, T., RILLIARD, A. & AUBERGE, V. (2009). Intercultural Perception of English, French and Japanese Social Affective Prosody. Dans HANCIL, S. (Ed.), *The Role of Prosody in Affective Speech*, 31-60. Bern : Peter Lang.

TORCK, D. (2006). Médisance? Paroles rapportées / représentées et construction d'une image négative de l'Autre dans l'écriture journalistique. Dans MOUGIN, S. (Dir.) (2006). *La Médisance* (115-138). Reims : Presses Universitaires de Reims.

TRAVERSO, V. (2012). Analyses interactionnelles : repères, questions saillantes et évolution. *Langue française* 175, 3-17.

VAISSIERE, J. (2006). *La phonétique*. Paris : Presses Universitaires de France.

WATTS, R., IDE, S. & EHLICH, K. (Eds.), (1992-2005 Second Edition). *Politeness in Language*. La Hague, Pays-Bas : Mouton de Gruyter.

WENNERSTROM, A. (2001). *Music of Everyday Speech*. Oxford : Oxford University Press.