

Université Paris Cité

Ecole doctorale 622 : Sciences du Langage

*Laboratoire de Linguistique Formelle (LLF), UMR 7110*

# **Production et perception des contours continuatifs français par des apprenants japonophones**

**Rachel ALBAR**

Thèse de doctorat de linguistique

Dirigée par Hiyon Yoo

Présentée et soutenue publiquement le 30 mars 2022

Devant un jury composé de :

<b>Hiyon Yoo</b>	MCF, Université Paris Cité	Directrice
<b>Élisabeth DELAIS-ROUSSARIE</b>	DR, Université de Nantes	Rapporteure
<b>Claire PILLOT-LOISEAU</b>	MCF, Université Sorbonne Nouvelle	Rapporteure
<b>Ingo FELDHAUSEN</b>	PR, Université de Lorraine	Examineur
<b>Takeki KAMIYAMA</b>	MCF, Université Paris 8	Examineur
<b>Philippe MARTIN</b>	PR Emérite, Université Paris Cité	Président



Except where otherwise noted, this work is licensed under  
<https://creativecommons.org/licenses/by-nc-nd/3.0/fr/>



---

**Titre : Production et perception des contours continuatifs français par des apprenants japonophones**

**Résumé :** En français, les frontières prosodiques sont associées à des montées de la fréquence fondamentale ainsi qu'un allongement vocalique plus ou moins importants selon la force de frontière. Par exemple, la fin d'un syntagme nominal sujet est généralement réalisé par l'un de ces mouvements prosodiques, autrement appelés « contours continuatifs ». Certains types de contours montants peuvent également avoir une fonction discursive et indiquer la continuité de l'acte de parole. En japonais, les frontières prosodiques ne sont généralement pas associées à un contour montant bien que des mouvements de pitch (BPM) peuvent être réalisés, notamment en parole spontanée. L'objectif de cette thèse est d'observer comment les apprenants japonophones perçoivent et produisent ces contours, et quelle est l'influence de la langue maternelle. Nos principaux résultats montrent que bien qu'elles ne soient pas présentes dans leur L1, les apprenants arrivent rapidement à produire les montées mélodiques. Cependant, des difficultés persistent dans la distinction des frontières prosodiques et des différentes forces de continuation qui leur sont associées, que ce soit au niveau perceptif ou dans la production des indices de durée de F0. Le traitement des indices prosodiques par les apprenants diffère aussi selon le type de tâche (lecture ou parole semi-spontanée), mais on observe également différentes stratégies chez les apprenants indépendamment de leur niveau de langue. Nos résultats mettent donc en lumière la grande variabilité de l'interlangue des apprenants. Les stratégies des apprenants, bien que pouvant être fortement liées à leur L1 (notamment dans la production des indices prosodiques), ne correspondent cependant pas obligatoirement à un phénomène de transfert direct.

**Mots clés :** prosodie L2, interlangue, contours continuatifs, frontières prosodiques, Français Langue Etrangère, apprenants japonophones

---

**Title : The production and perception of French continuation contours in Japanese learners**

**Abstract :** In French, prosodic boundaries are associated with a rise in F0 and a lengthening of the accented vowel, which vary to a greater or lesser extent depending on the strength of the boundary. For example, the end of a subject noun phrase is usually produced with one of these prosodic movements, named “continuation contours”. Certain rising contours also have a discursive function and indicate the continuity of speech. In Japanese, prosodic boundaries are not usually associated with a rise, even if some boundary pitch movements (BPM) can occur, especially in spontaneous speech. The goal of this thesis is to observe how Japanese learners of French perceive and produce these contours, and to examine to what extent their production is influenced by their L1. Our main results show that despite not being present in their L1, learners are quickly able to produce melodic rises in French. However, difficulties in distinguishing between prosodic boundaries and between the different levels of continuation that are associated with them remain, both in perception and in the production of prosodic cues (duration and F0). The use of these prosodic cues by the learners also differs depending on the type of task (reading task or semi-spontaneous speech) and we can also observe different strategies in the learners, independent from their proficiency level. Therefore, our results highlight the variability of the learners’ interlanguage. Learners’ strategies, while being strongly linked with their L1 (especially in their production of prosodic cues), are not entirely explained by a phenomenon of direct transfer.

**Key words :** L2 prosody, interlanguage, continuation contours, prosodic boundaries, French as second language, Japanese learners

# Remerciements

Il ne fait nul doute que ce projet de thèse n'aurait pu aboutir sans l'aide et le soutien de nombreuses personnes que je tiens à remercier ici.

Tout d'abord, mes remerciements vont à Hiyon Yoo pour m'avoir accompagnée dans ce projet de thèse et m'avoir poussée à me dépasser, en particulier en m'incitant à présenter mes recherches dans de nombreux colloques, ce qui fut à la fois motivant et formateur. Merci pour tes encouragements et tes conseils qui m'ont guidée tout au long de ce travail.

Je souhaite également remercier Élisabeth Delais-Roussarie, Claire Pillot-Loiseau, Ingo Feldhausen, Takeki Kamiyama et Philippe Martin de m'avoir fait l'honneur d'accepter de faire partie de mon jury, ainsi qu'à Harim Kwon pour sa participation en qualité de membre invité.

Merci à Harim Kwon et Takeki Kamiyama pour avoir fait partie de mon comité de suivi et pour m'avoir prodigué de nombreux conseils et aidée à évoluer en tant que chercheuse.

A Ioana Chitoran, qui m'a donné l'envie d'étudier la phonétique par son cours de première année de Master, et qui a toujours été présente pour tous les doctorants. Mais également à Philippe Martin et Georges Boulakia pour nos échanges lors de leurs fréquents passages au labo.

A tous les membres du labo « Phonétique-Phonologie » alliant des doctorants de LLF et CLILLAC-ARP, au bureau 529, en particulier à Hannah pour son aide précieuse à tous moments, mais aussi à Darya, Anqi, Qianwen, Anisia, Patricia et Ismaïl pour toutes ces joyeuses discussions autour d'un thé, lors de nos petites pauses collectives.

Merci à tous ceux qui ont prêté leur voix pour les expériences : Rémi, Patty, Chloé, et Alexandre. Je remercie aussi Fabián Santiago pour m'avoir partagé ses enregistrements de francophones natifs. Merci également à Giuseppina Turco pour son aide pour l'analyse statistique de mes données perceptives.

---

Je tiens également à remercier le laboratoire de phonétique de l'université Sophia à Tokyo pour m'avoir accueillie afin de réaliser mes expériences. Je remercie plus particulièrement Mafuyu Kitahara et Ayako Kitamura pour m'avoir gentiment accueillie et aidée à trouver des participants pour mes expériences et à les mettre en place. Par ailleurs, merci à tous les participants d'avoir passé ces expériences, parfois un peu répétitives, avec le sourire.

Merci à mes amis pour m'avoir soutenue moralement, en particulier Céline et Clémence qui ont toujours su me rassurer et me motiver dans les moments difficiles. Merci pour cette précieuse amitié !

Enfin, je tiens à remercier ma famille pour son soutien et sa patience, et en particulier ma mère pour sa relecture minutieuse. Merci infiniment à mon mari Yutaka pour avoir été à mes côtés toutes ces années et de m'avoir comblée par la naissance de notre fille Elena qui me donne tous les jours le courage et la force d'aller vers l'avant.

# Table des matières

<i>Table des figures</i>	<i>xvii</i>
<i>Liste des tableaux</i>	<i>xix</i>
<b>Introduction générale</b>	<b>1</b>
<b>I Etat de l'art</b>	<b>5</b>
<b>1 Acquisition L2 et interlangue</b>	<b>7</b>
1.1 L'erreur comme point de départ . . . . .	7
1.2 L'hypothèse de l'Analyse Contrastive . . . . .	9
1.3 Interlangue . . . . .	11
1.4 Modèles d'acquisition L2 : PAM et SLM . . . . .	15
1.4.1 Le Perceptual Assimilation Model (PAM) . . . . .	15
1.4.2 Le <i>Speech Learning Model</i> (SLM) . . . . .	17
1.5 Quel lien entre l'acquisition de la perception et de la production? . . . . .	19
1.6 Facteurs de variation . . . . .	24
1.6.1 L'âge de début d'apprentissage . . . . .	24
1.6.2 L'input . . . . .	26
1.6.3 La motivation . . . . .	29
1.6.4 Les capacités individuelles . . . . .	31
1.6.5 Le type de parole . . . . .	33
1.7 Synthèse . . . . .	35
<b>2 Systèmes prosodiques du français et du japonais</b>	<b>37</b>
2.1 Accentuation et rythme des deux langues . . . . .	37

2.2	Phrasé et hiérarchie prosodique . . . . .	43
2.2.1	Le système prosodique du français . . . . .	43
	Différentes descriptions prosodiques . . . . .	43
	<i>Delattre</i> . . . . .	46
	<i>Arbres et grilles métriques</i> . . . . .	46
	<i>Contours de Philippe Martin</i> . . . . .	48
	<i>Mertens</i> . . . . .	49
	<i>Di Cristo et Hirst</i> . . . . .	49
	<i>ToBI</i> . . . . .	50
	Les contours continuatifs et leur rôle dans la hiérarchie prosodique	50
	La question du syntagme intermédiaire . . . . .	54
	La syntaxe et la prosodie sont-elles isomorphes? . . . . .	58
	Récapitulatif : les contours continuatifs et leurs équivalences . . . . .	60
2.2.2	Le système prosodique du japonais . . . . .	61
2.3	Prosodie et interaction : le rôle du continuatif . . . . .	64
	Le rôle dialogique et interactionnel des contours continuatifs du	
	français . . . . .	64
	Des contours montants à valeur discursive en japonais – les <i>Boundary Pitch Movements</i> (BPM) . . . . .	71
2.4	Synthèse . . . . .	76
<b>3</b>	<b>Prosodie et contours montants en L2</b>	<b>79</b>
3.1	La prosodie des questions en L2 . . . . .	79
3.2	Accentuation, organisation prosodique et rythme en L2 . . . . .	84
3.3	Synthèse . . . . .	101
<b>4</b>	<b>Objectifs et questions de recherche</b>	<b>107</b>
<b>II</b>	<b>Etudes expérimentales</b>	<b>111</b>
<b>5</b>	<b>Expériences de production des contours continuatifs en parole lue et semi-spontanée</b>	<b>115</b>

5.1	Introduction . . . . .	115
5.2	Protocole expérimental . . . . .	118
5.2.1	Corpus . . . . .	118
5.2.2	Participants . . . . .	119
5.2.3	Protocole . . . . .	121
5.3	Expérience 1 : Lecture de textes . . . . .	122
5.3.1	Introduction . . . . .	122
5.3.2	Méthodologie . . . . .	123
5.3.3	Résultats . . . . .	125
	Observations générales . . . . .	125
	Résultats : Les montées mélodiques . . . . .	127
	Analyse des données de F0 . . . . .	133
	Analyse de l'indice de durée . . . . .	138
	Allongement vocalique et ratio . . . . .	141
5.3.4	Synthèse . . . . .	143
5.4	Expérience 2 : Description d'images . . . . .	147
5.4.1	Introduction . . . . .	147
5.4.2	Méthodologie . . . . .	148
5.4.3	Résultats . . . . .	150
	Observations générales . . . . .	150
	Résultats : les montées mélodiques . . . . .	152
	Analyse des données de F0 . . . . .	154
	Analyse de l'indice de durée . . . . .	155
	Allongement vocalique et ratio . . . . .	161
5.4.4	Synthèse . . . . .	165
5.5	Mise en parallèle et discussion . . . . .	167
<b>6</b>	<b>Expériences de perception et de reproduction d'énoncés</b>	<b>175</b>
6.1	Introduction . . . . .	175
6.2	Protocole . . . . .	176
6.2.1	Participants . . . . .	176

6.2.2	Conditions du recueil des données . . . . .	178
6.3	Expérience 3 : Reproduction de phrases au contour prosodique modifié . . .	178
6.3.1	Introduction . . . . .	178
6.3.2	Méthodologie . . . . .	179
	Matériel linguistique . . . . .	179
	Protocole . . . . .	182
	Participants . . . . .	182
6.3.3	Résultats . . . . .	182
6.3.4	Synthèse . . . . .	187
6.4	Expérience 4 : Perception des différents niveaux prosodiques : tâche de complétion de phrases . . . . .	189
6.4.1	Introduction . . . . .	189
6.4.2	Méthodologie . . . . .	190
	Matériel linguistique . . . . .	190
	Protocole . . . . .	192
	Participants . . . . .	193
6.4.3	Résultats . . . . .	193
6.4.4	Synthèse . . . . .	197
6.5	Discussion . . . . .	198
<b>III</b>	<b>Bilan et perspectives</b>	<b>203</b>
<b>7</b>	<b>Discussion générale</b>	<b>205</b>
<b>8</b>	<b>Proposition d'application didactique</b>	<b>211</b>
8.1	Méthodes d'entraînement en Prosodie L2 . . . . .	211
8.1.1	Les tâches de <i>shadowing</i> et de répétition . . . . .	211
8.1.2	La visualisation . . . . .	214
8.1.3	Les gestes . . . . .	218
8.2	La place actuelle de l'enseignement des continuatifs en FLE . . . . .	221
8.3	Idées et propositions didactiques : comment enseigner les continuatifs en FLE ? . . . . .	228

<b>Conclusion</b>	<b>239</b>
<b>Annexes</b>	<b>247</b>
<b>A Formulaires de consentement</b>	<b>249</b>
<b>B Questionnaire</b>	<b>255</b>
<b>C Expériences de production</b>	<b>257</b>
<b>D Expériences de perception</b>	<b>265</b>
<i>Bibliographie</i>	<i>269</i>



# Table des figures

2.1	Exemples issus de Mertens (1993) . . . . .	42
2.2	Différents schémas accentuels du AP selon Jun et Fougeron (2000) . . . . .	44
2.3	Equivalences des termes et expressions employés dans les travaux sur le phrasé prosodique du français par Di Cristo (2016) . . . . .	45
2.4	Les 10 intonations par Delattre (1966b) . . . . .	47
2.5	Exemple d'arbre métrique – Dell (1984) . . . . .	47
2.6	Exemple de grille métrique – Delais-Roussarie (2000) . . . . .	48
2.7	Exemple d'annotation par Martin (2018) . . . . .	48
2.8	Exemple d'annotation par Mertens 1993 . . . . .	49
2.9	Exemple de contraste de pente par Martin (2008) . . . . .	53
2.10	Exemples issus de Michelas (2011) : syntagmes intermédiaires de 2 (figure a.) et 3 AP (figure b.) . . . . .	56
2.11	AP accentué et non accentué (sans particule) – exemple issu du Handbook of Japanese Phonetics and Phonology, Igarashi (2015) . . . . .	62
2.12	Downstep en japonais – exemple issu du Handbook of Japanese Phonetics and Phonology, Igarashi (2015) . . . . .	62
2.13	Différents modèles prosodiques – Exemple issu du Handbook of Japanese Phonetics and Phonology, Igarashi (2015) . . . . .	63
2.14	Différents patrons intonatifs d'énumération – issu de Di Cristo 2016 . . . . .	70
2.15	BPM H% – exemple issu du Handbook of Japanese Phonetics and Phono- logy Igarashi (2015) . . . . .	72
2.16	BPM HL% – exemple issu de Sasaki-Hara 2004 . . . . .	73
3.1	Exemple issu de Anan 1981 : réalisation de la phrase « La télévision en couleur, une trouvaille » par un locuteur français et un locuteur japonais . . . . .	99

5.1	Exemple d'annotation Praat . . . . .	126
5.2	Pourcentages moyens de contours montants produits à différentes frontières prosodiques selon le niveau de langue. Les barres d'erreur représentent les écarts-types . . . . .	128
5.3	Pourcentage de contours montants produits aux frontières fortes (AP, ip, IP) selon le taux de montées réalisées par les natifs (items à 0% de montées mélodiques, 20-35%, 75-80% et 100%) . . . . .	129
5.4	Réalisation de l'énoncé « j'aimerais savoir s'il y aura un wagon restaurant » par JP5 . . . . .	130
5.5	Réalisation de l'énoncé « Ma soeur a une peur bleue de l'obscurité » par JP39	131
5.6	Réalisation de l'énoncé « La SNCF prévoit un wagon de jeux pour les enfants » par JP26 . . . . .	131
5.7	Réalisation de l'énoncé « Hier, je suis allé me promener dans les calanques » par JP20 . . . . .	132
5.8	Exemples de productions de contours montants de 4 apprenants et 3 francophones . . . . .	133
5.9	Exemple de production du locuteur JP9 . . . . .	133
5.10	Ecart de F0 suivant la frontière prosodique et le niveau de langue (en demi-tons) . . . . .	135
5.11	Ecart de F0 produit par 6 sujets aux frontières prosodiques de AP et IP (en demi-tons) . . . . .	135
5.12	Réalisation par un apprenant (JP5) de la phrase « Il fait un temps magnifique et la mer est très bonne » . . . . .	136
5.13	Réalisation par un francophone (FR11) de la phrase « Il fait un temps magnifique et la mer est très bonne » . . . . .	137
5.14	Exemple de réalisation de découpage en chunks suivis de pauses avec montées mélodiques importantes par l'apprenant JP26 phrase « Au moins, il la protégerait en aboyant au moindre danger » . . . . .	138
5.15	Exemple de production avec faible étendue mélodique, produit par l'apprenante JP39, phrase « Si je dois partir avant midi de Marseille » . . . . .	139

5.16	Durée des voyelles accentuées selon la frontière prosodique et le niveau de langue (en millisecondes) . . . . .	140
5.17	Durée des voyelles accentuées de 6 sujets aux frontières prosodiques de AP et IP . . . . .	141
5.18	Ratio voyelles accentuées/non accentuées selon le niveau de langue . . . .	142
5.19	Réalisation par une francophone (FR07) de l'énoncé « Si je dois partir avant midi de Marseille » . . . . .	143
5.20	Réalisation par une apprenante japonophone (JP15) de l'énoncé « Si je dois partir avant midi de Marseille » . . . . .	143
5.21	Exemple de segmentation Praat pour la tâche de description d'images . . .	150
5.22	Pourcentage de contours montants produits aux frontières de AP et IP selon le niveau de langue . . . . .	153
5.23	Pourcentage de contours montants produits aux frontières de AP et IP : exemple de 6 locuteurs . . . . .	153
5.24	Ecart de F0 suivant la frontière prosodique et le niveau de langue (en demi-tons) . . . . .	154
5.25	Ecart de F0 produit par 6 sujets aux frontières prosodiques de AP et IP (en demi-tons) . . . . .	155
5.26	Durée des voyelles accentuées selon la frontière prosodique et le niveau de langue (en millisecondes) . . . . .	156
5.27	Durée des voyelles accentuées de 6 locuteurs aux frontières prosodiques de AP et IP . . . . .	157
5.28	Exemple de production du sujet FR2, description de l'image 2. « Avec les couleurs typiques bleues et euh sur les murs et puis le plancher + et euh un grand lit + et pas mal de couleurs vives » . . . . .	159
5.29	Exemple de production du sujet JP9, description de l'image 2. « il y a un lit + il y a + il y a un peu de + un peu de tableaux + et + il y a deux + deux chaises + deux portes » . . . . .	160
5.30	Ratio voyelles accentuées/non accentuées selon le niveau de langue . . . .	161

5.31	Exemple de production d'un francophone (FR4) lors de la description de l'image 4. « Euh cette image est une photo + on dirait que c'est une photo de famille prise lors d'un anniversaire + parce que y'a un gâteau avec des bougies » . . . . .	163
5.32	Exemple de production d'un japonophone (JP26) lors de la description de l'image 4. « C'est la photo de la famille + il y a + six + six personnes + autour de la + gâteau de + anniversaire » . . . . .	164
5.33	Pourcentages moyens de contours montants aux frontières de AP et de IP dans les deux tâches de production . . . . .	169
5.34	Pourcentage de contours montants produits par les apprenants aux frontières de IP dans les deux tâches de production . . . . .	169
5.35	Ratio voyelle accentuée en frontière de AP et de IP / voyelle non accentuée dans les deux tâches de production . . . . .	171
5.36	Montées de F0 aux frontières de AP et de IP dans les deux tâches de production	172
6.1	Exemple de stimuli modifiés dans les quatre conditions de test . . . . .	181
6.2	Pourcentage moyen de schémas prosodiques grammaticaux par niveau de langue dans les quatre conditions prosodiques . . . . .	185
6.3	Répartition des moyennes de productions de schémas prosodiques grammaticaux par niveau de langue dans les quatre conditions prosodiques . . .	185
6.4	Pourcentage moyen de contours corrigés suivant le niveau de langue pour les stimuli 1 AP, condition L . . . . .	186
6.5	Répartition des moyennes de contours corrigés suivant le niveau de langue pour les stimuli 1 AP, condition L . . . . .	187
6.6	Proportions des schémas prosodiques produits suivant le niveau de langue et la condition prosodique . . . . .	188
6.7	Type de stimuli – Tâche de complétion de phrases . . . . .	191
6.8	Exemple de réalisation « non-prototypique » utilisée comme stimulus pour la condition AP . . . . .	192
6.9	Réponses globales correctes selon le niveau de langue . . . . .	194
6.10	Réponses correctes selon le niveau de langue et la frontière prosodique . .	195

6.11 Réponses correctes selon le niveau de langue et le nombre approximatif d'heures d'exposition au français / semaine en dehors des heures de cours .	197
6.12 Illustration hypothèse explicative concernant la tendance à de meilleurs résultats en condition ip dans l'expérience de complétion de phrases . . . .	202
8.1 Image activité didactique 4 . . . . .	237



# Liste des tableaux

2.1	Tableau des équivalences entre les différentes descriptions prosodiques . . .	60
5.1	Données sur les apprenants japonophones ayant participé à l'expérience de production . . . . .	120
5.2	Informations globales sur les données de lecture de textes . . . . .	126
5.3	Informations générales sur les productions de la tâche de description d'images	151
6.1	Données sur les apprenants japonophones ayant participé aux expériences de perception et de reproduction de phrases . . . . .	177
6.2	Stimuli expérience 3 . . . . .	181
6.3	Contrastes de l'interaction niveau de langue * condition prosodique, condi- tion 3 (H_L) . . . . .	184
6.4	Sommaire du modèle GLMM – résultats des effets fixes . . . . .	195
8.1	Méthodes de FLE analysées . . . . .	221
D.1	Stimuli Expérience 3 . . . . .	266



# Introduction générale

Manier la prosodie d'une langue étrangère est une tâche aussi difficile que nécessaire. Bien loin d'être superficielle, la prosodie joue un rôle primordial dans la compréhension et l'intelligibilité (Munro & Derwing, 2001 ; Quené & Van Delft, 2010 ; van Maastricht et al., 2016), et une « bonne » prosodie permet une communication fluide entre les locuteurs. Malgré son importance indéniable, la prosodie est de nos jours très peu enseignée dans les cours de FLE qui semblent privilégier l'acquisition de la grammaire et du lexique au détriment des compétences phonético-phonologiques (Sauvage & Billières, 2019).

Par ailleurs, le nombre d'études s'intéressant à la prosodie des apprenants en L2 reste limité par rapport aux études segmentales, et la majorité d'entre elles portent sur l'accentuation lexicale (*stress* de l'anglais, tons du mandarin, etc.), le plus souvent via des expériences contrôlées de lectures de phrases. Ainsi, jusqu'à présent, peu d'études ont été menées sur la prosodie dans un sens plus large et la réalisation des contours prosodiques, et c'est entre autres à ce manque que nous tentons de remédier par notre travail de thèse.

Lorsqu'on porte son attention sur l'intonation du français, on peut se rendre compte que celle-ci est caractérisée par un grand nombre de montées mélodiques. La langue française apparaît comme une « langue au sein de laquelle les montées mélodiques sont récurrentes » (D'Imperio et al., 2016). Cependant, il serait réducteur de résumer le français à une langue qui « monte ». D'une part, parce que ces montées bien que majoritaires ne sont pas nécessairement obligatoires, et que leur réalisation dépend de leur emplacement, de contraintes syntaxiques et métriques, du style de parole, ou de bien d'autres facteurs, linguistiques ou non. D'autre part, parce que la mélodie n'est pas le seul indice prosodique à prendre en considération et que la durée tient un rôle tout aussi important, voire plus, dans la réalisation de ces « montées » non-finales que nous nommons dans notre recherche **contours continuatifs**. Le principal rôle des contours continuatifs est démarcatif ; ceux-

ci permettent de structurer l'énoncé en unités prosodiques. Ces unités étant grandement liées à la syntaxe, les contours continuatifs facilitent ainsi la compréhension du discours. Les contours produits diffèrent dans leur réalisation selon le type de frontière prosodique : Delattre (1966b) distingue ainsi les continuations mineures des continuations majeures, ces dernières étant réalisées par un allongement final et une montée mélodique plus amples que les continuations mineures. A l'oral, les continuations majeures ont également un rôle discursif et permettent entre autres d'indiquer le maintien du tour de parole (Portes & Bertrand, 2005).

En raison des rôles essentiels que jouent les contours continuatifs du français et de leur omniprésence dans la langue, l'intérêt d'étudier ce phénomène chez les apprenants L2 nous semble indéniable. A notre connaissance, peu d'études se sont intéressées à la production et à la perception des continuatifs en français L2, et l'une des seules études conséquentes sur le sujet est la thèse de Santiago (2014) sur les apprenants hispanophones du français, qui intègre des analyses sur les contours continuatifs produits en frontière de syntagme intonatif. Quelques études portent également sur l'accentuation finale du français chez les apprenants japonophones, mais celles-ci ont une portée limitée et ne proposent que des analyses de lectures de phrases (Anan, 1981 ; Kamiyama, 2005 ; Lepetit, 1989) ou d'un court texte (Kojima, 1978).

Dans cette thèse, nous cherchons donc à élargir le champ des connaissances dans le domaine de la prosodie L2, en nous focalisant sur les contours continuatifs, un phénomène central dans la prosodie du français et jusqu'à présent peu étudié. Pour cela, nous présentons quatre expériences : deux expériences de production (tâche de lecture de textes et tâche de description d'images) et deux expériences liées aux compétences perceptives des apprenants (tâche de reproduction d'énoncés modifiés et tâche de complétion de phrases).

Cette thèse est divisée en trois parties. Dans la première partie, nous présentons les concepts fondamentaux et les connaissances nécessaires à la compréhension de notre sujet de recherche. Dans le Chapitre 1, nous nous intéressons à l'acquisition L2 et au concept d'interlangue. Dans le Chapitre 2, nous comparons les systèmes prosodiques du français et du japonais, et en particulier du point de vue des contours mélodiques non-finaux et de la hiérarchie prosodique. Ensuite, dans le Chapitre 3, nous faisons une revue de la prosodie L2. A la fin de cette première partie, dans le Chapitre 4, nous expliquons nos objectifs de

---

thèse et nous exposons les questions de recherche qui nous guideront tout au long de ce travail.

Dans la Partie 2 nous présentons nos expériences de production et de perception des contours continuatifs français par des apprenants japonophones. Les deux premières expériences, présentées dans le Chapitre 5, portent sur la production des contours continuatifs dans une tâche de lecture de textes, analysant ainsi la parole lue (Expérience 1), et de description d'images (Expérience 2), de parole semi-spontanée. Dans le Chapitre 6 sont présentées les deux expériences nécessitant l'usage de compétences perceptives : une expérience de reproduction d'énoncés modifiés (Expérience 3) et une expérience de complétion de phrases (Expérience 4).

Enfin, la Partie 3 conclut cette thèse en présentant la discussion générale – liant par la même occasion les résultats obtenus dans les expériences de production et ceux des expériences perceptives (Chapitre 7), suivie d'une proposition d'application didactique intégrant la revue des études testant des méthodes d'entraînement en prosodie L2, l'analyse de la place de l'enseignement des contours continuatifs du français dans les manuels de FLE, et terminant par la présentation d'idées d'activités didactiques pour leur enseignement (Chapitre 8).



# **Première partie**

## **Etat de l'art**



# Chapitre 1

## Acquisition L2 et interlangue

### 1.1 L'erreur comme point de départ

Les apprenants en langue étrangère produisent inévitablement ce qui est communément appelé des **erreurs**. L'erreur apparaît comme une méconnaissance des règles propres à la langue cible. Ainsi, Corder (1967) les distingue des disfluences involontaires que l'on peut aussi parfois observer dans les productions des natifs qu'il appelle *mistakes*, **fautes**. C'est le cas notamment des lapsus ou des oublis, qui peuvent être liés à des conditions extérieures telles que la fatigue du locuteur, mais en aucun cas n'expriment une méconnaissance des règles de la langue. De même, les erreurs des apprenants qui ne sont pas systématiques et peuvent être corrigées sont considérées comme des fautes. Il semble toutefois difficile de différencier les fautes des erreurs dans les productions des apprenants à moins de prendre en compte la fréquence des occurrences, et le terme « *mistake* » (faute) est actuellement généralement employé comme synonyme de « *error* » (erreur) dans la littérature en acquisition et enseignement L2.

Bien que l'erreur soit plus facilement associée à la production, puisque c'est ce qu'un locuteur natif va directement percevoir dans le signal, elle peut être également perceptive. En effet, une mauvaise perception et catégorisation des phonèmes de la L2 peut également être défini comme une erreur. Tout comme les erreurs de production, elles peuvent potentiellement entraver la compréhension et la communication entre l'apprenant et l'interlocuteur, et ce malgré l'aide du contexte. Ainsi, certains problèmes de communication

trouvent leur origine dans une (petite) erreur de perception comme l'explique Field (2003) :

*it is also a fact that many high-level breakdowns of communication originate in low-level errors. A second-language listener who hears I won't go to London as I want to go to London, is making a small mistake based on phoneme discrimination. But this mistake may impact upon the interpretation of what comes next, and may even influence understanding of the text as a whole. (p.325)*

Les erreurs segmentales et suprasegmentales chez les apprenants sont responsables de l'« accent étranger » qui est défini par Rasier et Hiligsmann (2007) en ces termes : « *the perception of general and discrete deviations from the generally accepted norm of pronunciation of a language that are reminiscent of another language.* » (p.28)

Les erreurs syntaxiques ou lexicales des apprenants sont généralement décrites par rapport aux règles normatives de la langue cible. Les réalisations effectives des locuteurs natifs peuvent cependant varier et s'écarter de cette norme. On peut observer différents types de **variation** selon la région (diatopique), la dimension sociale ou démographique (diastratique), stylistique ou situationnelle (diaphasique) (Gadet, 1996). À cela s'ajoute le moyen de communication (écrit ou oral) aussi appelé variation diamésique. Ainsi, un même énoncé peut être considéré comme acceptable dans un contexte particulier ou pour des locuteurs natifs d'une région spécifique. Un exemple de variation diamésique bien connu est la suppression du « ne » de négation à l'oral (par exemple : « *Je viens pas* » au lieu de « *Je ne viens pas* ») acceptée et maintenant largement majoritaire, tandis que son absence est plus facilement jugée « inacceptable » ou « agrammaticale » à l'écrit par les locuteurs natifs. Ainsi, il est fréquent de vérifier l'acceptabilité des énoncés par des locuteurs natifs à l'aide de test de jugement afin de repérer de potentiels usages qui diffèrent de la norme établie.

Cependant, si les erreurs syntaxiques et lexicales sont assez faciles à repérer, il est plus difficile de définir les « erreurs » responsables de l'accent étranger, en particulier concernant les données suprasegmentales telles que le rythme ou l'intonation. En effet, on retrouve les mêmes problématiques de la variabilité selon le lieu, le style de parole et le locuteur. Plus particulièrement, les variations prosodiques liées à « l'individu (jeune, vieux, homme, femme), [le] groupe social (prolétaire, bourgeois), ou [la] circonstance particulière (discours politique, sermon, etc.) » (Léon, 2009, p.160) sont appelés les **phonostyles**.

Mais, en dehors de cette variation stylistique, on observe aussi de la variabilité propre aux phénomènes suprasegmentaux qui permet une grande liberté au niveau du phrasé et différentes réalisations possibles. Le même énoncé peut ainsi être produit de différentes manières dans un certain contexte toute en étant « acceptable » et ne relève que du choix fait par le locuteur, qui n'est pas toujours systématique. C'est le cas de l'énoncé « Max adore les chocolats » présenté par Martin (2018) où l'accentuation peut indépendamment se produire sur la première syllabe (« **Max** ») ou la troisième (« **adore** »). Nous reviendrons sur ces difficultés propres à l'analyse suprasegmentale et ce que cela implique dans l'analyse des données des apprenants un peu plus tard.

## 1.2 L'hypothèse de l'Analyse Contrastive

Les différences existantes entre la langue source et la langue cible des apprenants ont très tôt été identifiées comme responsables de leurs erreurs, du moins en partie. Nous retrouvons cette idée dans la définition du crible phonologique donnée par Troubetzkoy (1949) :

Le système phonologique d'une langue est semblable à un crible à travers lequel passe tout ce qui est dit. Seules restent dans le crible les marques phoniques pertinentes pour individualiser les phonèmes. Tout le reste tombe dans un autre crible où restent les marques phoniques ayant une valeur d'appel; plus bas se trouve encore un crible où sont triés les traits phoniques caractérisant l'expression du sujet parlant. Chaque homme s'habitue dès l'enfance à analyser ainsi ce qui est dit et cette analyse se fait d'une façon tout à fait automatique et inconsciente. Mais en outre le système des cribles, qui rend cette analyse possible, est construit différemment dans chaque langue. L'homme s'approprie le système de sa langue maternelle. Mais s'il entend parler une autre langue, il emploie involontairement pour l'analyse de ce qu'il entend le « crible phonologique » de sa langue maternelle qui lui est familier. Et comme ce crible ne convient pas pour la langue étrangère entendue, il se produit de nombreuses erreurs et incompréhensions. Les sons de la langue étrangère reçoivent une interprétation phonologiquement inexacte, puisqu'on les fait pas-

ser par le « crible phonologique » de sa propre langue. (Troubetzkoy, 1949, p.54)

Vers les années 60 apparaît l'Hypothèse de l'Analyse Contrastive (à présent HAC), avançant que les erreurs des apprenants peuvent être prédites par les différences entre la langue maternelle et la langue cible (Lado, 1957). Ainsi, des phénomènes présents dans les deux langues seront « faciles » à acquérir pour l'apprenant tandis que des phénomènes différents « difficiles ».

Mais cette hypothèse ainsi que l'importance du transfert dans les productions L2 est remise en question. C'est notamment le cas en syntaxe où certaines études observent des résultats différents de ceux attendus par l'HAC. Certains travaux révèlent la présence de phénomènes qui semblent universaux, observables chez des locuteurs de L2 différentes, mais aussi chez les enfants (Dulay & Burt, 1974; Ervin-Tripp, 1974). L'HAC est donc supplantée par de nouvelles théories dont la *Creative Construction Hypothesis* (Dulay & Burt, 1974) faisant un parallèle entre l'acquisition L1 et L2. Selon cette théorie, les erreurs en L2 ne doivent pas être expliquées par le transfert de la L1 mais par un processus de « construction créative ». La théorie qui lui succède et qui aura une importance bien plus conséquente est l'approche des principes et des paramètres (Flynn, 1987; White, 1985). Cette approche chomskyenne est basée sur le concept de Grammaire Universelle qui permettrait d'expliquer les productions des apprenants et les phénomènes universaux que l'on peut observer, tout en prenant à nouveau en considération leur L1 par les paramètres activés lors de son acquisition qui auraient une influence sur leur apprentissage de la L2 lorsque celle-ci n'a pas les mêmes paramètres. White (1985) évoque notamment les difficultés des hispanophones à considérer comme agrammaticales les phrases en anglais avec omission du sujet puisque leur L1 active ce paramètre d'omission (pro drop). Si l'importance du transfert de la L1 a été fortement remise en question concernant les analyses syntaxiques, ce n'est pas le cas pour la phonologie L2 où le transfert a toujours été considéré comme primordial (Corder, 1967; Rasier & Hiligsmann, 2007; Young-Scholten, 1993). Richards (1971) évoque une dichotomie entre l'analyse L2 syntaxique et phonologique : « *Studies of second language acquisition however have tended to imply that contrastive analysis may be most predictive at the level of phonology, and least predictive at the syntactic level* » (p.2).

Malgré cela, à l'instar des observations faites en syntaxe L2, il apparaît que certains ré-

sultats ne peuvent s'expliquer par le transfert phonologique et phonétique de la L1. Brière (1966) a jugé les productions d'apprenants américains de phrases d'une langue composite (à base de français, arabe et vietnamien) et constate ainsi que certains phonèmes sont facilement acquis bien qu'ils ne soient présents dans leur L1.

Tarone (1980), quant à elle, affirme qu'entre 10 et 47% des erreurs produites par les apprenants ayant le portugais, cantonais ou coréen comme L1 ne peuvent s'expliquer par le transfert. Selon elle, il s'agit d'une préférence générale pour les syllabes CV chez tous les apprenants et la L1 n'influencerait que le type de stratégie employée par les apprenants (épenthèse, élision consonantique...). Il apparaît donc que l'analyse contrastive montre également ses limites dans la description des erreurs phonétiques et phonologiques des apprenants.

## 1.3 Interlangue

Le terme d'**interlangue** désigne le système linguistique de l'apprenant L2 situé entre sa langue maternelle et la langue (Selinker, 1972).

Cette notion d'interlangue se retrouve sous un nom différent dans d'autres théories de la même époque : « système approximatif » (Nemser, 1971) ou « compétence transitionnelle » puis « dialecte idiosyncratique transitionnel » (Corder, 1967). Richards (2015), quant à lui, utilise le terme « *interlanguage errors* » pour désigner les erreurs de transfert qu'il distingue des « *intralingual and developmental errors* », c'est-à-dire les erreurs développementales non liées au transfert et qui se retrouvent dans les productions d'apprenants de différentes L1. Cependant, c'est le terme d'interlangue (incluant les erreurs dues au transfert de la L1 et celles résultant d'autres facteurs) qui fut finalement retenu par la suite.

Selinker (1972) cite cinq facteurs d'erreur chez les apprenants L2 :

1. le transfert de la L1 (*language transfer*)
2. la méthode d'apprentissage (*transfer-of-training*)
3. les stratégies d'apprentissage de la L2 (*strategies of second-language learning*)
4. les stratégies de communication L2 (*strategies of second-language communication*)

5. la surgénéralisation des règles et formes linguistiques (*overgeneralization of TL linguistic material*).

Selinker illustre le facteur d'erreur lié à la méthode d'apprentissage par l'exemple des apprenants Serbo-Croates de l'anglais qui ne font pas la distinction entre *he* et *she*, non pas en raison de l'absence d'un tel contraste dans leur L1, mais par l'absence de phrases employant le pronom *she* dans leurs manuels de langue.

Parmi les stratégies d'apprentissage de la L2, Selinker évoque la stratégie de simplification du système linguistique de la langue cible et cite Jain (1969) qui observe l'emploi de verbes soit toujours transitifs, soit toujours intransitifs dans l'anglais des locuteurs indiens.

Selinker cite également Coulter (1968) qui reporte un certain nombre d'erreurs caractéristiques de deux locuteurs russes parlant anglais. Par exemple, il observe une tendance à ne pas produire les articles, le pluriel et le marquage du passé. Si ces erreurs peuvent être associées à une simplification du système linguistique de la langue cible et être considérées comme des stratégies d'apprentissage, Coulter les considère comme des stratégies de communication. En effet, ces erreurs ne semblent pas gêner la compréhension des interlocuteurs et n'auraient donc tout simplement pas été corrigées.

Une autre stratégie fréquente en L2 est la stratégie **d'évitement** (*avoidance*) que nous pouvons aussi classer comme stratégie de communication L2. Cette stratégie est mise en lumière par Schachter (1974) dans son étude sur les relatives chez les apprenants de l'anglais persanophones, arabophones, sinophones et japonophones. En extrayant et analysant les relatives de 50 compositions par groupe de langue (25 de niveau intermédiaire et 25 de niveau avancé), elle observe que les persanophones et les arabophones ont produit significativement plus d'erreurs que les sinophones et les japonophones. Si la simple observation du taux d'erreur des apprenants donne l'impression que les relatives sont plus faciles pour les sinophones et les japonophones, elle met cependant en avant une différence importante entre les deux autres groupes et ceux-ci : les deux groupes avec un taux d'erreur élevé (persanophones et arabophones) ont produit environ deux fois plus de relatives que les deux autres. Selon Schachter, il s'agit d'un phénomène d'évitement de la part des locuteurs sinophones et japonophones. En effet, dans leur L1 les relatives sont pré-nominales tandis qu'elles sont post-nominales en anglais, et les apprenants préfèrent donc éviter de les produire en raison de cette difficulté conséquente. Cependant, si ce phénomène d'évi-

tement existe bien en syntaxe, Schachter émet des doutes quant à la possibilité qu'il soit également présent lors de l'acquisition phonétique et phonologique de la L2 :

*It is quite possible that the avoidance phenomenon does not occur in the acquisition of the phonological subcomponent of the target language, and that there is a qualitative difference in the acquisition of the phonological as opposed to the syntactic subcomponents. In the syntactic subcomponent, the possibility of paraphrase exists, and the student can take advantage of paraphrase relations to avoid constructions he finds difficult, while still getting his idea across. [...] However, there is no such thing as phonological paraphrase, and therefore the avoidance phenomenon is difficult, if not impossible. (p.212)*

En effet, si l'on peut envisager un tel phénomène d'évitement en parole libre par l'usage d'un lexique dont la prononciation est plus aisée, celui-ci n'aurait certainement qu'une faible importance et ne serait limité qu'à certains phonèmes ou structures relativement peu fréquents.

La dernière stratégie qu'évoque Selinker (1972) est la surgénéralisation des règles et formes linguistiques. Selinker donne comme exemple l'emploi du morphème du passé -ed utilisé dans une condition où il n'est pas demandé : « *What did he intended to say?* » Il cite également Jain (1969) qui, toujours en anglais L2, observe un emploi généralisé du verbe *drive* à tous les moyens de transports. Il s'agit ici d'exemples morphosyntaxiques et lexicaux, mais cette stratégie de surgénéralisation est également fréquemment observée en prosodie L2, notamment en ce qui concerne l'accentuation et la production de schémas accentuels (Pešková, 2019 ; Ramírez Verdugo, 2006 ; Rasier & Hiligsmann, 2007).

Dans ses travaux analysant 40 courtes conversations de vingt apprenants hispanophones de l'anglais, Ramírez Verdugo (2006) observe qu'à la différence des natifs les apprenants ne produisent pas de contours bas montants (*low rises*) sur les informations précédemment données ou supposées. Au lieu de cela, ils produisent le même contour descendant habituellement produit sur les éléments nouveaux. Ainsi ils surgénéralisent ce contour descendant en le produisant quel que soit le type d'information.

Dans son étude sur les propriétés prosodiques des énoncés vocatifs produits par les apprenants locuteurs du tchèque L1 de l'espagnol et espagnol L2, Pešková (2019) observe

également une surgénéralisation de contours prosodiques. Elle a enregistré la production de 20 apprenants tchèques de l'espagnol L2 et 20 apprenants de l'italien L2 à l'aide d'une tâche discursive où l'apprenant doit produire un énoncé guidé par un contexte donné. Parmi les résultats obtenus Pešková observe une forte tendance des apprenants en italien L2 à produire des contours H\*+LL% au lieu des L+H\* !H% habituellement produits dans le cas des énoncés vocatifs. La présence de ce contour n'était pas prédite par l'analyse contrastive des deux langues et ne peut s'apparenter à un phénomène de transfert puisqu'il n'est pas utilisé dans la langue maternelle des apprenants, le tchèque. Pešková explique donc l'utilisation de ce contour par une surgénéralisation de ce contour qui est produit dans de nombreux autres énoncés en italien (qu'elle observe tout autant en italien L1 que L2).

Selon Selinker, certaines erreurs produites par les apprenants et étant dues à des causes pouvant être différentes (transfert de la L1, stratégie de communication...) résistent à l'apprentissage et se retrouvent dans les productions des apprenants, même après plusieurs années, et cela quel que soit l'âge de l'apprenant ou la quantité de son exposition à la L2. C'est ce qu'il nomme *fossilisation*. Plus tard, Selinker ajoute à cette définition un critère de durée (de deux à cinq ans). Cependant, cette notion de fossilisation a été critiquée (voir Long 2003), notamment parce qu'il existe peu d'études longitudinales sur plus de trois ans traitant de ce phénomène et qu'il est difficile de vérifier objectivement la potentielle fossilisation d'une erreur. Long préfère alors parler de *stabilisation*, évinçant le problème de permanence sous-entendu par la notion de fossilisation et évoquant des phénomènes testables et qui ne sont pas remis en cause.

Au-delà de ces débats théoriques, cette notion de fossilisation est sans nul doute liée au constat que certaines erreurs sont plus difficilement corrigées que d'autres par les apprenants et tendent à persister chez les locuteurs de niveau avancé, tandis que d'autres notions semblent plus facilement acquises.

La notion d'interlangue permet ainsi de prendre en compte tous les phénomènes présents en L2, sans se limiter au transfert de la L1 vers la L2. Elle offre ainsi l'avantage de s'intéresser aux stratégies universelles des apprenants et de pouvoir décrire l'acquisition de la prosodie L2 dans toute sa complexité.

## 1.4 Modèles d'acquisition L2 : PAM et SLM

L'influence de la L1 sur la L2 est souvent visualisée à travers le prisme du transfert direct de la première langue sur la seconde. Ainsi, si un phénomène présent en L2 n'existe pas dans la L1, les apprenants auront des difficultés à le produire et le percevoir. Cependant, les modèles d'apprentissage PAM (Best) et SLM (Flege) prédisent les difficultés des apprenants face à des segments (et supra-segments) non-natifs, non pas uniquement par l'absence ou la présence de ceux-ci dans la L1 de l'apprenant, mais aussi par la proximité des autres segments natifs par rapport au segment non-natif concerné. Dans cette section, nous présentons ces deux modèles.

### 1.4.1 Le Perceptual Assimilation Model (PAM)

Le Perceptual Assimilation Model de Best (1995) définit la similarité entre les segments natifs et non-natifs par la proximité des lieux d'articulation et des gestes articulatoires, qui déterminerait l'assimilation perceptive des segments non-natifs à des catégories natives. Il est également possible que le son non-natif soit interprété comme un son faisant partie de la parole par ses propriétés sans être assimilé à une catégorie non-native. Il peut aussi dans un cas extrême ne pas être interprété comme un son de la parole, comme un claquement de doigts ou une toux qui sont des sons habituellement non associés au langage.

Les gestes auxquels Best (1995) fait référence (citant Browman et Goldstein 1986, 1989, 1990, 1992) ne correspondent pas à un seul geste unique mais à un groupe de gestes (*gestural constellation*), qui peuvent même s'étendre aux niveaux syllabiques et accentuels, bien que le modèle PAM se base principalement sur les contrastes segmentaux. Certains gestes peuvent être présents dans une langue au niveau phonétique sans l'être forcément au niveau phonologique. Best donne l'exemple du français et de l'anglais qui ont en commun un geste d'ouverture du vélum. En français, ce geste a une valeur phonologique distinctive entre les voyelles nasales et orales. En anglais, ce geste est uniquement observé en coarticulation et ne sert donc aucune valeur distinctive ; il ne fait donc pas partie de *l'espace phonologique* de l'anglais. Un autre critère sur lequel Best attire notre attention est le *phasing* entre les différents gestes. En effet, de mêmes gestes peuvent avoir un *phasing* différent. Best donne l'exemple des consonnes plosives non-voisées, aspirées en anglais et

non-aspirées en français, qui diffèrent principalement par le *phasing*.

Ainsi, Best définit les segments non natifs comme des éléments gestuels ou phasing entre les éléments gestuels qui ne correspondent pas exactement à ceux d'un groupe gestuel natif (catégorie).

Selon le PAM l'assimilation des segments est liée aux contrastes présents en L1 et L2. Plus l'écart entre un son non-natif et les groupes gestuels présents dans la L1 est important, plus l'apprenant est capable de le différencier des catégories natives. Dans certains cas où l'écart est très important, il peut être capable de le considérer comme élément de la parole mais ne l'associer à aucune catégorie native, ce qui permet une bonne discrimination de ce son. Ainsi, il existerait différents schémas d'assimilation des contrastes non-natifs :

- *Two category assimilation* (TC Type) : Chaque segment non-natif est assimilé à une catégorie non native. -> Bonne discrimination.
- *Category-Goodness Difference* (CG Type) : Les deux sons non-natifs sont assimilés à la même catégorie native, mais l'un diffère la prononciation « idéale » native (l'un est acceptable, l'autre non). -> Moyenne à très bonne discrimination
- *Single-Category Assimilation* (SC Type) : Les deux sons non-natifs sont assimilés à la même catégorie mais sont également éloignés ou proches de l'« idéal » natif (les deux sont également acceptables, ou non) -> Mauvaise discrimination attendue.
- *Both Uncategorizable* (UU Type) : Les deux sons non-natifs sont considérés comme faisant partie du système phonétique mais ne sont assimilés à aucune catégorie native -> Discrimination de mauvaise à bonne suivant leur proximité dans le système phonologique natif.
- *Uncategorized versus Categorized* (UC Type) : Un son non-natif est assimilé à une catégorie native, l'autre à aucune d'entre elles. -> Très bonne discrimination attendue
- *Nonassimilable* (NA Type) : Les deux catégories non-natives se retrouvent en dehors du domaine de la parole. -> Bonne à très bonne discrimination

Best donne différents exemples qui attestent la validité de son modèle. Une étude sur la perception de clics zoulous par des anglophones<sup>1</sup> atteste la facilité des locuteurs à percevoir et distinguer des sons qui ne font pas partie du système de leur L1 considérés comme « en dehors du domaine de la parole » (NA Type). D'autres contrastes zoulous sont ensuite testés sur les anglophones. Tout d'abord, un contraste de type TC avec deux consonnes latérales fricatives (voisée et non-voisée) correspondant à deux catégories de l'anglais. Ils ont également testé un contraste de type CG avec une vélaire non voisée aspirée et une consonne vélaire éjective assimilée au /k/ de l'anglais (avec l'une des réalisations les plus proche de l'« idéal » du /k/ anglais) et un contraste de type SC ou léger CG avec deux bilabiales (plosive et non plosive) qui devraient être assimilées à la catégorie /b/ de l'anglais sans que l'une des réalisations ne soit plus proche de l'« idéal ». Tous les résultats ont confirmé le modèle : le dernier contraste a bien donné plus de difficultés aux locuteurs (bien que leur performance ait dépassé le seuil de chance) et le contraste de type TC où les deux sons sont assimilés à deux catégories distinctes en L1 a obtenu de bien meilleurs résultats que le contraste de type CG et SC.

#### 1.4.2 Le *Speech Learning Model* (SLM)

Le *Speech Learning Model* (SLM) de Flege (1995) est principalement basé sur la perception. Selon ce modèle, le système phonétique est adaptable tout au long de la vie en réagissant aux sons perçus en L2 et en créant de nouvelles catégories phonétiques ou en modifiant d'anciennes catégories. Ces catégories phonétiques L1 et L2, sont présentes sur un même plan phonologique.

La principale hypothèse de ce modèle est que plus la différence phonétique perçue entre le son en L2 et son équivalent le plus proche en L1 est grande, plus le son « nouveau » sera facile à discerner car associé à une nouvelle catégorie. A l'inverse, si le son L2 est proche du son L1, la création d'une nouvelle catégorie sera plus difficile pour l'apprenant qui aura donc plus de difficultés à les distinguer. Enfin, si les sons L1 et L2 sont semblables (diaphones), le mécanisme de création d'une nouvelle catégorie est bloqué et leur production sera elle aussi semblable.

---

1. Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts : Zulu click discrimination by English-speaking adults and infants. *Journal of experimental psychology : human perception and performance*, 14(3), 345.

Un des exemples de Flege illustrant ce modèle est celui de japonophones et de leur perception du contraste /l/ - /ɹ/ en anglais. Il évoque la plus grande proximité perceptive du /ɹ/ japonais avec le /l/ anglais qu'avec le /ɹ/ anglais. Le /ɹ/ anglais et le /ɹ/ japonais étant plus éloignés perceptivement, cela permettrait la création d'une nouvelle catégorie phonétique et donc une meilleure perception du phone. Flege appuie cette hypothèse par les résultats d'une étude (citant Flege et al 1995) qui a testé la capacité de deux groupes d'apprenants (un groupe « inexpérimenté » ayant vécu moins de trois ans aux États-Unis et un groupe « expérimenté ») à identifier les liquides de l'anglais. Cette étude a montré que dans les deux groupes d'apprentissage le /ɹ/ anglais était plus facilement identifié que le /l/ anglais. Le /ɹ/ anglais a en effet été correctement identifié dans 92% des cas pour le groupe expérimenté et 76% des cas pour le groupe inexpérimenté, contre respectivement 76 et 63% pour le /l/ anglais.

Cependant, même lorsqu'une nouvelle catégorie est créée par l'apprenant en L2 il est possible qu'elle ne corresponde pas à la catégorie native. Par exemple, ces catégories peuvent être basées sur des paramètres phonétiques différents que pour les natifs. C'est par exemple le cas des locuteurs suédois et finlandais (Flege & Hillenbrand, 1986) qui distinguent le contraste /s/-/z/ en anglais inexistant dans leur L1, en position finale (*peace-peas*) en utilisant les paramètres de durée de la voyelle mais pas la durée de la fricative alors que les natifs utilisent les deux. Cette catégorisation différente de celle des natifs peut donc ainsi engendrer des difficultés de production et un usage non-natif du paramètre de durée.

Le SLM prend également en compte l'âge d'acquisition de la L2. Plus l'apprentissage de la L2 débute tôt, plus la distance perceptive nécessaire entre deux sons pour la création d'une nouvelle catégorie peut être réduite. Ainsi, un jeune enfant est plus facilement capable de créer deux catégories différentes pour deux sons perceptivement très proches tandis qu'un adulte utilisera la même catégorie pour les deux. L'idée du SLM est cependant que le système phonétique peut évoluer avec le temps et que l'apprentissage n'est donc pas limité par un âge critique (voir 1.6 pour plus de détails).

Bien que la plupart des études sur l'acquisition d'une L2 se centralisent sur l'effet de la L1 sur la L2, Flege met en avant le phénomène inverse qui est l'effet de l'acquisition d'une

L2 sur la L1. Dans une de ses études citée<sup>2</sup> il observe que les apprenants anglophones de français de niveau avancé produisent les consonnes /ptk/ en anglais avec un VOT plus court (plus proche du français) que celui produit par les anglophones monolingues. Il observe également cela dans l'autre direction ; en effet, les francophones apprenant l'anglais L2 de niveau avancé produisent des VOT plus longs en français que les francophones monolingues.

Pour finir, il est important de noter que le SLM ne prétend pas que toutes les erreurs sont attribuables à la perception. Par exemple, Flege note que les erreurs observées peuvent aussi être liées à l'instruction en L2 ou au niveau de langue. Il donne l'exemple des sujets français scolarisés qui tendent à remplacer le /θ/ anglais par /s/ tandis que les sujets non scolarisés le remplacent plus fréquemment par /t/.

Si ces deux modèles d'apprentissage ont le mérite de montrer que l'influence de la L1 ne se résume pas à un transfert simple de la L1 à la L2 (en particulier au niveau perceptif), nous constatons qu'ils sont principalement axés sur l'apprentissage des phénomènes segmentaux. Le PAM a toutefois été élargi à l'étude des tons (So & Best, 2014), mais son application montre des limites et, parmi les prédictions faites par le modèle, certaines n'ont pas pu être vérifiées par les études (Wu, 2017).

## 1.5 Quel lien entre l'acquisition de la perception et de la production ?

Nous avons vu que certains modèles d'acquisition sont basés sur la perception, notamment le SLM de Flege ayant pour postulat que les erreurs de production sont généralement issues d'une difficulté à percevoir les sons non-natifs. Mais est-ce vraiment le cas ? Est-ce que la perception précède toujours la production ?

De nombreuses études ont tenté d'éclaircir le lien entre la perception et la production, et une grande partie d'entre elles se sont intéressées en particulier à la distinction /r/-/l/ chez les japonophones et coréanophones.

Sheldon et Strange (1982) ont enregistré 6 apprenants japonophones et 4 anglophones natifs aux États-Unis produisant 16 paires minimales incluant le contraste anglais /r/-/l/

---

2. Flege, J. E., & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, 6(3), 185-202.

dans quatre positions différentes (en position intervocalique, en deuxième position d'un cluster, en position initiale et finale). Les mots produits en isolation ont ensuite été utilisés pour créer un test d'identification. Les sujets enregistrés ont ensuite participé au test de perception, une partie des stimuli étant donc leurs propres productions. Ils observent que les réalisations des apprenants d'un bon niveau d'anglais sont correctement identifiées par les anglophones natifs. En revanche, ces apprenants ont des difficultés à percevoir et identifier leurs propres réalisations ainsi que celles des anglophones natifs. Sheldon and Strange concluent donc qu'ils ont une meilleure capacité à produire les contrastes qu'à ne les percevoir.

Borden et al. (1983) ont quant à eux étudié la production et la perception de ce même contraste chez 13 apprenants coréanophones, enregistrés aux États-Unis. Les apprenants ont passé des expériences avant et après une session d'entraînement de 45 minutes : une expérience d'identification et une expérience de discrimination utilisant un continuum [ra]-[la] de synthèse, ainsi qu'une expérience d'auto-perception où les apprenants doivent imiter des phrases et juger immédiatement la justesse de leur réalisation. Bien qu'ils observent une amélioration dans les résultats de perception et la production du contraste /l/ - /r/ entre le premier et le dernier test, celle-ci n'est pas homogène. Cependant, Borden et al. observent que les apprenants ayant le plus de facilités à s'auto-juger sont également ceux qui ont le mieux progressé en production. Ils avancent donc l'hypothèse que la perception de ses propres réalisations (self-perception) est un pré-requis pour l'amélioration de la production.

Akahane-Yamada et al. (1996) ont également comparé les productions d'apprenants avant et après un entraînement spécifique. Eux aussi se sont intéressés au contraste /l/ - /r/ de l'anglais, chez les apprenants japonophones. Ils ont enregistré les productions de 23 apprenants japonophones dans une tâche de répétition de 55 paires minimales. Ils ont également passé un test d'identification. 11 d'entre eux ont ensuite été entraînés à identifier le contraste durant 45 sessions réparties sur 15 jours, recevant comme seul *feedback* un signal indiquant que la réponse donnée est soit juste soit fausse. Les apprenants restants n'ont pas eu cet entraînement et servent de groupe contrôle. Tous les apprenants ont ensuite été enregistrés suivant le même protocole que le premier test, puis à nouveau 3 et 6 mois plus tard. L'intelligibilité de leurs productions a été évaluée par 10 anglophones natifs. A cela

s'ajoute l'évaluation comparée des productions du pré-test et des post-tests par 10 autres anglophones qui devaient choisir celle qui leur paraissait la meilleure. Akahane-Yamada et al. observent que, contrairement au groupe contrôle, l'intelligibilité des productions du groupe d'apprenants ayant reçu un entraînement est bien plus élevée lors du premier post-test, et les productions des apprenants des post-tests sont préférées à celles du pré-test, même après 3 et 6 mois. Leurs résultats montrent ainsi un effet positif de l'entraînement perceptif sur l'amélioration des productions des apprenants, qui s'inscrirait dans la durée.

Comme Borden et al, Baker et Trofimovich (2006) ont identifié un lien entre l'auto-perception et la capacité à produire correctement la L2. Ils ont fait passer un test de production et de perception à 40 apprenants coréanophones habitant aux États-Unis (depuis 1 à 10 ans) en utilisant des mots de l'anglais ayant une structure CVC avec différentes voyelles. Pour le test de perception, les apprenants devaient identifier des mots produits par des anglophones natifs avec choix multiple et pour le test de production des images leurs étaient présentées et ils devaient dire le mot correspondant. 10 anglophones natifs ont ensuite identifié les mots produits dans cette tâche. Les résultats ont montré une corrélation entre les résultats de perception et de production, à l'exception du groupe d'apprenants de niveau intermédiaire où la moitié d'entre eux ont obtenu de meilleurs résultats en production qu'en perception, et le résultat inverse pour l'autre moitié. Baker et Trofimovich ont ensuite cherché à savoir dans quelle mesure les apprenants étaient capables de s'auto-percevoir et si ces résultats étaient liés à ceux de perception et de production. Trois mois plus tard, les apprenants ont donc passé un troisième test où ils ont identifié les mots produits par les apprenants lors de la tâche de production, incluant leurs propres réalisations. Lorsque les réponses étaient similaires à celles des anglophones natifs, elles ont été notées comme « correctes », donnant ainsi une note de justesse d'auto-perception. Les auteurs observent que les apprenants ayant les meilleurs résultats de production sont également ceux qui ont une bonne auto-perception et que ceux ayant des résultats faibles en production ont également un faible score d'auto-perception. Ils observent également que les apprenants ayant obtenu de meilleurs résultats en production qu'en perception ont une très bonne capacité d'auto-perception, ce qui expliquerait ce résultat inattendu.

Okuno et Hardison (2016) se sont intéressés à l'impact de plusieurs types d'entraînements perceptifs sur la perception et la production du contraste de durée vocalique en

japonais. Ils ont recruté 48 apprenants anglophones du japonais et les ont répartis en trois groupes : un groupe qui suivra un entraînement audio-visuel (test d'identification à choix multiple exposant les participants au son du stimuli et à son spectrogramme), un groupe qui suivra un entraînement uniquement audio (test d'identification à choix multiple avec uniquement le son des stimuli présentés) ainsi qu'un groupe contrôle qui ne suivra aucun entraînement. Les entraînements ont été répartis en 8 sessions de 25 minutes, sur une durée de deux semaines. Tous les apprenants ont été évalués en perception et en production avant et après ces deux semaines à l'aide d'un test à choix multiple et de la lecture à voix haute de stimuli écrits en *romaji* (retranscription du japonais par l'alphabet latin). Les productions des apprenants ont été évaluées par des japonophones natifs à l'aide d'un test d'identification à choix multiple. Okuno et Hardisson observent une amélioration en perception et en production entre le pré-test et le post-test dans les groupes audio-visuel et audio, et une absence d'amélioration dans le groupe contrôle. Si le groupe audio-visuel semble être associé une amélioration plus importante de la perception que le groupe audio, celle-ci n'est pas significative.

La grande majorité des études montrent donc un fort lien entre la perception et la production, que confirme l'influence d'entraînements perceptifs sur les productions des apprenants. Ces résultats corroborent le postulat du SLM avançant que les erreurs de production trouvent leurs origines dans des difficultés perceptives. Detey et Racine (2015) observent cependant un décalage entre l'acquisition de la perception et celle de la production. Dans leur étude, ils ont testé la perception et la production de trois voyelles nasales du français par 22 apprenants japonophones après 4 mois d'apprentissage. Ils ont utilisé un test AXB pour étudier leur capacité à percevoir les voyelles nasales / $\tilde{a}$ /, / $\tilde{e}$ /, / $\tilde{o}$ /, ainsi qu'une tâche de répétition et une tâche de lecture pour tester leur production. Les stimuli utilisés dans les trois tâches étaient des mots monosyllabiques comportant les voyelles étudiées. Leurs résultats montrent des différences entre les tâches de production et le test AXB. En effet, les apprenants japonophones parviennent bien plus facilement à distinguer le contraste / $\tilde{a}$ / - / $\tilde{e}$ / que le contraste / $\tilde{a}$ / - / $\tilde{o}$ /. Ainsi, la voyelle / $\tilde{e}$ / semble plus facile à discriminer, alors qu'en production aucune différence entre les voyelles n'a pu être montrée dans la tâche de répétition, et la voyelle / $\tilde{e}$ / a moins bien été produite par les apprenants dans la tâche de lecture que ne l'ont été les voyelles / $\tilde{a}$ / et / $\tilde{o}$ /. La conclusion de cette étude

est que la perception de cette voyelle précède donc sa production.

Si la majorité des études testent le lien production-perception par un paradigme pré-test, entraînement perceptif suivi d'un post-test, quelques études récentes ont choisi d'analyser le potentiel impact d'un entraînement en **production** sur la perception. C'est le cas notamment de l'étude de Kartushina et al. (2015) sur 27 francophones entraînés à produire des voyelles du danois, langue qu'il leur était jusqu'alors inconnue. Les participants ont été enregistrés lors d'une tâche de répétition de voyelles isolées et leur capacité à percevoir les voyelles a été évaluée par un test AXB, avant et après l'entraînement en production. Deux groupes de participants ont été créés : un groupe de 15 francophones, qui se sont entraînés à reproduire les voyelles en recevant un *feedback* leur permettant de comparer leurs productions et valeurs F1 et F2 aux voyelles cibles produites par des natifs, et un groupe contrôle ne recevant qu'un *feedback* pour le F1, alors insuffisant. Cinq séances d'entraînement de 45 minutes ont été suivies par les participants. Kartushina et al. observent que les participants du groupe ayant reçu le *feedback* complet (F1 et F2) ont amélioré leur production des voyelles entre le pré et le post-test, alors que ce ne fut pas le cas du groupe contrôle. De plus, ils observent une amélioration de la perception avec de meilleurs résultats en post-test qu'en pré-test uniquement dans le groupe expérimental. L'entraînement en production semble donc être associé à une amélioration de la perception et ces résultats renforcent l'idée d'un fort lien existant entre la perception et la production.

Cependant, les résultats de l'étude de Tateishi et Winters (2013) vont dans une direction opposée. Ils se sont intéressés au contraste /l/ - /r/ de l'anglais chez les apprenants japonophones et ont mené leur étude sur 10 apprenants. Leur perception du contraste a été évaluée à l'aide d'un test d'identification de paires minimales, et leur production par un test impliquant la reproduction de paires minimales présentées à l'écrit mais également à l'oral. Après des explications générales sur l'articulation, les apprenants ont ensuite participé à 5 sessions d'entraînement de 30 minutes où ils étaient à nouveau exposés à des paires minimales et des syllabes comportant le contraste étudié, tout en ayant accès à un exemple de production native par des images ultrason. Ils pouvaient ensuite s'entraîner à reproduire les cibles en observant leur propre articulation à l'aide de l'ultrason. Au début de chaque séance, ils recevaient également par l'instructeur un *feedback* sur la séance précédente et des conseils afin d'améliorer leur production du contraste. Après les séances d'entraîne-

ment, la production et perception des participants était à nouveau évaluée. Les résultats obtenus ne montrent aucune influence de l'entraînement articulatoire sur la perception du contraste, alors qu'ils observent bien une amélioration concernant la production, notamment pour le /l/ devenant acoustiquement plus proche de la cible native.

## 1.6 Facteurs de variation

En dehors de l'influence de la L2, divers facteurs jouent un rôle dans l'acquisition de la L2 et peuvent expliquer les résultats obtenus au niveau individuel. En effet, un apprenant ayant commencé à apprendre la langue cible plus précocement ou ayant une motivation plus forte, par exemple, obtiendra potentiellement de meilleurs résultats que les autres. Dans cette section, nous détaillons ainsi plusieurs facteurs de variation : des facteurs de variation individuelle comme l'âge de début d'apprentissage, l'input auquel les apprenants ont été exposés dans leur apprentissage, la motivation, les capacités individuelles, mais également le type de parole (variation situationnelle).

### 1.6.1 L'âge de début d'apprentissage

De nombreuses recherches ont mis en avant l'importance de l'âge dans l'acquisition L2, notamment concernant la possibilité d'atteindre une prononciation native dans une langue étrangère. La notion d'âge critique en acquisition L2 est introduite par Lenneberg (1967), qui explique cette notion par des raisons neurobiologiques et associe la capacité de l'être humain à apprendre une langue à la latéralisation de son cerveau. Selon lui, une fois cette latéralisation terminée (au moment de la puberté) il n'est plus possible de commencer une nouvelle langue et de la parler sans accent étranger. Selon Lenneberg, la période précédant la latéralisation du cerveau chez l'enfant serait une « période critique » et son hypothèse sera donc reprise dans les études qui lui succèdent sous le nom d'Hypothèse de la Période Critique (HPC).

L'âge critique varie selon les études. Selon Patkowski (1994) on observerait une discontinuité dans la capacité à apprendre une langue étrangère aux alentours de 15 ans. Selon Long (1990) l'âge critique serait différent suivant les compétences linguistiques concernées : pour les compétences phonologiques celui-ci se situerait aux alentours de 6 ans pour la plupart des enfants ainsi que de 12 ans pour les autres, tandis que pour les compétences

morphologiques et syntaxiques celui-ci se situerait autour de 15 ans. Il précise également qu'un âge inférieur à 6 ans lors du début de l'apprentissage de la langue étrangère garantit une prononciation native.

Cependant, d'autres études ne montrent pas une telle discontinuité et observent une tendance linéaire (Flege et al., 1995; Flege et al., 1999; Oyama, 1976). De plus, certaines études contredisent l'affirmation de Long et montrent qu'un âge inférieur à 6 ans lors de l'arrivée en pays étranger ne corrèle pas toujours avec l'absence d'accent étranger.

C'est le cas de l'étude de Flege et al. (1995) sur l'accent étranger de 240 italiens arrivés au Canada entre 2 et 23 ans. Les participants devaient répéter cinq phrases lues et entendues puis leurs productions ont été jugées par 10 anglophones. Les participants ont été groupés selon leur âge d'arrivée au Canada formant 10 groupes distincts. Flege et al. observent que l'un des juges anglophones natifs est capable de percevoir un accent étranger dans le premier sous-groupe où les locuteurs sont arrivés très tôt au Canada (à l'âge de 3 ans en moyenne) et à le distinguer du groupe contrôle ne comportant que des locuteurs natifs.

A l'inverse, d'autres études montrent la présence d'apprenants tardifs avec une prononciation native ou quasi-native. Bongaerts s'est particulièrement intéressé à ce phénomène. Dans une première étude Bongaerts et al. (1997) font juger les productions d'apprenants néerlandophones de 6 phrases en anglais par 13 locuteurs anglophones natifs. Trois groupes de participants ont passé l'expérience : un premier groupe de 11 apprenants néerlandophones de niveau très avancé, un deuxième groupe de 20 apprenants néerlandophones de niveau varié ainsi qu'un groupe contrôle de dix locuteurs natifs. Dans le groupe d'apprenants de niveau très avancé, 5 d'entre eux obtiennent des résultats semblables aux natifs pour au minimum 5 phrases sur 6. Parmi ces 5 apprenants, 3 montrent même des résultats comparables aux natifs pour toutes les phrases, sans exception.

Dans une autre étude (Bongaerts et al., 2000), ils enregistrent la lecture de 10 phrases en français par 9 apprenants hollandais spécialement choisis pour leur bonne maîtrise de la langue, un groupe de 18 apprenants hollandais plus variés ainsi qu'un groupe contrôle de 9 natifs francophones. Ils font évaluer le degré d'accent étranger de ces productions par 10 francophones natifs et observent que parmi les 9 apprenants avancés 3 d'entre eux ont été jugés comme natifs. Bongaerts et ses collègues jugent ainsi que la présence d'apprenants tardifs avec un niveau de prononciation équivalent à celui des natifs n'est pas un

phénomène isolé et que cela remet en cause la validité de l’HPC.

L’HPC est également remise en question dans le SLM de Flege (1995) qui considère que, si le début d’apprentissage a un rôle important dans la phonétique L2, tout être humain a accès aux mécanismes d’apprentissage L1 qui lui permettent la création de nouvelles catégories phonologiques.

Si l’HPC et sa généralisation peuvent être remises en question, l’âge d’arrivée ou de début d’apprentissage reste – comme de nombreuses études peuvent en témoigner – le facteur principal en prononciation L2 (Flege et al., 1999 ; Oyama, 1976 ; Tahta et al., 1981). Bien que certaines exceptions existent, les enfants ayant commencé l’apprentissage de la langue avant leurs 6 ans ne conservent que très rarement un accent étranger, tandis qu’à l’inverse les apprenants tardifs ayant débuté leur apprentissage après 15 ans et capables de parler la langue cible avec un accent natif restent de rares exceptions. Cependant, d’autres facteurs se confondent avec celui de l’âge d’arrivée dans le pays étranger ce qui peut remettre en cause son importance, ou du moins en partie (Wode, 1994). Flege et Fletcher (1992) évoquent ainsi la présence de facteurs cognitifs, sociaux et psychologiques.

### 1.6.2 L’input

Selon Flege (2008), l’input joue un rôle primordial dans l’acquisition de la L2 et il réfute le point de vue de DeKayser (2000, cité par Flege 2008) considérant l’input comme n’ayant qu’un rôle « très limité » puisqu’il ne varierait que peu là où l’on observe le plus d’effet de l’âge : dans le domaine phonologique. Pour Flege, l’input que reçoit un apprenant en L2 est bien varié et il n’est pas rare qu’il s’agisse au départ de celui reçu par un professeur ayant lui-même un accent étranger, puis, même après avoir immigré dans un pays où la L2 est parlée, l’apprenant va s’adresser à des locuteurs natifs parlant avec des variétés différentes ou à des locuteurs eux aussi non-natifs et ayant un accent L2. Pour les apprenants tardifs, cet input pourra donc être très différent de celui reçu par les enfants en L1.

Flege (2008) compare le poids de l’input par rapport aux autres facteurs en acquisition de phonologie L2 dans une revue de ses études sur le sujet. Il s’agit de trois études jugeant de l’accent étranger sur des phrases en L2 (Flege et al. 1995, Flege et al. 1999, Piske et al. 2001), une étude d’identification de consonnes plosives (Mackay et al. 2001), une étude de discrimination de voyelles (Flege et MacKay 2004) et une étude de production de voyelles

(Flege et al 2003). Flege fait deux observations principales :

- Les groupes de participants s'étant auto-évalués comme utilisant peu leur L1 montrent de meilleurs scores en L2
- L'âge d'acquisition reste le facteur principal et permet d'expliquer en moyenne 57,4% de la variance tandis que l'usage de la L1 n'en explique que 8,4%.

Cependant, Flege attire l'attention sur le fait que le facteur d'âge d'acquisition englobe différentes variables qui sont le niveau de développement neurologique et cognitif lorsque l'apprentissage a débuté, le niveau de développement des catégories phonologique L1, le type d'input, et même la fréquence d'utilisation de la L2 ou la L1. Il s'agirait donc d'une « macrovariable » qui ne peut être comparée directement à la variable de l'usage de la L1. Il explique ainsi que le pourcentage de variation expliquée par l'usage de la L1 (8,4%) pourrait être plus important qu'il n'y paraît.

Selon Krashen (1985), l'input doit être compréhensible pour l'apprenant tout en étant assez complexe afin qu'il puisse progresser. Le *teacher-talk* serait donc un meilleur input pour les apprenants débutants puisqu'il leur est plus facilement compréhensible que l'input du « monde réel », mais lorsque le niveau d'apprentissage de l'apprenant est plus élevé celui-ci devient moins adapté.

Le *feedback* du professeur joue également un rôle primordial dans l'apprentissage et la correction (ou non) de certaines erreurs. Ainsi, Nystrom (1983 cité par Moyer 2004) explique que : « *some errors are ignored, while others of a similar type receive significant amounts of attention in the teacher's response* » (p.170).

De nombreuses études ont cherché à évaluer l'importance du type d'instruction sur l'accent étranger et l'apprentissage phonologique de la L2. Moyer (2004) observe qu'un apprentissage basé sur la traduction et la grammaire mène à une moins bonne fluence et qu'un apprentissage segmental et suprasegmental est associé à un accent plus proche de celui des natifs (ce n'est pas le cas quand celui-ci n'est que segmental).

Saito (2012) compare différentes études sur l'efficacité de méthodes d'instruction qu'il classe suivant trois types d'instruction :

1. *focus-on-form* (FonF) : où l'enseignement de la prononciation est mis en place à la fois par des exercices contrôlés (par exemple de répétition) mais également à tra-

vers des activités communicatives où l'attention n'est pas uniquement dirigée vers la prononciation mais s'intègre dans les activités (Saito donne l'exemple d'un professeur qui corrige la prononciation du /ɪ/ anglais lors d'une tâche de prise de parole argumentative)

2. *focus-on-formS* (FonFS) : l'enseignement de la prononciation est uniquement mis en place par des exercices contrôlés dont le seul but est d'atteindre une bonne prononciation
3. *focus-on-meaning* (FonM) : l'enseignement se focalise sur des tâches communicatives sans enseignement de la prononciation

Saito observe que les six études associées au type d'instruction FonF montrent une amélioration sur des tâches contrôlées et que deux d'entre elles montrent aussi une amélioration en parole spontanée. Pour les études associées au type d'instruction FonFS, huit études montrent bien une amélioration sur des tâches contrôlées. En revanche, les trois études intégrant une tâche de parole spontanée n'ont pas montré d'amélioration à ce niveau-là. Ainsi, le type d'instruction FonFS, uniquement basé sur des exercices contrôlés, serait moins performant concernant l'amélioration de la prononciation en tâche spontanée. Saito évoque aussi une différence concernant l'importance de l'amélioration mais également la durée des bénéfices acquis par l'instruction entre les deux tâches, observée dans l'étude de Abe (2011, cité par Saito 2012), FonF étant encore une fois plus efficace que FonFS en étant associé à une amélioration plus conséquente et à une plus longue durée des effets de l'instruction.

En prosodie L2, des observations similaires sont faites par McKinnon (2017). Dans son étude, il compare deux types d'instruction sur la production de déclaratives et d'impératives en espagnol L2 : une instruction basée sur la grammaire (qu'il nomme FOG – Focus on Grammar) et un type d'instruction similaire où il ajoute uniquement des explications et un *feedback* prosodique (FOG + I). La méthode d'instruction utilisée dans les deux cas est l'approche communicative à travers un jeu de rôle où les apprenants sont chargés de reproduire une recette de cuisine en alternant ordres et narration des actions face au professeur, puis en binôme. Leur évolution est mesurée à l'aide d'un pré-test et d'un post-test où les apprenants doivent effectuer une tâche de complétion guidée par un contexte et la

présence de la phrase cible écrite. Les deux groupes d'apprenants ne distinguaient pas les déclaratives des impératives dans le pré-test. Les apprenants du groupe ayant reçu un *feedback* prosodique (FOG + I) ont modifié les contours produits ainsi que leur empan de F0 afin de faire une distinction entre les deux types de phrases. Ce n'est pas le cas du groupe n'ayant reçu que des instructions grammaticales (FOG). Cependant, l'auteur observe beaucoup de variation chez les apprenants et certaines modifications du pitch ne sont pas celles produites par l'enseignant. Les apprenants ont donc pris conscience de l'importance de la prosodie pour distinguer ces deux types de phrases, mais n'ont pas la même perception et/ou le même réemploi des indices prosodiques.

### 1.6.3 La motivation

La motivation est un facteur assez complexe qui peut avoir différentes interprétations. Par exemple, lorsqu'on parle de *motivation*, on peut parler d'objectif ou du but que l'apprenant souhaite atteindre. Pour un apprenant vivant dans le pays de sa langue cible, ce peut être la volonté de savoir appréhender les situations quotidiennes et pouvoir communiquer avec les autres habitants ; il privilégiera donc un apprentissage basé sur des situations concrètes et attachera certainement plus d'importance à sa compétence communicative. Pour un étudiant spécialisé dans la littérature de la langue cible, ce ne seront pas l'oral ni la communication qui seront mis en avant mais les compétences écrites, voire uniquement la compréhension écrite. D'autres n'ont pas d'objectif particulier à atteindre et sont simplement guidés par leur intérêt pour la langue.

Mais la motivation c'est également le sentiment qui pousse les apprenants à étudier la langue quelles que soient les raisons d'apprentissage et leurs objectifs. Nous imaginons bien qu'un apprenant démotivé qui se force à apprendre une langue (parce qu'il y est obligé, par exemple en raison de son travail ou de ses études) ne montrera pas la même progression que s'il était *motivé* dans son apprentissage.

Bien que nous ayons évoqué le fait que certains apprenants ne recherchent pas à améliorer les mêmes compétences, il faut également noter que même les apprenants qui ont pour objectif de développer leur compétence orale ne chercheront pas toujours à atteindre le même degré de *bonne prononciation*. En effet, si leur intérêt principal est la communication, il est possible qu'ils ne cherchent qu'à atteindre le seuil minimal leur permettant

d'être compris par les locuteurs natifs avec lesquels ils conversent. D'autres, par intérêt ou dans une optique d'intégration ou d'assimilation de l'identité liée à la langue cible, chercheront à parler avec le moins d'accent possible. Ainsi, Purcell et Suter (1980) montrent dans leur étude que la variable « *degree of concern for pronunciation accuracy* » semble jouer un rôle important dans la perception d'un accent étranger dans la production L2.

Plusieurs études ont cherché à définir l'importance de la motivation dans l'apprentissage.

Bongaerts et al. (2000), dans la quatrième étude portant sur les apprenants très performants ont fait lire dix phrases à deux groupes de locuteurs : un groupe de 10 néerlandophones natifs et un groupe de 30 apprenants de néerlandais de niveau avancé, de différentes L1. 21 néerlandophones natifs ont jugé l'accent des participants. Parmi les 30 apprenants L2, ils notent la présence de 2 apprenants très performants (*highly successful*) dont le niveau d'accent étranger perçu correspond à celui des natifs (ou tout du moins aux locuteurs natifs les moins « performants » ayant obtenu des notes légèrement plus basses). Ils attribuent la réussite de ces apprenants à leur très grande motivation et suggèrent qu'une forte motivation et un entraînement intensif en perception et production des sons de la L2 sont des facteurs de réussite pour atteindre une prononciation native après la période critique.

Flege et al. (1995) a mené une étude de production sur 240 locuteurs anglais L2 vivant au Canada (110 hommes et 130 femmes) afin d'analyser les facteurs associés à l'accent étranger. Son étude consiste en la production de mots et phrases isolées dans une tâche de répétition différée. 24 anglophones natifs ont également été enregistrés et les enregistrements obtenus ont tous été modifiés par un filtre basse-pass puis jugés par 10 anglophones natifs. Les principaux facteurs observés étaient sans surprise liés à l'âge de début d'apprentissage (voir 1.6) et à l'utilisation de la L2 (voir 1.2). Cependant, il a également pu observer une influence du facteur de motivation, notamment dans le groupe de locuteurs L2 hommes. En effet, chez ces locuteurs, la motivation dite « *integrative* » (l'envie d'apprendre l'anglais pour des raisons sociales ou culturelles) semble jouer un rôle positif sur le niveau d'accent étranger.

Moyer (2004) a mené une étude sur 25 immigrants vivant à Berlin et âgés de 25 à 35 ans et ayant un niveau d'éducation élevé puisqu'ils ont tous obtenu l'équivalent allemand

du baccalauréat. Moyer leur a fait compléter un questionnaire avec de nombreuses questions sur leur rapport à la langue allemande, leur motivation, le type d'instruction reçue, la fréquence des contacts avec les natifs, etc. Ensuite, ils ont participé à une expérience de production (parole contrôlée et semi-contrôlée ainsi qu'une interview) et 3 germanophones natifs ont évalué leur accent. Les résultats montrent un effet de la motivation sur le degré d'accent étranger, notamment la corrélation positive entre la faiblesse de l'accent étranger perçu et la constance de la force de la motivation, la motivation personnelle ainsi que l'intensité de la motivation. La motivation *uniquement* professionnelle (c'est-à-dire sans être également associée à une motivation personnelle) n'a pas montré d'effet significatif. Les résultats montrent également une corrélation entre la motivation et les contacts avec des locuteurs natifs ; des contacts peu fréquents sont associés à une motivation basse, tandis que des contacts fréquents sont associés à une forte motivation et une perception positive de la justesse de son propre accent. Moyer évoque cependant un contre-exemple notable : une locutrice L2 d'origine polonaise ayant obtenu des résultats semblables à ceux des natifs et n'ayant que pourtant peu de contacts avec les locuteurs natifs. En revanche, elle a une très forte motivation et a une quantité importante d'input passif (lecture, visionnage de films allemands...). Elle montre également un fort sentiment d'identité allemande bien qu'elle ne rejette pas son identité polonaise. Moyer attribue donc sa réussite exceptionnelle à un *engagement optimal* qui se manifeste par une recherche active d'input et de *feedback*.

### 1.6.4 Les capacités individuelles

Nous avons vu précédemment que certains apprenants avaient plus de facilités à atteindre une prononciation comparable à celle des locuteurs natifs, malgré un apprentissage tardif (après l'âge de 15 ans). Ces apprenants sont jugés comme « exceptionnels » et ne représentent qu'une petite minorité de personnes. Si la motivation, liée à une recherche active d'input, semble être un facteur expliquant ce phénomène (voir 1.6.2), d'autres études mettent en avant l'importance des capacités individuelles dans l'acquisition de la L2 et de sa prononciation.

Dans leurs études, Purcell et Suter (1980) ont évalué l'importance de 12 facteurs sur la prononciation de l'anglais, jugés significatifs dans une première étude intégrant un total de

20 facteurs. Les deux études se sont basées sur les données de 61 apprenants de différentes L1 (japonais, thaï, arabe, perse). Les participants ont passé différentes expériences et rempli un questionnaire afin d'obtenir les différentes variables. La capacité à imiter les sons (*aptitude for oral mimicry*) a été définie par une expérience où les participants devaient répéter plusieurs fois des sons non-natifs après l'expérimentateur. La justesse de ces réalisations ainsi que le nombre de répétitions avant d'atteindre une réalisation correcte ont été pris en compte pour l'évaluation de cette capacité. Les participants ont ensuite été enregistrés dans une tâche de parole spontanée où ils devaient décrire une fête ou un jour férié existant dans leur pays natal. Le modèle statistique avec trois variables retenues comporte les variables suivantes :

1. La langue (L1) de l'apprenant
2. La capacité à imiter les sons (oral mimicry)
3. La durée de résidence dans un pays utilisant la langue cible et/ou avec une personne native dans cette langue

Un deuxième modèle ajoutant une quatrième variable est également retenu. Cette quatrième variable est le degré de préoccupation pour la prononciation (*strenght of concern for pronunciation*). La capacité des locuteurs à imiter les sons semble donc jouer un rôle essentiel dans leur production de la langue cible puisqu'elle fait partie des trois principales variables expliquant les résultats obtenus.

Kendrick (1997) a fait évaluer les productions de 8 apprenants de l'anglais de différentes L1 (japonais, russe, coréen, taïwanais, thaïlandais) par 20 personnes. Ils étaient enregistrés dans plusieurs tâches allant de la parole spontanée à la lecture de textes et d'une liste de mots. Kendrick a également cherché à évaluer différents facteurs tels que la motivation, l'utilisation de l'anglais, et les capacités linguistiques. Ces dernières ont été évaluées par des tests (de perception, rythme...), incluant deux tests de capacité à imiter les sons :

- Un test de répétition de *chunks* en anglais de plus en plus longs
- Un test de répétition de phrases en non-mots

La capacité des apprenants à imiter est un facteur significatif et les apprenants ayant une meilleure capacité à imiter sont également ceux qui ont la meilleure prononciation en anglais.

Granena et Long (2013) ont cherché à évaluer quatre facteurs associés au degré d'accent étranger perçu dans les productions L2 : l'âge de début d'apprentissage, la durée de résidence, la capacité linguistique (*language aptitude*). Pour ce faire, ils ont enregistré 61 apprenants sinophones à qui ils ont demandé de lire trois lignes de texte en espagnol incluant des sons habituellement difficiles pour les sinophones. Ces enregistrements ont été jugés par 12 juges natifs. La capacité linguistique des apprenants a été mesurée à l'aide du test LLAMA comportant différentes tâches dans une langue inventée : d'acquisition de vocabulaire, de grammaire, de reconnaissance de sons et de correspondance entre des sons et leur graphie. Granena et Long observent dans leur étude une corrélation significative entre la capacité linguistique des apprenants et leur prononciation dans le groupe d'apprenants ayant commencé leur apprentissage entre 16 et 26 ans. Les auteurs relativisent cependant ces résultats et supposent que cette corrélation ne serait pas observable en parole spontanée.

Ainsi, différentes études montrent une corrélation plus ou moins conséquente entre les capacités individuelles (capacité à reproduire des sons, à apprendre une langue inconnue) et le niveau de prononciation atteint par les apprenants. Cependant, l'effet de facteur reste limité comparé à celui d'autres facteurs tels que l'âge de début d'apprentissage ou la durée de résidence.

### 1.6.5 Le type de parole

A la différence des facteurs évoqués précédemment, le type de parole n'est pas un facteur de variation lié à l'apprentissage en tant que tel (c'est-à-dire concernant l'apprenant ou les conditions de son apprentissage) mais permet d'expliquer certaines différences observées dans les productions des apprenants. Il peut également être lié aux précédents facteurs notamment la motivation ou l'input. En effet, certains locuteurs seront plus à l'aise dans certaines tâches que d'autres (lecture de textes, parole spontanée...) en raison de la nature de la tâche mais aussi des autres facteurs évoqués. La plupart des études s'intéressant à l'influence du type de tâche sur les productions L2 oppose les tâches de lecture à

voix haute aux tâches de parole spontanée. Mais les résultats de ces études sont parfois contradictoires.

Selon Tarone (1982), il existe un phénomène de *style-shifting* et dans certains styles de parole l'attention portée à celle-ci est plus ou moins importante. Elle cite comme exemple Dickerson et Dickerson (1977) et son étude sur la prononciation du /r/ anglais par des apprenants japonophones dans trois types de tâches (parole libre, lecture de dialogue et lecture d'une liste de mots) qui a montré que le /r/ a correctement été produit dans presque 100% des cas lors de la lecture de mots isolés tandis que dans seulement 50% des cas en parole libre. Ainsi, l'attention à la forme serait plus importante en lecture de mots isolés et permettrait une meilleure production de la L2.

Shizuka (1992) observe des résultats similaires dans son étude sur la production du /r/ anglais chez 60 apprenants japonophones. Les participants ont été enregistrés dans trois tâches différentes : environ deux minutes de parole libre (se présenter), la lecture d'un texte puis la lecture d'une liste de mots isolés. Shizuka observe de meilleurs résultats avec la tâche de lecture de mots isolés que dans les deux autres tâches. Il note cependant que 12 des 60 apprenants montrent de meilleurs résultats en parole libre qu'en lecture de texte. Selon lui, ce résultat s'expliquerait par les mots choisis par l'apprenant en parole libre, plus faciles et fréquents que ceux présents dans le texte.

Oyama (1976) fait des observations différentes dans son étude. Elle a enregistré 60 immigrants italiens aux États-Unis ayant appris l'anglais après leur arrivée. Les participants ont réalisé deux tâches de production : une tâche de lecture de textes à voix haute ainsi qu'une courte tâche de parole spontanée où ils devaient décrire un moment qui les a effrayés dans leur vie. Le degré d'accent étranger a ensuite été jugé par deux locuteurs natifs. Oyama observe de meilleurs résultats pour la tâche de parole spontanée que pour la lecture de textes. Elle attribue cette différence inattendue au fait que les participants ressentiraient plus d'anxiété en lisant les textes en étant conscients de leur accent étranger et de leurs difficultés que lorsqu'ils parlent librement en parole spontanée.

Díaz-Campos (2004) s'intéresse également à l'influence du type de parole en production L2. Il a analysé les productions de 60 apprenants anglophones étudiant l'espagnol (26 apprenants en échange universitaire en Espagne et 20 apprenants aux États-Unis) de niveau intermédiaire dans deux tâches : une tâche de lecture à voix haute et une tâche de

production orale (interview). Ses analyses portent sur la production des consonnes occlusives non-voisées à l'initiale du mot, les consonnes fricatives voisées intervocaliques, les consonnes latérales en fin de syllabe ainsi que les consonnes nasales palatales. Ses résultats diffèrent selon les phonèmes observés mais il constate une plus grande quantité de productions correctes dans la tâche spontanée que dans la tâche de lecture de textes.

Ainsi, l'effet du type de tâche sur la production L2 n'est pas clair; certaines études observent une meilleure production dans les tâches contrôlées (lecture de textes, de mots) tandis que d'autres, à l'inverse, montrent des résultats plus proches de ceux des natifs dans les tâches plus spontanées (interview, raconter un épisode de sa vie).

## 1.7 Synthèse

Dans ce chapitre, nous avons mis en lumière la complexité du processus d'acquisition de la L2, et plus particulièrement de l'acquisition phonétique et phonologique. Les difficultés et facilités rencontrées par les apprenants peuvent s'expliquer par une multitude de facteurs tels que leur langue première, l'âge de début d'apprentissage, l'input auquel ils sont exposés, voire des facteurs personnels comme leur motivation ou leurs capacités.

Nous avons vu également que si le lien entre la perception et la production semble relativement fort, il arrive que ces capacités évoluent différemment et que les compétences perceptives ne soient pas alignées avec la production. Si certains parlent d'un décalage entre les deux compétences et expliquent que la perception précède la production, d'autres études montrent qu'une amélioration en production n'est pas toujours associée à une amélioration perceptive. Ces deux systèmes, bien qu'étant liés, doivent donc être pris séparément. Ainsi, des conclusions tirées d'une analyse perceptive ne sont pas directement applicables à la production, et inversement. Parmi les facteurs jouant un rôle dans l'acquisition L2 que nous avons détaillés dans ce chapitre, la L1 de l'apprenant reste l'un des facteurs les plus importants qui permet d'expliquer une grande partie des erreurs *phonétiques* et *phonologiques* des apprenants. Cette influence de la L1 sur la L2 c'est ce qu'on appelle le phénomène de **transfert** et qui a été fortement remis en question dans les autres domaines (acquisition de la syntaxe, du lexique...) mais qui est toujours d'actualité en phonétique et phonologie L2. L'hypothèse de l'interlangue insiste cependant sur la présence de stratégies chez les apprenants (comme la surgénéralisation) qui ne peuvent s'expliquer par ce

phénomène. L'interlangue est la langue transitoire de l'apprenant où l'on peut donc observer l'influence de la langue première voire des autres langues en cours d'acquisition, mais également toutes les stratégies mises en place par l'apprenant pour s'approprier la L2. Celle-ci évolue au cours de l'apprentissage dans la direction de la langue cible. Ainsi, si le transfert de la langue première ne permet pas systématiquement d'expliquer les erreurs observées en L2, son importance reste difficilement contestable. C'est pourquoi, dans le prochain chapitre, nous allons détailler les systèmes prosodiques du japonais et du français, afin de pouvoir comparer les deux langues et anticiper les difficultés que rencontreront les apprenants japonophones dans leur production et perception des continuatifs.

## Chapitre 2

# Systemes prosodiques du français et du japonais

### Introduction

Lors de l'apprentissage de la prosodie la langue maternelle de l'apprenant influence la production et la perception des segments non-natifs, bien que cela n'explique qu'en partie l'interlangue des apprenants comme nous l'avons vu dans le chapitre précédent. Toujours est-il que cette influence de la langue maternelle peut nous permettre d'expliquer certains résultats et c'est pour cela que dans ce chapitre nous allons comparer les systèmes prosodiques du français et du japonais. Dans un premier temps, nous allons nous intéresser à la hiérarchie prosodique des deux langues ainsi que leurs différents constituants et les réalisations prosodiques qui leur sont associées, dont le contour continuatif pour le français, et dans un deuxième temps nous allons décrire le rôle discursif et interactionnel des contours montants dans ces deux langues.

### 2.1 Accentuation et rythme des deux langues

La première différence entre le japonais et le français se situe au niveau rythmique. En effet, tandis que le japonais est généralement considéré comme une langue moraïque (*mora-timed*, Labruno 2006 et Ladefoged 1975), le français serait une langue dite syllabique (*syllable-timed*, Pike 1945). A la différence du système syllabique, la more en japonais peut

correspondre à la première partie d'une gémignée (/Q/), la seconde partie d'une voyelle longue (/R/) ou bien la nasale /N/. Le mot *hontō* (/hoNtoR/, « vraiment ») comporte ainsi quatre mores, alors qu'il ne compterait que deux syllabes. Il faut cependant être prudent sur ces définitions qui sont relativement contestées. Tout d'abord, pour le japonais, de nombreux travaux réfutent cette classification et considèrent que le japonais est une langue syllabique (voir Kawahara (2016) pour une revue des arguments en faveur de la présence de syllabes en japonais). Malgré tout, des études montrent que les japonais perçoivent la more en tant que plus petite unité, notamment les travaux de Otake et al. (1993) montrant que les japonophones réagissent plus vite pour l'identification des mores que des syllabes. En effet, leur temps de réaction pour identifier la première syllabe d'un stimulus (CVC, par exemple /taN/ dans *tanshi*) est significativement plus long que pour l'identification de sa première more (CV /ta/). Nous choisissons donc dans cette thèse de conserver cette appellation de *more*, qui reste très généralement répandue dans la littérature récente de la phonologie japonaise, malgré les débats au sujet de cette notion. Concernant le français, sa potentielle isochronie syllabique est remise en question, tout comme de ce fait l'appellation de « langue syllabique ». Pour Wenk et Wioland (1982), l'idée d'isochronie syllabique en français n'est pas cohérente puisque le système prosodique du français implique la présence de syllabes accentuées, associées à un allongement. Elle cite également Delattre (1966b, p.190) qui note que : « *it is French which of the four languages [English, German, Spanish and French], shows by far the widest ratio of lengths from stressed to unstressed syllables.* » Quant à la durée de l'intervalle entre deux accents, Fant et al. (1991) ont montré que celle-ci était équivalente en suédois, anglais et français. L'isochronie moraique du japonais est moins fortement rejetée que l'isochronie syllabique du français, mais il est intéressant de noter qu'elle n'est néanmoins pas toujours attestée (Arai & Greenberg, 1997 ; Beckman, 1982)

Si le point évoqué précédemment et la dichotomie langue syllabique / langue moraique ne font pas consensus, les différences accentuelles entre le japonais et le français sont quant à elles indiscutables. En effet, le système prosodique du japonais diffère grandement de celui du français par la présence d'accent lexical. Il existe de nombreuses paires minimales où seul l'accent permet d'indiquer quel est le mot auquel se réfère le locuteur, par exemple dans la paire minimale *Ame* (HL, « pluie ») / *aME* (LH, « bonbon »). Cet accent lexical est

principalement réalisé par une différence de F<sub>0</sub>, et non de durée ou d'intensité, à la différence des langues à *stress* comme l'anglais. Le japonais est donc considéré comme une langue à hauteur mélodique (*pitch accent language*). Cet accent lexical est sujet à la variation à travers les différents dialectes ou accents régionaux japonais. Par exemple, la paire minimale évoquée précédemment (**A**me / a**ME**) est inversée dans le dialecte de Kagoshima où **A**me signifie « bonbon » et a**ME** « pluie » (Kubozono, 2014). Mais les différences prosodiques entre les dialectes japonais ne se résument pas à un accent lexical et des paires minimales différentes, mais parfois à tout un autre système intonatif. Dans cette thèse, lorsque nous évoquons le système intonatif du japonais, nous nous référons uniquement au dialecte dit « standard » de la région de Tokyo, le *hyōjungo*. En japonais standard, la more accentuée est associée à une chute mélodique (de haut à bas) qui la suit. Certains mots dit « atones » (ou non-accentués) n'ont pas de more accentuée. On peut observer différents schémas accentuels :

1. accent initial, (*atamadaka-gata*), exemple : **NA**mida « larme »
2. accent interne, *nakadaka-gata*, exemple : ko**KO**ro « coeur »
3. accent final, *odaka-gata*, exemple : kaga**MI** « miroir »
4. atone, exemple : sakura<sup>o</sup> « cerisier »

(exemples issus de Labrune 2006)

Les emprunts étrangers suivent quant à eux l'accentuation « par défaut » avec l'accent produit sur la more antépénultième, ou pré-antépénultième lorsque la more antépénultième est une more spéciale (/N, /Q/, /R/). La dernière syllabe des mots atones est généralement prononcée dans un registre haut ce qui rend sa réalisation proche du schéma d'accentuation finale. La principale différence entre un mot à accent final et un mot atone s'observe lorsqu'il est suivi d'une particule (par exemple *ga*). En effet, l'ajout de la particule permet d'observer la chute mélodique suivant la more accentuée, avec la particule produite dans un registre plus bas. Dans le cas d'un mot atone, on n'observe pas de forte chute mélodique et la particule est produite dans un registre de hauteur proche de celui de la dernière more du mot atone. Dans l'annotation J\_ToBI (Venditti, 2005), la particule suivant le mot

est associée à un ton bas dans les deux cas (que le mot soit atone ou accentué), la principale différence étant que ce ton bas est précédé par la chute mélodique de l'accent – notée H\*L – lorsque le mot est accentué (Venditti, 2005). Au niveau phonétique, une différence au niveau de la courbe mélodique peut être observée entre un mot atone et mot avec accent final : la montée mélodique sur la dernière more est plus importante pour les mots accentués que les mots atones.

A l'inverse du japonais, le français ne possède pas d'accent lexicalement distinctif. Il est même parfois considéré comme une langue « sans accent » (Rossi & Di Cristo, 1980). D'autres préfèrent utiliser le terme d'« accent fixe » (Martinet, 1968) puisque celui-ci est principalement situé à la fin de mot, sur la dernière syllabe pleine et donc qui n'est pas porteuse de schwa final. Ce refus de considérer l'accent lexical en français s'explique par le fait que celui-ci s'amalgame avec l'intonation aux frontières des niveaux prosodiques supérieurs. Ainsi, selon Rossi (1980) : « le français est une langue sans accent, en ce sens que l'accent et l'intonation ne constituent, ni par leur nature, ni par leur fonction, deux unités distinctes. » (p.39).

Vaissière (2010), quant à elle, préfère décrire le français comme une langue à « accent de frontière » puisque ces prééminences sont généralement associées aux frontières prosodiques, et c'est ce lien entre accentuation et phrasé que nous allons détailler dans la prochaine section.

Si l'accent du français est souvent réduit à cette accentuation finale, il existe aussi une **accentuation initiale** parfois nommée accent secondaire, d'insistance, d'emphase, ou même ictus mélodique (Rossi, 1999). Cependant, si ces termes ont été utilisés indifféremment dans un premier temps, il est maintenant admis que l'on distingue au minimum deux types d'accentuation initiale : l'accent initial emphatique et l'accent initial non-emphatique. L'accent initial emphatique est celui de l'instance, parfois associé à la focalisation mais qui peut également avoir une fonction expressive. L'accent non-emphatique, quant à lui, a un rôle rythmique. Associé à l'accent final, on lui attribue une fonction démarcative (Astésano & Bertrand, 2016; Jun & Fougeron, 2000); l'accent final marque la frontière droite d'un constituant tandis que l'accent initial est à la frontière gauche de celui-ci. En effet, l'accent initial est généralement produit sur la première syllabe accentuable (sur le premier mot lexical) d'un groupe prosodique. C'est le cas dans l'exemple

suivant :

### 1. le **PRE**sident de la répu**BLI**QUE

Ce phénomène particulier est également connu sous le nom d'arc accentuel (Fónagy, 1980). Lorsque le syntagme est relativement court, il est fréquent que la présence d'accent initial soit une manière d'éviter la collision accentuelle, c'est-à-dire la succession de deux syllabes accentuées (Di Cristo, 2016). Ainsi, « une che**MISE NOIRE** » serait produit avec un déplacement de l'accent final de « chem**ise** » à l'initiale du mot : « une **CHE**mise **NOIRE** ». Cependant, selon Martin (2018) ce recul d'accent en cas de collision accentuelle n'est pas la norme et il observe que lorsque l'intervalle entre les deux syllabes accentuables est supérieur à 250 ms elles sont toutes les deux accentuées. Dans le cas où cet intervalle est inférieur à 250 ms, il observe soit :

- un recul de l'accent avec restructuration (arc accentuel) ou sur une autre syllabe
- un effacement du premier accent

L'accent final en français est principalement caractérisé par l'allongement de la syllabe accentuée. Selon Delattre, « si l'on veut évaluer l'accent français par l'un de ses éléments acoustiques, c'est la durée qu'il faut choisir » (1966b, p.70).

L'accent final ne semble cependant pas être associé à une quelconque variation d'intensité (Delattre, 1938). Comme pour l'accent initial, l'accent final est généralement réalisé par une prééminence mélodique. Ces deux types d'accent se différencient cependant par le paramètre de durée : l'accent initial, contrairement à l'accent final, n'est pas associé à un allongement syllabique (Jun & Fougeron, 2000 ; Rossi, 1999).

En français, le statut phonétique et phonologique des items lexicaux ne fait pas consensus. En effet, beaucoup s'accordent à dire qu'ils n'ont pas de véritable statut phonologique puisque c'est en premier lieu le groupe accentuel qui porte l'accentuation finale et que les accents des items lexicaux peuvent ne pas être réalisés selon leur place dans l'organisation prosodique. Cependant, certains chercheurs considèrent que ces accents lexicaux ne disparaissent pas totalement dans la parole et qu'il ne s'agit que d'une « désaccentuation incomplète » (Delattre, 1966a). Ce principe est également présent dans l'hypothèse de Dell : « En français, chaque mot porte un accent dont la force est proportionnelle à l'importance

de la coupe syntaxique qui suit ce mot » (Dell, 1984, p.77). Dans leur analyse de 13 heures de parole journalistique, Adda-Decker et al. (2009) montrent que les mots lexicaux sont effectivement prononcés avec une hauteur mélodique plus élevée sur leur dernière syllabe forte. Ainsi, les mots lexicaux seraient donc associés à une accentuation mélodique finale.

Certains mots du français sont considérés comme n'étant pas accentuables, ou *accentogènes* et c'est notamment le cas des clitiques : les pronoms (je, me, tu, te, nous, vous, il(s), elle(s), se, le, la, les, lui, leur, y, en), le morphème de négation « ne » et les déterminants monosyllabiques (un, des, le, les, ce, etc.). Les prépositions monosyllabiques (de, à, en) ainsi que les adverbes monosyllabiques sont aussi parfois considérés comme clitiques. Les pronoms toniques (toi, moi, eux, etc.) sont quant à eux généralement classés comme non clitiques. Mertens (1993) distingue les mots clitiques des mots non clitiques en leur attribuant la propriété [ $\pm$ clitique]. Il décrit l'unité accentuelle comme un mot non clitique associé aux clitiques qui s'y rattachent. Selon lui, l'accentuation n'est cependant pas liée à la propriété non clitique puisqu'un mot clitique peut être accentué lorsqu'il se situe à la fin d'une unité accentuelle. C'est le cas dans l'exemple « demande-le-lui », où l'accent final se trouve sur le clitique « lui ». L'accent initial peut également être produit sur un clitique.

(47)	<b>il le lui donne</b>
<b>AL</b>	. . . o
<b>GA</b>	. . . -
<b>GI</b>	b.....b HB-
(48)	<b>il a balbutié une excuse</b>
<b>AL</b>	. o . . o . . o
<b>GA</b>	. - . . - . . -
<b>GI</b>	bb BB b b HH b b B-B-

FIGURE 2.1 – Exemples issus de Mertens (1993); les syllabes accentogènes sont représentées par 'o' tandis que les syllabes non-accentogènes, dont celles des clitiques, par ''

Pour conclure, si les différences rythmiques entre le japonais et le français (langue syllabique VS moraïque) ne sont pas clairement attestées, il ne fait pas de doute que le système accentuel est très différent. Il apparaît notamment que les corrélats acoustiques associés à cette accentuation diffèrent entre les deux langues (hauteur mélodique uniquement pour le japonais et hauteur mélodique et durée pour le français). De plus, l'accentuation du français est principalement finale (avec toutefois la présence facultative d'un accent ini-

tial) tandis qu'en japonais l'accent lexical se manifeste par une chute mélodique interne au groupe accentuel. Ces différences au niveau de l'accentuation nous font supposer des difficultés chez les apprenants japonophones à réaliser l'accent final français. Dans la section suivante, nous verrons que, au-delà de l'accentuation, les différences entre le français et le japonais se manifestent également dans leur organisation prosodique.

## 2.2 Phrasé et hiérarchie prosodique

### 2.2.1 Le système prosodique du français

Pour saisir la problématique des contours continuatifs français chez les apprenants japonophones, il est important de comparer les systèmes intonatifs des deux langues – le français et le japonais. Nous allons donc à présent faire l'état des différentes descriptions prosodiques du français et du japonais, et de la place de ce que nous appelons contours continuatifs au cœur de ces différentes théories.

#### **Différentes descriptions prosodiques**

Il existe de nombreuses descriptions du système prosodique du français dans la littérature, et le nombre de niveaux de phrasé prosodiques varie de l'une à l'autre. Néanmoins, deux niveaux sont généralement admis. Un niveau de syntagme accentuel, qui correspond généralement à une tête lexicale et sa projection maximale (Delais-Roussarie & Post, 2008). Le deuxième niveau de phrasé que l'on retrouve dans la majorité des descriptions prosodiques du français est le syntagme intonatif. Un syntagme intonatif comporte généralement plusieurs syntagmes accentuels et est caractérisé par un allongement de la syllabe accentuée plus conséquent et (lorsqu'il est montant) d'une montée mélodique plus importante en comparaison avec le syntagme accentuel. Il est aussi souvent suivi d'une pause. Cependant, il reste, selon Delais-Roussarie et Post, défini de façon relativement « floue » dans les différentes descriptions.

Selon Jun et Fougeron (2000), le schéma prototypique de l'AP est /LHiLH\*/, comportant l'accent initial (Hi) et l'accent final (H\*). Le premier ton L, quant à lui n'est pas produit lorsque l'accent initial est sur la première syllabe, et cet accent est lui-même facultatif. Les réalisations de surface qu'ils observent, qui ne correspondent ainsi pas toujours à ce

schéma, sont présentées en Figure 2.2. La plupart des schéma accentuels du AP terminent par un ton H\*, à l'exception du dernier schéma terminant par L\*. Selon Jun et Fougeron, ces schémas peuvent principalement s'observer lorsque le AP est en position finale d'un IP porteur d'un ton de frontière L%, et que la frontière de AP coïncide donc avec la frontière de ce IP :

The last pattern [LHiL] is the most common pattern when an AP is final in an Intonation Phrase whose boundary tone is L%. Since the AP final syllable is also the Intonation Phrase final syllable, the AP final H\* and Intonation Phrase final L% are supposed to be realized on the same syllable. In this case, the AP final H\* is preempted by the higher level (Intonation Phrase) boundary tone. (p.216)

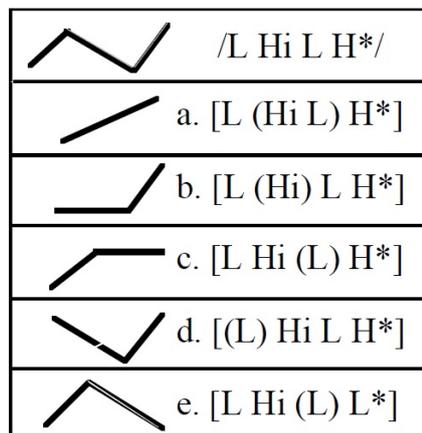


FIGURE 2.2 – Différents schémas accentuels du AP selon Jun et Fougeron (2000)

Le syntagme accentuel décrit par Jun et Fougeron serait comparable au Groupe Intonatif (GI) de Mertens (Mertens, 1993), puisqu'il correspond au domaine comportant à la fois l'accent primaire et l'accent secondaire. En effet, le Groupe Accentuel (GA) décrit par Mertens est une unité plus petite que le syntagme accentuel de Jun et Fougeron (2000) qui correspond à chaque unité regroupant une syllabe accentogène (non-clitique), que celle-ci soit accentuée ou non. Mertens affirme donc qu'« il y aura autant de ces groupes dans l'énoncé qu'il compte de syllabes non clitiques. La syllabe finale de GA sera accentuable, par où on entend qu'elle peut recevoir un ton du paradigme de l'accent final ».

Di Cristo (2016) utilise également une unité inférieure au syntagme accentuel dans son modèle : l'unité tonale (UT). Cette unité prosodique correspond à toute syllabe accentuée, qu'elle soit associée à un accent initial ou final. Dans son modèle, Di Cristo distingue ainsi trois niveaux prosodiques : l'unité tonale (UT), le syntagme prosodique (SP) dont l'accent final se distingue de celui de l'UT par la présence d'un allongement, et l'unité intonative (UI). Il propose toutefois également des niveaux supérieurs à l'UI pour l'analyse de phénomènes dépassant le cadre d'analyse de la phrase traditionnelle : la Macro Unité Intonative (M.UI), le Paraton et le Paragraphe prosodique.

Un niveau prosodique supérieur au syntagme accentuel mais inférieur au syntagme intonation – le syntagme intermédiaire – est ajouté par Jun et Fougeron (2000) à leur description. Ce niveau intermédiaire sera également repris dans le F\_ToBI (Delais-Roussarie et al., 2015).

Ainsi le nombre et la définition des constituants prosodiques varient grandement selon les auteurs, bien que l'on puisse observer des parallèles entre des constituants de nom différent à travers les études. Dans le schéma présenté ci-dessous (2.3), Di Cristo (2016) montre les équivalences des termes et expressions employés dans les travaux sur le phrasé prosodique du français.

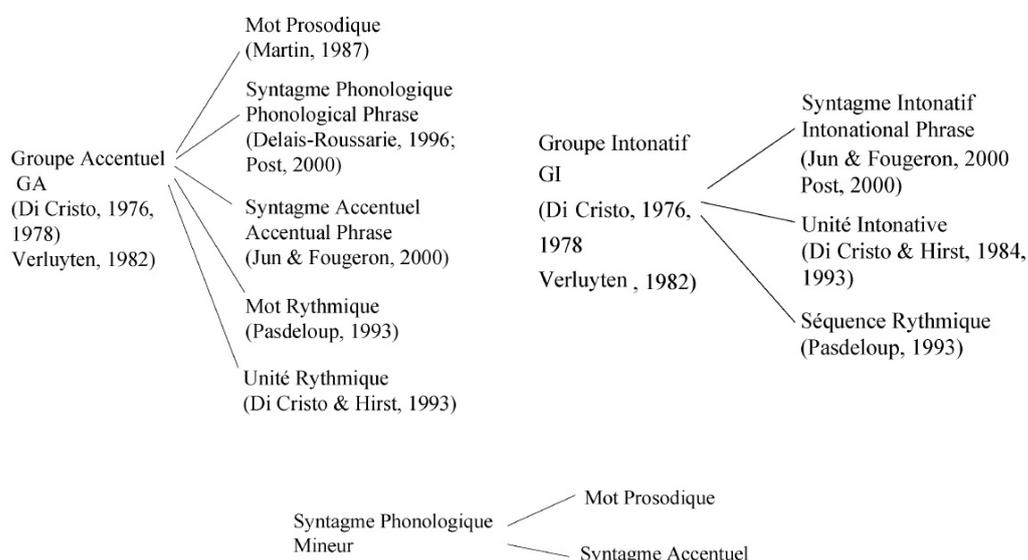


FIGURE 2.3 – Equivalences des termes et expressions employés dans les travaux sur le phrasé prosodique du français par Di Cristo (2016)

Certains auteurs ont une approche phonologique et décrivent ainsi des phénomènes in-

tonatifs qui sont leur base théorique. D'autres adoptent une approche morpho-syntaxique en mettant en avant les liens existant entre les structures syntaxique et prosodique. Ces différentes approches se retrouvent dans la manière de représenter le système prosodique du français. Encore une fois, il existe une grande variété de représentations qui sont intrinsèquement liées à l'approche théorique des auteurs.

### ***Delattre***

Delattre (1938, 1939, 1966b) est l'un des premiers à avoir proposé une description de l'intonation du français. Il présente dans ses travaux un inventaire de dix patrons intonatifs, « intonèmes » (voir 2.4). Ces intonèmes s'opposent phonologiquement, permettant de créer des « paires minimales », soit des énoncés produits par des contours différents qui n'ont ainsi pas le même sens. Ces intonèmes se distinguent les uns des autres par leur niveau de hauteur (allant de 1 à 4) et la forme de leur courbe. Bien que des études plus récentes ont depuis nuancé ces travaux, ils restent, même aujourd'hui, une référence pour de nombreuses études en prosodie du français.

### ***Arbres et grilles métriques***

La structure prosodique peut être représentée par une grille métrique (Prince 1983) ou un arbre métrique (Lieberman et Prince). Les grilles métriques se focalisent sur les prééminences accentuelles, qui sont représentées par un nombre de croix (x) plus ou moins important selon la force de la prééminence. Les arbres métriques, quant à eux, rappellent les arbres syntaxiques en rendant visible la hiérarchie prosodique et les niveaux de prééminence par différents branchements. Un niveau fort (« *strong* », « s ») et un niveau faible (« *weak* », « w ») sont distingués. Des exemples en français de grille et d'arbre métriques sont présentés en Figures 2.6 et 2.5.

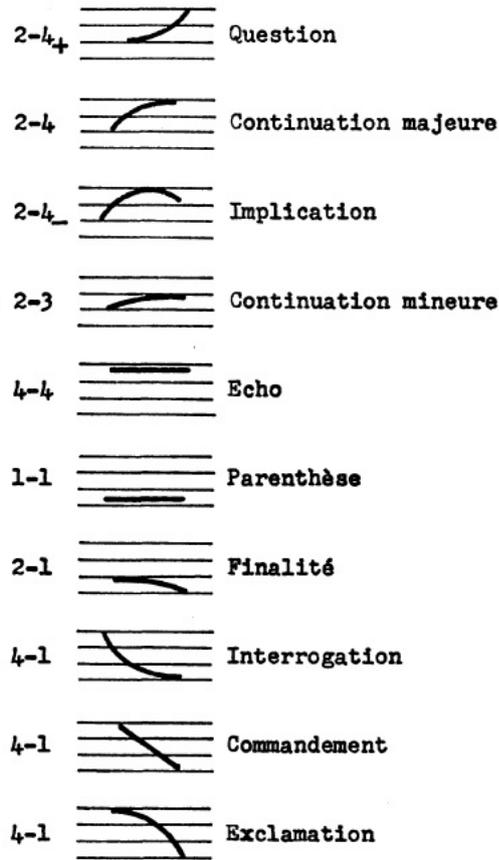


FIGURE 2.4 – Les 10 intonations par Delattre (1966b)

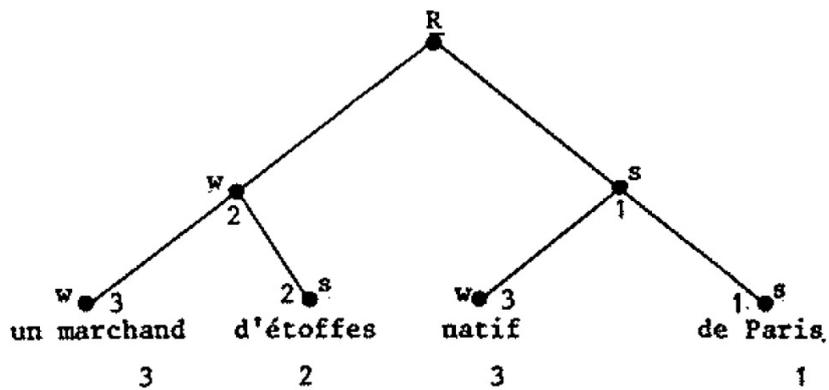


FIGURE 2.5 – Exemple d'arbre métrique – Dell (1984)

	*							*	
	*			*				*	
	*			*		*		*	
*	*	*	*	*	*	*	*	*	<*>
le	kUR	RƏ	pRã	dRõ	lə	ɥi	sep	tã	<brƏ>

FIGURE 2.6 – Exemple de grille métrique – Delais-Roussarie (2000) « Les cours reprendront le huit septembre »

**Contours de Philippe Martin**

La particularité du système de contours proposé par Martin est que celui-ci se réfère aux variations observées sur la voyelle accentuée (en fin de groupe accentuel) et non sur l'ensemble du groupe prosodique, contrairement au système proposé par Delattre. Martin désigne par C0 le contour terminal conclusif (C0 déclaratif et sa variante Ci interrogative), C1 le contour de continuation majeure, C2 le contour de continuation mineure et Cn le contour neutralisé. Il met en avant les dépendances qui existent entre les différents contours. Le contour neutralisé ne peut exister que si un contour continuatif mineur ou majeur ou un contour terminal est planifié après celui-ci. Un contour continuatif majeur, quant à lui, doit être précédé par un contour continuatif mineur.

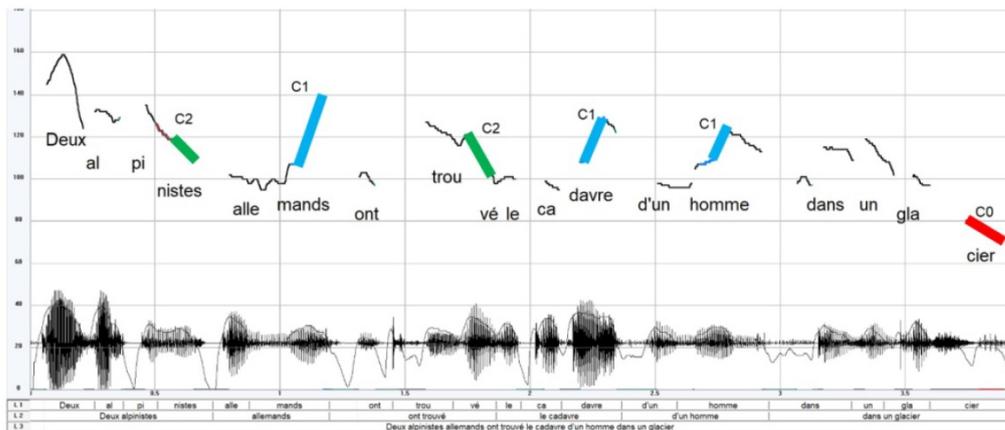


FIGURE 2.7 – Exemple d'annotation par Martin (2018)

**Mertens**

Le système de représentation prosodique de Mertens (1993) est basé sur quatre niveaux de hauteur : le bas, le haut, l'infra-bas et le suraigu, notés par les symboles h, b, b-, h+. Ce sont des niveaux de hauteur *relatifs* puisqu'ils représentent non pas des valeurs absolues mais des changements mélodiques que l'on peut observer d'une syllabe à une autre. Les syllabes accentuables sont, quant à elles, associées à un redoublement du ton pour représenter l'allongement vocalique : HH, BB, B-B-, H+H+. Enfin, les « petits intervalles mélodiques » de 2 à 3 demi-tons et qui n'indiquent pas de réel changement de niveau de hauteur sont représentés par les signes / et \ (signalant respectivement un rehaussement et un abaissement mélodiques).

```
(a) un professeur de français canadien
    b.....b /BB \b...b HH b.b B-B-
(b) un professeur de français canadien
    b.....b HH b...b /BB b.b B-B-
```

FIGURE 2.8 – Exemple d'annotation par Mertens 1993

Dans le cas d'exemples réels Mertens associe cette annotation à la courbe de F0 réalisée. Plus récemment, Mertens a choisi d'utiliser un système de transcription semi-automatique, Prosogramme (Mertens, 2005). Ce logiciel permet la visualisation de la durée des syllabes et des voyelles, et stylise les courbes de F0 en prenant en compte le seuil de glissando. Ainsi, les changements de pente mélodique inférieurs à ce seuil sont considérés comme n'étant pas percevables par l'être humain et sont représentés par une ligne horizontale. Les changements mélodiques supérieurs à celui-ci sont quant à eux représentés par des lignes inclinées, pouvant former des courbes lorsque plusieurs changements mélodiques s'observent sur un même noyau syllabique. Il est également possible d'associer une transcription phonologique à cette annotation semi-automatique en utilisant des tirs.

**Di Cristo et Hirst**

Di Cristo et Hirst utilisent un autre système d'annotation, basé sur l'algorithme MOMEL

(Hirst & Espesser, 1993) qui a pour fonction de styliser la courbe de F0 et de repérer automatiquement les points cibles. Ils associent cette annotation phonétique de la parole à une notation phonologique, Intsint (Hirst et al., 2000), transformant les points cibles détectés par l'algorithme MOMEL en tons : M(id), T(op) et B(ottom) pour représenter les valeurs de hauteur absolues et H(igher), L(ower), S(ame), U(ptepped) et D(ownstepped) pour les valeurs de hauteur relatives.

### **ToBI**

La transcription ToBI (Tone and Break Indices) est récemment très largement utilisée dans les travaux sur la prosodie du français. Cette transcription consiste en plusieurs niveaux :

- une transcription orthographique
- un niveau tonal où alternent tons L (low, bas) et H (high, haut). Les frontières des constituants prosodiques sont représentées par le signe % et les syllabes accentuées par l'ajout du signe \*. Les accents initiaux sont quant à eux transcrits « Hi ».
- un niveau de notation des indices de rupture permettant de distinguer les différents niveaux de frontière par une échelle allant de 0 à 4
- un niveau de notation *miscellaneous* pour les commentaires, par exemple la présence de rires, de souffle, etc.

Ce système de transcription a été pensé pour représenter les distinctions phonologiques, et non les réalisations phonétiques, notamment afin de faciliter les comparaisons avec les autres langues romanes (Delais-Roussarie et al., 2015). Bien que d'autres modèles prosodiques aient été pris en compte dans sa création, il est principalement basé sur la théorie auto-segmentale métrique (Beckman & Pierrehumbert, 1986 ; Ladd, 2008).

### **Les contours continuatifs et leur rôle dans la hiérarchie prosodique**

En français, les contours non-terminaux sont appelés **contours continuatifs** et s'observent à la fin des constituants prosodiques. Delattre (1966b) identifie deux types de

contours continuatifs : la continuation **mineure** et la continuation **majeure**. Cette distinction continuation mineure / majeure se retrouve dans la plupart des descriptions prosodiques et est assimilée à la différente réalisation du contour final suivant le type de constituant prosodique. Ainsi, la continuation mineure correspondrait au contour produit en frontière de syntagme accentuel (AP) tandis que la continuation majeure serait équivalente aux contours observés aux frontières de syntagme intermédiaire (ip) et intonatif (IP).

D'après Delattre, ces deux continuations se distinguent par le niveau de hauteur qu'elles atteignent. En effet, les continuations mineures et majeures commenceraient dans les deux cas au niveau 2, ensuite, les continuations majeures atteindraient le niveau 4 tandis que les continuations mineures n'atteindraient que le niveau 3. Delattre distingue également les contours continuatifs majeurs des montées mélodiques des questions, atteignant un même niveau 4, par la forme des contours produits. En effet, le contour de continuation serait de forme convexe tandis que le contour de question serait concave.

Cependant, ce critère est réfuté par Di Cristo (2016) qui observe la présence de contours continuatifs majeurs concaves ou convexes dans ses données et affirme ainsi que « il ne semble pas que la forme du contour nucléaire constitue un trait principal de l'identification du patron Continuatif Majeur » (p.209).

La différence de hauteur de registre décrite par Delattre est également critiquée par Di Cristo. En effet, il montre qu'en terme de valeurs de F0 relatives l'accentuation en fin de syntagme prosodique (équivalent syntagme accentuel) peut être sur une même hauteur mélodique que l'accent présent en fin de syntagme intonatif, en raison du phénomène naturel de déclinaison. Cependant, il observe bien une différence dans l'écart mélodique entre la syllabe prétonique et l'accent final ; **l'écart mélodique** produit en fin de syntagme intonatif étant plus important que les autres prééminences, à l'exception de certains cas d'emphase.

L'écart mélodique n'est cependant pas le seul marqueur permettant de différencier les accents finaux des syntagmes accentuels de ceux des syntagmes intonatifs. En effet, **la durée** joue également un rôle primordial. Selon Vaissière (2010), la durée seule permet de distinguer les phrases « Ton tonton Thon tond tonton » et « tonton Thon tond ton thon Thon », produites en synthèse vocale avec un F0 plat. Elle note également que « de façon

générale, plus le mot est indépendant du mot suivant, et plus l’allongement final est important (et plus le contour de F0 est montant, et/ou haut) » (p.15). Selon Delais-Roussarie et Post (2008) les syntagmes intonatifs sont en effet associés à un allongement « très important » de la syllabe accentuée, tandis que pour les syntagmes accentuels la dernière syllabe est allongée mais dans une moindre mesure (Feldhausen & Delais-Roussarie, 2012).

La présence d’une **pause** à la suite de l’accent final est également associée aux syntagmes intonatifs (et intermédiaires). Cette pause est cependant considérée comme facultative. Ainsi, selon Fougeron et Jun (1998) le IP est « *optionally followed by a pause* » (p.220). Delais-Roussarie et Post (2008) reprennent également ce critère de pause facultative dans leur description du syntagme intonatif. De manière similaire, Astésano et al. (2012) notent la présence fréquente d’une pause après le syntagme intermédiaire qui « [renforce] la perception d’une frontière forte » (p.358).

Le contour par « défaut » de frontière de syntagme accentuel, soit de la continuation mineure, est un contour montant (Jun & Fougeron, 2000). Delattre (1966b) avait cependant déjà remarqué que certaines continuations mineures étaient réalisées par une descente : « [la continuation mineure] peut descendre aussi bien que monter lorsqu’elle précède une courbe à niveau plus élevé ; par exemple, lorsqu’elle précède la continuation majeure. » (p.10). Selon lui, cette variation n’est pas significative et servirait à « briser la monotonie, [à] donner de la variété à la ligne mélodique » (p.10). Martin (1975) s’est beaucoup intéressé à ce phénomène de forte dépendance entre les continuations mineures (« C1 » dans sa représentation prosodique) et majeures (« C2 »), qu’il nomme « contraste de pente ». Ainsi, selon lui, lorsqu’un contour C1 est suivi d’un contour C2 montant, celui-ci serait réalisé par une descente mélodique. Il est toutefois important de noter que dans le système d’analyse de Martin, seuls les contours produits sur la syllabe accentuée sont pris en compte et cette descente ne concerne donc pas obligatoirement l’ensemble du constituant prosodique. Dans l’exemple présenté ci-dessous (Figure 2.9, le contour sur la syllabe accentuée à la fin du syntagme accentuel « le coléreux » est décrit comme descendant par Martin, tandis qu’il correspondrait à une cible haute (H\*) et donc à une montée mélodique dans Jun et Fougeron (2000) ou toute autre représentation basée sur un système d’annotation semblable à ToBI.

Dans certains énoncés, la distinction des continuations (mineures et majeures) est pri-

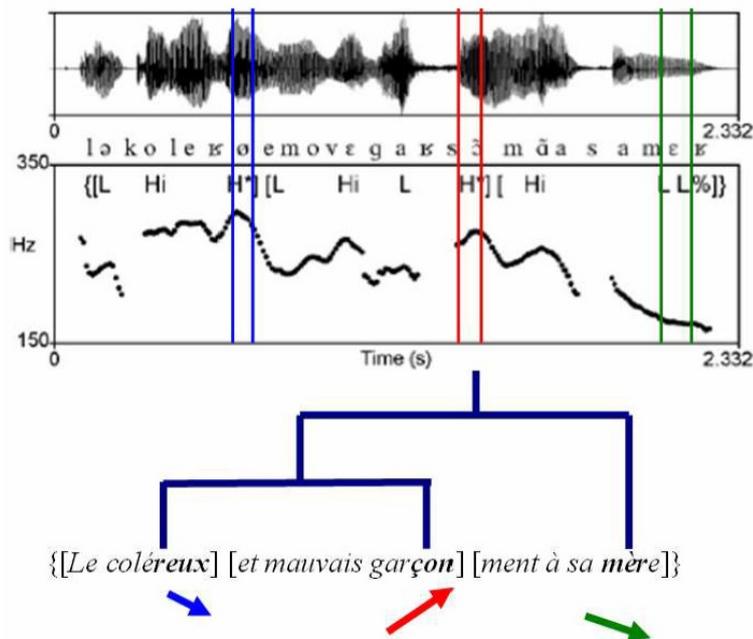


FIGURE 2.9 – Exemple de contraste de pente par Martin (2008)

mordiale pour accéder au sens de l'énoncé. En effet, selon Delattre « le rôle du niveau 4 est clairement de réunir de petites unités de sens en une grande unité de sens qui n'est pas la dernière de la phrase. [...] l'intonation de continuation majeure (4) fait sentir que toutes les unités mineures qui la précèdent, et qui ont l'intonation de continuation mineure (3), appartiennent à la grande unité de sens qui se termine au niveau 4 » (Delattre, 1966b, p.10). Delattre donne comme exemple les deux énoncés « Il a demandé / qui parlait // à Marie. » et « Il a demandé // qui parlait / à Marie » (les symboles « / » et « // » symbolisent respectivement une continuation mineure et une continuation majeure) où dans le premier cas le sujet demande à Marie qui parlait, et dans le deuxième il demande qui parlait à Marie à une tierce personne (1966, p.9). Mertens évoque aussi cette fonction groupante des différents niveaux de continuation : « Lorsqu'une frontière faible est suivie d'une frontière forte, leurs groupes forment une unité plus grande, tout en préservant le rapport de force entre les deux parties. » (Mertens, 2008, p.100). Il nomme ces grandes unités « paquets intonatifs » et insiste sur les relations entre l'organisation intonative et syntaxique en présentant l'exemple de « L'institut de phonétique de l'université de Bruxelles. » comportant deux « paquets intonatifs » délimités par des frontières fortes (HH sur « phonétique » et -B-B sur Bruxelles) contrastant avec les frontières faibles (/BB sur « institut » et « université »).

En effet, le groupement prosodique correspond dans ce cas à la structure syntaxique, et la production d'une frontière forte sur « institut » (HH) mais faible sur « phonétique » (/BB) serait impossible.

### La question du syntagme intermédiaire

Nous avons vu dans les sections précédentes que certaines descriptions prosodiques du français prenaient en compte un niveau intermédiaire, entre le syntagme accentuel et le syntagme intonatif : le syntagme intermédiaire ("si" ou "ip", *intermediate phrase*; Delais-Roussarie et al., 2015; Jun & Fougeron, 2000). Ce niveau prosodique correspond au niveau « 3 » du F\_ToBI. Le syntagme intermédiaire a longuement été décrit dans la thèse de Michelas (2011). Le syntagme intermédiaire est décrit comme suivant deux contraintes (D'Imperio & Michelas, 2009; Michelas, 2011) :

- MIN-BIN : le syntagme intermédiaire doit comporter au minimum deux syntagmes accentuels (AP)
- (ALIGN-XP,R, ip, R). : sa frontière droite est alignée avec celle d'une projection syntaxique majeure

Ses caractéristiques prosodiques sont les suivantes :

- un allongement de sa dernière syllabe plus important que celle d'un syntagme accentuel
- un retour mélodique à la ligne de référence marqué par le premier LH\* du syntagme intermédiaire, tandis que l'accent d'un AP interne au ip est rabaissé par rapport au premier accent
- un *reset* partiel après la frontière d'ip, l'accent final de l'AP du prochain ip étant moins haut que la ligne de référence de l'ip qui le précède.

Ces différentes caractéristiques peuvent s'observer sur les exemples en Figure 2.10. Nous pouvons voir le retour à la ligne mélodique de référence à la fin du ip, et inversement

observer un rabaissement sur l'accent final du AP interne au ip (deuxième AP dans la figure b.). L'allongement syllabique en fin de ip est aussi visible, notamment lorsqu'on le compare avec l'accent de l'AP interne. L'accent final du deuxième AP diffère grandement dans sa réalisation prosodique (durée et hauteur mélodique) selon le fait qu'il soit en frontière de AP (figure b.) ou de ip (figure b.). Dans sa thèse, par une tâche de complétion de phrases où les énoncés étaient tronqués à la fin du deuxième AP (cible), Michélas a montré que les francophones étaient ainsi capables de percevoir cette différence. On observe également la présence d'un *reset* partiel avec un accent final de AP suivant le ip (que Michélas signale par une flèche) plus bas que la ligne de référence.

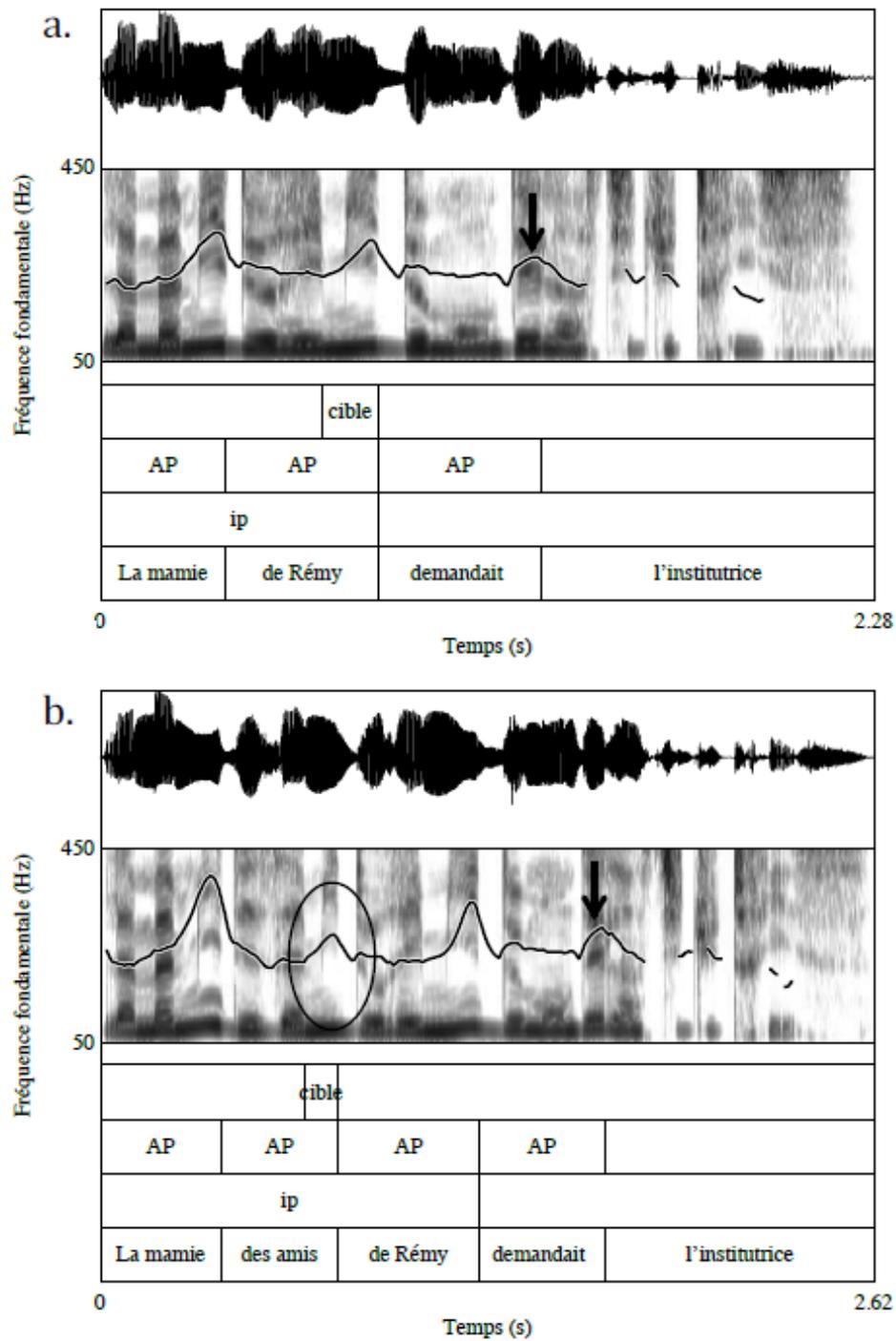


FIGURE 2.10 – Exemples issus de Michelas (2011) : syntagmes intermédiaires de 2 (figure a.) et 3 AP (figure b.)

Cependant, beaucoup contestent l'ajout de ce niveau intermédiaire ; c'est le cas notamment de Di Cristo (2016) qui considère que ses caractéristiques décrites par Michelas ne

permettent pas de le différencier du syntagme intonatif. Selon Di Cristo, les syntagmes intermédiaires correspondraient donc à des syntagmes intonatifs (IP).

Delais-Roussarie et al. (2020) ont récemment proposé une catégorisation des syntagmes intermédiaires, en distinguant :

- les *ips* métriques : structures semblables à celles traitées par Michelas, par exemple un syntagme nominal avec au minimum deux AP, ou un syntagme nominal objet dans le cas des constructions avec deux compléments d'objet ou avec un complément d'objet complexe
- les *ips* incidents : les ajouts à S et cadratifs (« Chaque lundi,... »), les expressions parenthétiques (« François, d'après ce qu'on m'a dit, va partir en vacances en Grèce. »), les dislocations (« A Paul, je lui ai donné un livre. »)
- les *ips* correspondant aux constituants de début de phrases, comprenant le focus informationnel de l'énoncé (je voudrais des oranges, s'il vous plaît madame.)

Il faut noter que ces différentes structures présentées comme syntagmes intermédiaires ont très souvent été analysées comme syntagmes intonatifs. C'est, comme nous l'avons évoqué précédemment, le cas de toutes ces structures (*ips* métriques inclus) selon l'approche de Di Cristo (2016). Mais c'est également le cas des cadratifs et expressions parenthétiques, ainsi que les constituants comportant le focus informationnel, dont les frontières droites sont souvent analysées comme une frontière de IP (Delais-Roussarie & Post, 2008 ; Mertens, 2008).

Selon elles, seuls les *ips* métriques peuvent être analysés comme des syntagmes intermédiaires, tandis que les *ips* incidents sont classés comme « IP mineurs ». Les IP majeurs correspondent, quant à eux, aux propositions ou clauses indépendantes et incluent les *ips* comportant le focus informationnel. Cette classification est justifiée par le fait que les contours finaux des *ips* métriques ne sont pas dépendants des autres contours et sont très généralement produits montants, à la différence des contours finaux des IP mineurs pouvant être descendants s'ils sont suivis par une frontière de IP majeur montante. Delais-Roussarie et al. illustrent ce phénomène par l'exemple suivant : [[(les amis)AP(du mari)AP

(de Valérie)AP\\_]Minor-IP [(je les ai appelés)AP] Major-IP [...et nous nous sommes rencontrés..] Major-IP. Le syntagme nominal disloqué « les amis du mari de Valérie », associé à la proposition indépendante (IP majeur) formant « Les amis du mari de Valérie je les ai appelés » finissant par une montée continuative majeure, est ainsi produit avec une descente mélodique à sa frontière.

### La syntaxe et la prosodie sont-elles isomorphes ?

De nombreuses approches théoriques sur l'intonation mettent en avant les relations étroites que celle-ci entretient avec la syntaxe. Ainsi, les constituants prosodiques sont très souvent définis par leurs relations avec la structure syntaxique ou macro-syntaxique. Delais-Roussarie et al. (2011) ont montré que leurs prédictions, basées sur la syntaxe, de réalisations de frontières de syntagme prosodiques et intonatifs dans un corpus de français oral (interviews) pouvaient se vérifier dans 90% des cas.

Cependant, bien que ce taux d'alignement avec la syntaxe soit très élevé, il n'en reste pas moins que dans 10% des cas les réalisations phonologiques attendues n'ont pu être observées. Ainsi, la structure syntaxique à elle seule ne nous permet pas toujours de prédire le phrasé prosodique. D'autres nombreux phénomènes ont aussi une influence sur le phrasé, c'est le cas notamment de la métrique. Le principe métrique le plus connu est l'**eurythmie**, pouvant être à l'origine de phrasés syntaxiquement incongruents. Voici un exemple de Martin (2018) illustrant ces propos, l'énoncé « Max adore les chocolats » pouvant être produit de deux manières :

1. congruent à la syntaxe, avec l'insertion d'une pause après le sujet « Max » :
  - [ **Max #** ] [ adore les chocolats ]
2. incongruent à la syntaxe en étant associé au verbe pour former des groupes plus équilibrés (eurythmie) :
  - [ Max **adore** ] [ les chocolats ]

Ainsi, la métrique joue un rôle primordial dans le phrasé. Rappelons également la définition donnée des syntagmes intermédiaires associant contraintes syntaxiques (Align-XP) et contraintes métriques (MIN-BIN).

La longueur des constituants joue également un rôle important dans le phrasé. Mertens (2009) illustre ce phénomène en comparant les deux énoncés « un enfant qui joue, oublie le monde extérieur » et « un petit enfant qui joue avec ses jouets, oublie le monde extérieur », ayant une même structure syntaxique – un groupe nominal suivi d’une relative. Il remarque que l’accentuation sur « enfant » est moins probable dans le premier cas que dans le deuxième, et explique cela par l’association de la contrainte de congruence syntaxique et de celle de la longueur du groupe intonatif :

Pourquoi la relative et son antécédent forment-ils tantôt un seul groupe intonatif, tantôt deux? Ceci résulte de l’interaction de deux contraintes : l’une liée à la congruence, l’autre à la longueur du groupe intonatif (le nombre de syllabes). La contrainte sur la congruence autorise les deux réalisations alternatives. En effet, comme la relative est régie par son antécédent, ils peuvent former ensemble un seul groupe intonatif ou être prononcés en deux groupes intonatifs successifs, participant au même paquet. Cependant, étant donné la longueur de la séquence « un petit enfant qui joue avec ses jouets », un deuxième accent apparaît facilement à la frontière syntaxique la plus importante au sein de la séquence. La fin de l’antécédent ne s’accompagne donc pas nécessairement d’une frontière prosodique, et certainement pas d’une frontière prosodique majeure. Il suffit que cette frontière soit inférieure ou égale à celle après la relative. (p.26)

Mertens note donc toutefois qu’une telle réalisation, bien qu’issue de contraintes rythmiques et non syntaxiques, reste congruente avec la syntaxe. En effet, la production à la fin de l’antécédent d’une frontière « inférieure ou égale à celle après la relative » permettrait de conserver l’unité syntaxique. Il note également qu’une continuation majeure ne peut toutefois pas être produite après l’antécédent, alors que ce pourrait être le cas à la fin de la relative. Dans ce cas la frontière en fin d’antécédent serait obligatoirement inférieure à celle-ci.

Martin (2006), citant Wioland (1985), évoque la règle des 7 syllabes contraignant les syllabes accentuables à n’être « obligatoirement accentuées que si elles empêchent une succession de 7 syllabes inaccentuées successives » (p.145). Il remarque cependant que ce

nombre de 7 syllabes n'est pas toujours respecté et qu'il peut aller jusqu'à 10 syllabes par groupe accentuel avec un débit de parole très rapide. Ainsi, plus qu'un nombre de syllabes maximal, ce serait une contrainte de durée maximale entre deux syllabes accentuées, que Martin (2018) définit à 1250 ms. Il met aussi en évidence le lien existant entre les ondes cérébrales delta et les groupes accentuels, puisque la période maximale de l'onde delta (de 1250 ms) correspond à la durée maximale possible entre deux syllabes accentuées et que l'on peut observer une synchronisation entre syllabes accentuées et ondes deltas.

La congruence entre prosodie et syntaxe est également étroitement liée au type de tâche réalisée. En effet, il est généralement admis que lors d'une lecture oralisée on observe une congruence plus importante que pour une tâche de parole spontanée. Martin (2018) évoque ainsi « la nécessité pour le lecteur d'évaluer la syntaxe imposée par le texte avant d'y adapter une intonation de phrase, alors que pour le locuteur spontané, c'est la syntaxe qui va être adaptée à l'intonation qu'il planifie en premier lieu » (p.38).

### Récapitulatif : les contours continuatifs et leurs équivalences

Dans les sections précédentes, nous avons vu que les contours continuatifs du français n'étaient pas nécessairement au coeur des descriptions prosodiques du français. Certaines de ces descriptions mettent en avant les liens entre syntaxe et métrique en proposant des niveaux de constituants prosodiques définis par ces critères, associés à des réalisations prosodiques particulières. Le nombre et les noms de ces constituants varient grandement entre les différentes descriptions, nous proposons ici pour plus de clarté un tableau récapitulatif essayant de mettre en évidence les équivalences entre ces descriptions, du point de vue des contours non-finaux.

Niveau annotation ToBI	1	2	3	4
Description prosodique				
Jun & Fougeron (2000)	-	Accentual phrase (AP) LH*	Intermediate Phrase (ip) LH*H-	Intonation Phrase (IP) LH*H%
Michelas et d'Imperio (2010)	Mot prosodique	Accentual phrase (AP)	Intermediate Phrase (ip)	Intonation Phrase (IP)
Mertens (2008)	Groupe accentuel (GA)	Groupe intonatif (GI)	« paquet intonatif »	
Di Cristo (2016)	Unité Tonale (UT)	Syntagme Prosodique (SP)	Unité Intonative (UI)	
Martin (1987, 2018)	-	Mots prosodiques ou groupes accentuels - le contour continuatif associé étant C2	Syntagme intonatif - le contour continuatif associé étant C1	
Delais-Roussarie et al (2020)	-	AP	ip métrique	IP mineur / IP majeur
Delattre (1966)	-	Continuation mineure	Continuation majeure	

TABLEAU 2.1 – Tableau des équivalences entre les différentes descriptions prosodiques

Notons que certaines correspondances sont approximatives. En effet l'unité tonale de

Di Cristo (2016) n'est pas tout à fait équivalente au GA de Mertens (1993), puisque le premier correspond à une unité plus petite pouvant marquer indifféremment l'accent initial ou l'accent final. De même, le niveau ip (Jun & Fougeron, 2000 ; Michelas, 2011) est associé à « continuation majeure » de Delattre bien qu'ils ne soient pas définis par les mêmes indices (retour au niveau de référence mélodique pour le ip et mouvement mélodique du niveau de hauteur 2 à 4 pour la continuation majeure de Delattre). Dans cette thèse, nous choisissons d'utiliser le système d'annotation ToBI et de distinguer ainsi trois niveaux prosodiques au-dessus du mot, mais nous utilisons les termes de « continuation mineure » et « continuation majeure », plus parlant de notre point de vue pour décrire les contours montants produits respectivement aux frontières de AP et de IP.

### 2.2.2 Le système prosodique du japonais

Nous avons vu dans la section 2.1 que l'accentuation du japonais différait grandement du français puisque le japonais est caractérisé par la présence d'accent de hauteur mélodique tandis que le français porte son accent principal en fin de groupe. Selon l'analyse prosodique hiérarchique de Beckman et Pierrehumbert (1986), l'accent tonal se réalise au niveau du syntagme accentuel, en combinant un ton H au début et un ton bas à la fin ; ainsi, le syntagme accentuel est toujours marqué par une chute, qu'il porte un accent tonal (%L (H-) H\*+L L%) ou qu'il soit considéré comme non accentué (%L H- L%) (voir Figure 2.11). Igarashi (2015) note également que si un AP classique est composé d'un mot lexical et d'une particule, cette construction n'est pas la seule possible : un AP peut contenir plusieurs mots lexicaux et certaines études démontrent l'existence d'AP correspondant à une particule seule (il cite entre autres Kubozono 1988/1993, Maekawa et Igarashi 2007, et Vance 2008). Un AP peut également comporter plusieurs mots lexicaux : lorsque le premier mot est non accentué il peut se rattacher à celui qui le suit pour ne former qu'un seul AP.

Que l'AP soit accentué ou non, on peut observer que dans les deux cas il finit toujours par un ton bas (L%), bien que nous ayons vu dans la première section de ce chapitre (2.1) que le niveau de hauteur mélodique sur la particule (en fin de AP) est semblable à celui de la dernière more du mot non accentué, soit dans un registre relativement haut. (voir Figure 2.11).

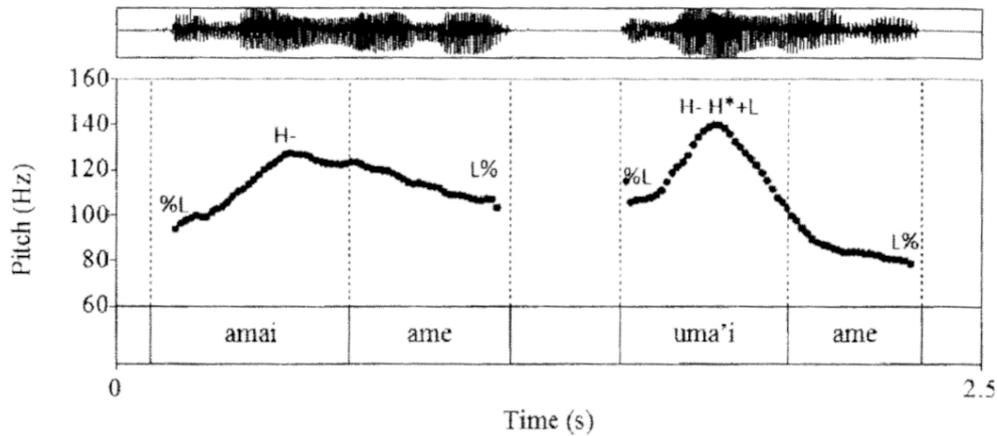


FIGURE 2.11 – AP accentué et non accentué (sans particule) – exemple issu du Handbook of Japanese Phonetics and Phonology, Igarashi (2015); Traduction : « (un/des) bonbon(s) sucré(s) » (à gauche) et « (un/des) bon(s) bonbon(s) » (à droite)

Le syntagme intonatif japonais peut être composé d'un ou de plusieurs AP. La particularité du syntagme intonatif est que c'est un domaine prosodique dans lequel un *pitch range* est associé, et le locuteur choisit un nouveau *pitch range* au début de chaque IP indépendamment de celui qui précède (Venditti, 2005). Il est très fortement lié au phénomène que l'on retrouve en japonais de *downstep* ou *catathesis* entre les différents AP, où chaque AP a un pic de F0 plus bas que celui qui le précède, donnant une courbe de F0 en forme d'escaliers. En effet, ce *downstep* ne se produit qu'à l'intérieur d'un même IP et il est bloqué lorsqu'un nouveau *pitch range* est choisi à l'IP suivant, ce que l'on appelle « *pitch reset* ». Ce phénomène est illustré en Figure 2.12.

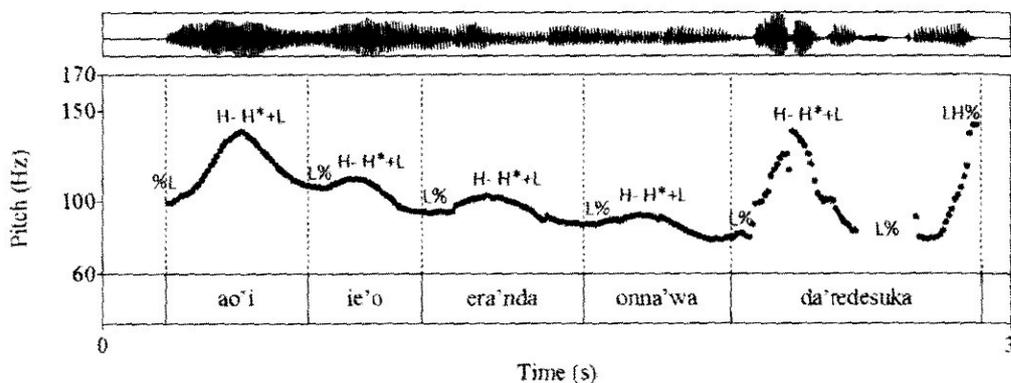


FIGURE 2.12 – Downstep en japonais – exemple issu du Handbook of Japanese Phonetics and Phonology, Igarashi (2015); glose : bleue maison-ACC choisir-PASSE femme-TOPIC qui-COP, POL-Q

Si la plupart des descriptions prosodiques du japonais décrivent deux niveaux de constituants (AP et IP pour J\_ToBI), certains modèles intègrent un niveau supérieur à IP (Utterance) ou considèrent qu'il existe deux niveaux prosodiques entre Utterance et AP. Les principaux modèles existants sont présentés par Igarashi. Selon Kawahara et Shinya (2008), il existe quatre niveaux de constituants : *Minor Phrase* (MiP), *Major Phrase* (MaP), *Intonational Phrase* (IP) et Utterance. Dans leur étude portant sur la production de phrases lues (dont des phrases coordonnées), ils observent certaines différences prosodiques entre les différents niveaux de constituants. Chaque phrase syntaxique (*clause*) constitue un IP, caractérisé par un abaissement tonal, la présence de voix craquée et une pause à sa frontière droite, ainsi qu'un *pitch reset* et une montée initiale plus importante que pour un MaP. Les frontières gauches de MaP correspondraient quant à elles aux frontières gauches de SV (syntagme verbal), et les frontières gauches de MiP aux frontières gauches de SN (syntagme nominal).

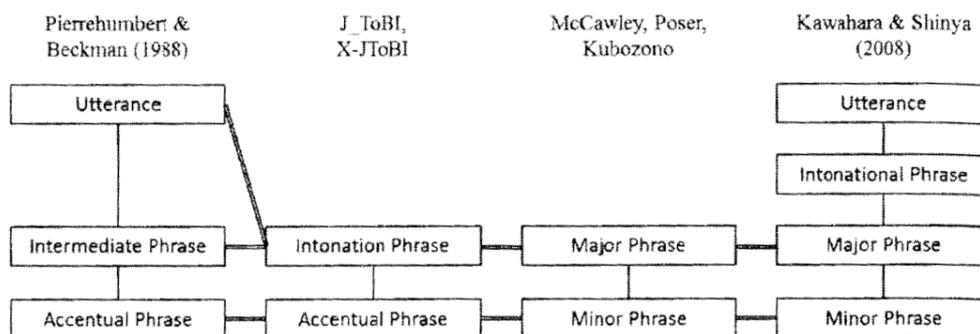


FIGURE 2.13 – Différents modèles prosodiques – Exemple issu du Handbook of Japanese Phonetics and Phonology, Igarashi (2015)

Nous avons principalement présenté ici la description ToBI du système prosodique du japonais que nous choisissons de retenir pour nos discussions théoriques dans cette thèse. Ce choix est en grand partie motivé par le fait que la notation ToBI puisse être utilisée dans la description de la hiérarchie prosodique du français et du japonais, facilitant ainsi les comparaisons entre les deux langues. D'autres descriptions du japonais existent néanmoins ; il s'agit principalement de descriptions basées sur les contours prosodiques que l'on observe en japonais. C'est notamment le cas des célèbres travaux de Kawakami (1963), faisant l'inventaire des contours prosodiques (de question, non-finaux, etc.) qui

ont servi de base théorique pour l'étude de la prosodie du japonais, notamment en ce qui concerne les Boundary Pitch Movements (BPM) que nous présentons dans la section suivante.

## 2.3 Prosodie et interaction : le rôle du continuatif

Dans la section précédente, nous avons vu que le contour continuatif français était fortement lié à la hiérarchie prosodique en étant associé aux frontières des constituants. Cependant, les continuatifs français ne sauraient être réduits à ce rôle démarcatif. En effet, Portes, Bertrand et al. (2012; 2007; 2005; 2007) se sont intéressés au rôle discursif et interactionnel des continuations du français, notamment lors des conversations. Leurs recherches remettent en question la position de Delais-Roussarie (2005) considérant que les contours continuatifs du français sont des mouvements mélodiques par défaut contraints par des facteurs syntaxiques et métriques. En effet, elles considèrent que si c'est bien le cas des continuations mineures, cette définition de mouvement mélodique par défaut ne s'appliquerait pas aux continuations majeures du français qui auraient un véritable rôle discursif.

### **Le rôle dialogique et interactionnel des contours continuatifs du français**

Dans une de leurs études, Portes et Bertrand (2005) ont analysé 10 minutes de dialogue entre deux locutrices, issues du CID (*Corpus of dialogal interaction*, Bertrand et al. 2008) en identifiant deux niveaux de groupes prosodiques : RG (*Rhythmic Group*, groupe rythmique) et IP (*Intonational Phrase*, syntagme intonatif). Elles ont ensuite identifié auditivement les contours mélodiques aux frontières de ces groupes prosodiques, distinguant continuations majeures et mineures (mc et MC), contours montants (R), contours descendants (F), contours montants-descendants (RF1 et RF2) et contour « parenthétique » avec une intonation plate (PAR). Elles ont ensuite choisi d'ajouter une catégorie pour l'énumération (ENUM). D'autre part, elles ont aussi annoté les événements conversationnels selon le cadre de l'analyse conversationnelle (*Conversational Analysis* Selting 2000) en relevant les tours de parole dans l'extrait analysé. Selon cette analyse, les tours de parole sont constitués d'unités appelées *turn-constructional units* (TCU) et les auteurs reprennent la définition de Selting (2000) les qualifiant de : « *the smallest interactionally relevant complete*

*linguistic unit* ». Ce sont en effet des unités qui doivent être complètes syntactiquement et prosodiquement mais aussi pragmatiquement. Toujours en suivant les définitions de Selting (2000), les auteurs choisissent de distinguer deux types de TCU : les TCU non-finaux et les TCU finaux. La particularité des TCU finaux est qu'ils sont associés à une idée de complétion et sont potentiellement suivis par un changement de locuteur ; ils finissent par ce que l'analyse conversationnelle nomme TRP (*transition relevance place*). Cela ne veut pas dire d'un changement de locuteur aura nécessairement lieu à cet endroit mais que cela est possible. Les TCU non-finaux, quant à eux, sont dits « incomplets », par exemple pour des raisons sémantiques ou pragmatiques, et ne finissent pas par des TRP. Lors d'une conversation, l'interlocuteur utilise souvent de courts mots ou sons signifiant qu'il est bien attentif au discours du locuteur (« OK », « oui », « mhm »...) ou même des gestes (un hochement de tête, par exemple). Il s'agit de *backchannels*, terme introduit par Yngve (1970 cité par Schegloff 1982) et que Schegloff nomme *continuers*. Ces *backchannels* peuvent en effet indiquer au locuteur que l'interlocuteur est attentif à ses propos en attendant leur suite, et ne cherche donc pas à prendre le tour de parole.

Les continuations majeures sont plus fréquemment associées à des *backchannels* de la part de l'interlocuteur que les continuations mineures (32% contre 11,6%). Les auteurs notent par ailleurs que dans le cas des continuations mineures ces *backchannels* ont plus souvent pour fonction de relancer l'interlocuteur lors de disfluences que d'approuver et soutenir ses propos. Les continuations majeures sont associées dans 36 % des cas à la fin d'un TCU non final et dans 42,7 % des cas à un TCU final, tandis que les continuations mineures ne sont jamais associées à des TCU. Parmi les continuations majeures relevées dans leur étude, Portes et Bertrand (2005) observent toutefois que 18,4 % d'entre elles ne sont associées ni à un TCU ni à des *backchannels*, et n'ont pas de rôle interactionnel mais plutôt une fonction structurelle. En effet, elles apparaissent par exemple à la fin d'adverbes isolés ou d'expressions qui « introduisent un changement dans l'activité discursive ou le début d'un nouveau tour de parole » (notre traduction).

Au vu de ces résultats, les auteurs concluent donc qu'il est nécessaire de conserver la dichotomie continuation mineure / majeure puisque seules les continuations majeures semblent avoir un rôle discursif et interactionnel. Ce rôle serait le suivant : « *to inform the addressee that the content of the next utterance is to be understood related to the current one* »

(p.12)

Dans une étude postérieure, Portes et al. (2007) distinguent les continuations majeures en position focale des continuations en position non-focales. Il s'agit ici d'une autre notion de finalité que celle abordée précédemment (TCU finaux et non-finaux). En effet, Portes et al. considèrent les contours en position non-focale comme correspondant aux continuations majeures définies par Delattre ayant comme rôle de rassembler plusieurs unités en un plus grand groupe de sens, par exemple à la fin d'un syntagme nominal ou en frontière d'un élément disloqué. Les continuations majeures aux positions focales sont quant à elles *inales* et ont le rôle défini par Pierrehumbert et Hirschberg (1990) de « *to be interpreted with respect to a succeeding phrase* » (p.305). Le corpus utilisé dans cette étude est à nouveau le *Corpus of Interactional Data* (Bertrand et al., 2008) mais les auteurs ont analysé cette fois-ci un total de 8 heures de dialogues. Ils ont annoté la présence de contours continuatifs montants en position focale (493 contours) et non-focale (896 contours) ainsi que les contours d'énumération (148 contours) et relevé la durée de la syllabe accentuée de chaque contour, l'écart mélodique entre la cible L de départ et la cible H à la fin du contour ainsi que la différence temporelle entre les deux cibles tonales L et H. Les résultats de leurs analyses montrent qu'il n'y a pas de différence phonétique entre les contours continuatifs montants en position focale et non-focale. Ainsi, les seules différences entre ces deux contours sont phonologiques puisque seuls les contours continuatifs en position focale peuvent avoir un rôle discursif.

Cette notion de focalité pour distinguer les contours continuatifs montants du français n'est cependant pas reprise dans une autre étude des auteurs (Bertrand et al., 2007), mais comme dans la première étude que nous avons présentée (Portes & Bertrand, 2005) c'est la distinction entre les TCU finaux et non finaux qui est mise en avant. Alors que Portes et Bertrand (2005) distinguent deux niveaux de continuation (continuations mineures, mc, et continuations majeures, MC), ces niveaux sont remplacés par trois catégories dans Bertrand et al. (2007) : les contours mineurs (mc), les contours continuatifs majeurs montants (*Rising Major Continuations*, RMC) et les contours majeurs montants (*terminal rising*, RT). Cette classification est liée à la distinction faite entre les TCU finaux et non-finaux. En effet, selon Bertrand et al. (2007) : « par définition, les /RMC/ ne peuvent pas coïncider avec une TRP puisqu'elles contribuent à projeter une suite, à savoir les étapes attendues du récit »

(p.71). Or, les TCU finaux sont nécessairement associés à une TRP, et ainsi les contours montants majeurs produits à cette position sont considérés comme terminaux (RT) et non continuatifs. Ces deux catégories distinctes n'en formaient qu'une (*Major Continuations*, MC) dans Portes et Bertrand (2005). Les /RMC/ correspondent indifféremment aux continuations majeures montantes produites en position focale ou non-focale, les auteurs rappelant que « ce contour [peut] occuper *sans variation de forme* deux distributions différentes : à l'intérieur des énoncés d'une part et en fin d'énoncé d'autre part. » mais que « le 'continuatif' prend une valeur discursive interactionnelle en fin d'énoncé qu'il n'a pas sinon. ». Bertrand et al. (2007) ont analysé un extrait de 10 minutes de conversation du *Corpus of Interactional Data* et se sont intéressés à la distribution syntaxique et discursive des différents contours mélodiques. Ils distinguent notamment les macro-séquences « dialogales » des macro-séquences « monologiques », les premières correspondant par exemple à des interactions de type question-réponse tandis que les secondes correspondent généralement à des séquences narratives *quasi* monologiques où l'interlocuteur n'intervient que par *backchannels*, en coénonciation ou pour demander des précisions ou faire des commentaires. Aux frontières des macro-séquences dialogales, ils observent que les contours terminaux sont majoritairement montants ou de très faible variation mélodique (/fl/, *flat*). A la fin des macro-séquences monologiques, ils observent également beaucoup de contours montants /RT/ mais aussi des contours de liste « virtuelle » /RL/ qui marquent la fin d'un paragraphe ou une séquence et demandent à l'interlocuteur d'imaginer la suite. A l'intérieur des séquences dialogales, presque aucun contour montant de continuation /RMC/ n'est observé, tandis qu'il est très majoritaire dans les macro-séquences monologiques et à la fin des micro-séquences les composant.

Ces différentes études prouvent ainsi le rôle interactionnel de la continuation majeure en français, et la nécessité de continuer à distinguer les continuations majeures des continuations mineures en français. Les contours continuatifs majeurs montants sont séparés en deux catégories : les contours en position dite « non-focale » à l'intérieur d'un énoncé et les contours en position « focale » en fin d'énoncé (Portes et al., 2007). Bien que leur réalisation phonétique ne semble différer, seuls ces derniers ont un véritable rôle interactionnel. Ainsi, les contours continuatifs majeurs montants sont plus souvent associés à des *backchannels* quand ils sont en fin de construction verbale que lorsqu'ils sont en fin

de pré-noyau (Bertrand et al., 2007). Ces *backchannels* sont principalement des *continuers*. Selon Bertrand et al. (2007) les contours continuatifs majeurs ont donc ici pour rôle d'« empêcher que les points de complétion syntaxique soient considérés comme des TRP » et par la même occasion de « suspendre et de reporter le moment de l'interprétation » (p.75). Ils appuient cette hypothèse par le fait que ces contours apparaissent majoritairement à la fin des macro-séquences narratives qui impliquent plusieurs étapes discursives. Nous concluons donc que ces contours indiquent à l'interlocuteur la nécessité de continuer à être attentif et écouter la suite énoncée par le locuteur, sans l'interrompre – ce qui implique également un rôle de maintien du tour de parole (*turn holding cue*) du continuatif majeur montant français.

Les études présentées précédemment ont aussi questionné le rôle des énumérations et, plus particulièrement en quoi celles-ci diffèrent ou non des continuations majeures. Selon Portes et al. (2007) l'énumération s'apparenterait au contour continuatif montant majeur puisqu'elle indiquerait que la phrase où elle apparaît doit être interprétée avec celle qui suit : « *it also gives the instruction to delete the final paradigm and to interpret the current clause with the following one when appearing in focal position* » (p.6). L'énumération est ainsi décrite comme une « sous-catégorie » de la continuation majeure, ayant comme caractéristique formelle spécifique une pente mélodique (*slope*) plus petite. Les auteurs notent l'existence de « listes virtuelles » où seul le premier objet de la liste est cité et la suite de la liste est laissée à la libre imagination de l'interlocuteur. Dans son étude portant sur une locutrice francophone native, Kaminskaïa (2016) s'est exclusivement intéressée au contour d'énumération. Elle a relevé 22 contours d'énumérations dans une interview de 1202 secondes ; 18 « marquées émotionnellement » et 4 « non-marquées émotionnellement ». Elle observe un débit de parole plus lent pour les énumérations que pour les continuations : 3,26 et 3,96 syllabes / seconde pour les énumérations (non-marquées émotionnellement / marquées) contre 3,82 et 5,99 syllabes / seconde pour les continuations (*idem*), ainsi qu'une syllabe accentuée proportionnellement plus longue pour les énumérations. Concernant les valeurs de F0, elle note une montée mélodique plus longue pour les énumérations avec les cibles tonales L et H plus espacées que pour les continuations, ainsi qu'un écart de F0 plus réduit. Ainsi, Kaminskaïa confirme les observations faites par Portes et al. (2007) décrivant les contours d'énumération comme étant caractérisés par une

pente mélodique moins abrupte (« *smaller slope* »). De plus, la cible tonale H tend à être produite plus tardivement sur la syllabe accentuée des énumérations. Kaminskaïa propose alors, si ces observations peuvent être confirmées par la suite à plus grande échelle, de retranscrire les contours énumératifs du français par L\*H- et de conserver la notation LH\* pour les contours continuatifs. Lorsqu'ils évoquent les contours d'énumération, Portes et Bertrand (2005) et Kaminskaïa utilisent le terme de « cliché » mélodique. Si ces deux études font plutôt état de contours énumératifs montants (à l'exception des contours énumératifs « non-marqués émotionnellement » de Kaminskaïa qu'elle décrit comme étant relativement « plats » au niveau intonatif), Di Cristo (2016) décrit une plus grande variété de patrons mélodiques d'énumération. En effet, bien qu'il reconnaisse que le contour montant d'énumération est « prototypique », il fait état de contours énumératifs descendants. En tout, six types de contours énumératifs sont décrits :

1. le contour énumératif montant (voir schéma *a* de l'image 2.14) et le contour énumératif descendant (schéma *b*)
2. un contour énumératif montant qualifié de « cliché mélodique » associé à un fort allongement (reporté comme « exceptionnel ») de la voyelle accentuée (schéma *c*)
3. un contour énumératif (cliché mélodique) qui comporte un « rehaussement de la prétonique et un allongement de la tonique finale dans un registre mi-haut formant un plateau » (schéma *d*)
4. un contour énumératif (cliché mélodique) décrit comme portant un ton haut à la fin du syntagme verbal suivi d'un plateau mélodique qui s'étend sur le groupe complément (schéma *e*)
5. un contour énumératif (cliché mélodique) qui comporte un rehaussement de la prétonique et un contour descendant étendu, mais descendant dans un registre moins bas qu'un patron mélodique conclusif (schéma *f* ; le schéma *g* représente un contour conclusif à but comparatif)

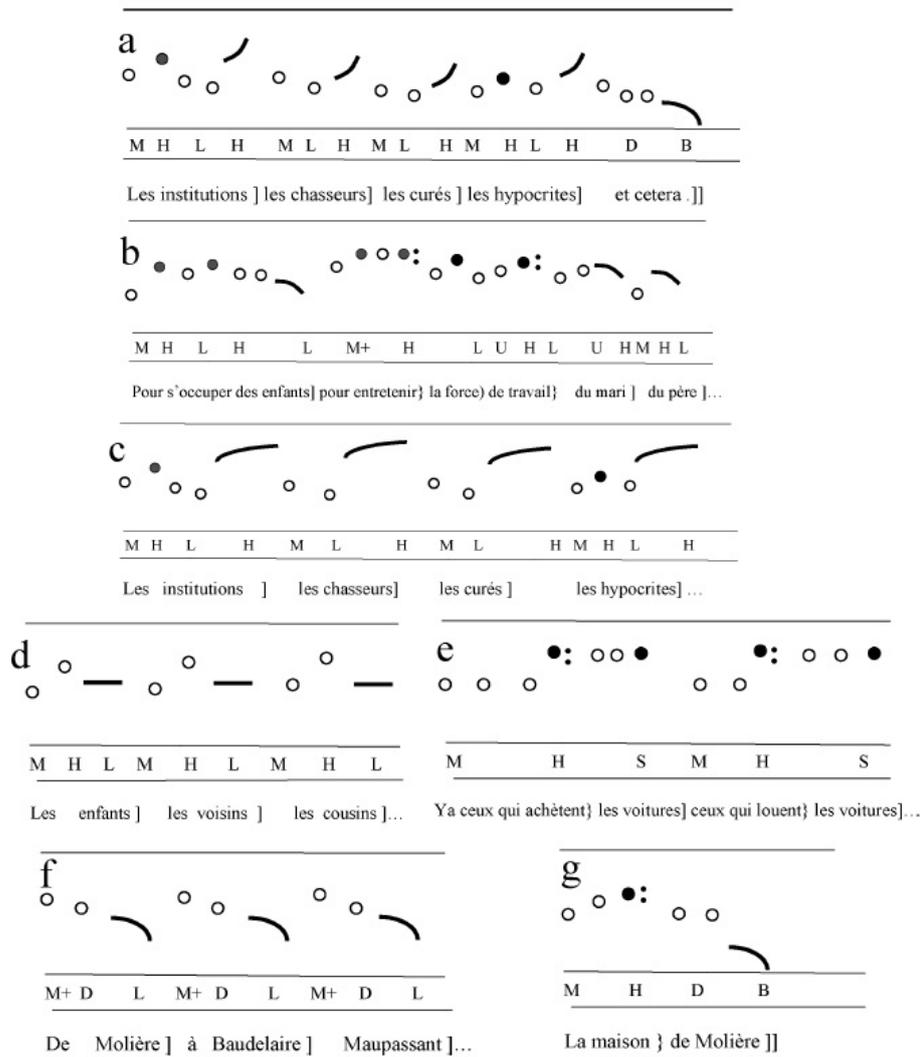


FIGURE 2.14 – Différents patrons intonatifs d'énumération – issu de Di Cristo 2016

Le cliché mélodique de l'énumération associé à un fort allongement vocalique sur sa syllabe accentuée (schéma *c*) nous semble en effet bien représentatif de l'énumération en français, notamment lorsque l'on cherche à l'expliquer ou l'enseigner. Dans les faits, comme le décrit Di Cristo, la voyelle accentuée n'est cependant pas toujours aussi fortement allongée, mais cela correspond aux résultats de Kaminskaïa qui observe un allongement plus marqué pour les énumérations. Il est aussi possible que le contexte discursif joue un rôle important dans le choix du patron mélodique énumératif. Par exemple, nous pouvons imaginer qu'une situation de communication plus relâchée favorise l'utilisation de clichés mélodiques, tandis que les contours montants « prototypiques » seraient privilégiés lors d'une conversation plus formelle. Dans tous les cas, nous sommes en accord avec

Portes et al. (2007) et Kaminskaïa (2016) qui considèrent le contour énumératif comme étant une « sous-catégorie du contour continuatif » et qui partage avec ce dernier une même fonction discursive continuative.

Nous avons vu que la continuation phonologique en français n'avait pas d'équivalent en japonais où seuls le *pitch accent* et le *downstep* interne caractérisaient les différents niveaux prosodiques. En français, la dichotomie contour mineur et contour majeur reflète des types de réalisations de continuations qui diffèrent suivant la force de frontière prosodique ; typiquement une continuation mineure en fin de syntagme accentuel mais majeure en fin de syntagme intonatif. Mais en dehors de ce rôle purement phonologique, les travaux de Portes, Bertrand et al ont également montré le rôle de la continuation et les différences entre contours mineurs et majeurs dans l'interaction.

#### **Des contours montants à valeur discursive en japonais – les *Boundary Pitch Movements* (BPM)**

S'il n'y a pas d'équivalent à la continuation phonologique du français en japonais, il existe cependant des montées de F0 aux frontières prosodiques en parole spontanée appelées les *Boundary Pitch Movements* (BPM). Les BPM sont des mouvements de F0 apparaissant à la fin des groupements prosodiques japonais, généralement sur la dernière more. Ils peuvent donc apparaître en fin d'énoncé, notamment pour marquer les questions, ou à l'intérieur de l'énoncé. Les BPM sont généralement produits sur le dernier AP d'un IP, mais peuvent parfois être observés en fin de AP sans qu'ils ne coïncident avec une frontière de IP (Igarashi & Koiso, 2012). Les BPM ont été décrits dans de nombreuses études, notamment celles de Kawakami (1963). Kawakami est l'un des premiers à s'intéresser à ce sujet, et a différencié les types de BPM suivant leur forme de contour mais aussi leur fonction pragmatique.

Dans le modèle intonatif J\_ToBI ainsi que sa version élargie pour l'analyse de la parole spontanée X\_JToBI (Maekawa et al., 2002), les BPM y sont décrits dans une liste où l'on retrouve la grande majorité des BPM décrits dans les études précédentes. Là où les premières études sur les BPM leur donnaient un nom différent à chacun (utilisant parfois même plusieurs noms pour décrire le même BPM, et variant aussi selon les auteurs), ils sont ici décrits par des tons (H) et bas (L), comme le demande la transcription ToBI. Par

exemple le contour montant (*jōshōchō*) de Kawakami (1963) est représenté par H%.

- LH% : BPM généralement produit en fin d'énoncé pour marquer l'interrogation ou l'incrédulité

Phonétiquement, il diffère du BPM H% par sa montée de F0 qui ne commence qu'à partir du milieu de la dernière more de l'AP, et non pas au début, ainsi que par sa forme plus concave.

- H% : BPM marquant l'insistance ou l'importance de l'information donnée. Il a également pour fonction de segmenter la parole en *chunks* (Venditti et al., 2008). Il est produit par une simple montée mélodique. Il existe cependant plusieurs types de contours montants H% en japonais qui n'ont pas exactement la même fonction pragmatique (Venditti et al., 1998). Pour mieux classifier ces BPM, le symbole « > » est proposé dans le X\_JToBI (Maekawa et al., 2002). Ce symbole permet de noter H%> l'un des variants de H%, produit de façon identique dans sa première partie (montée de F0) mais ensuite suivi par un plateau mélodique ainsi qu'un allongement vocalique.

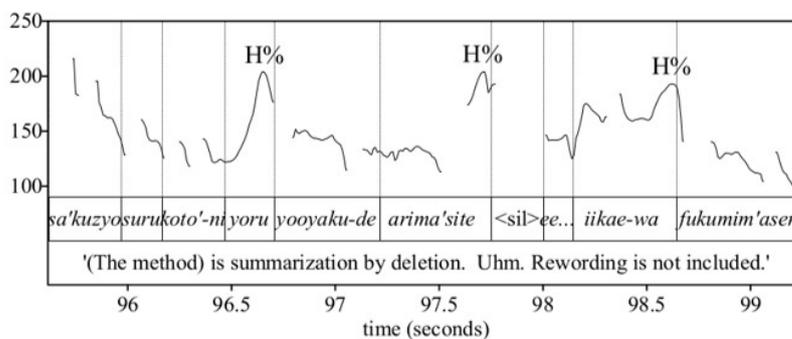


FIGURE 2.15 – BPM H% – exemple issu du Handbook of Japanese Phonetics and Phonology Igarashi (2015)

- HL% : BPM plus fréquemment utilisé à l'oral, notamment lorsque le locuteur donne des explications. Il diffère du BPM H% par la présence d'une chute mélodique après sa première partie montante, ainsi que par un allongement de la dernière more plus important. Il partage la plupart des fonctions du BPM H% – notamment la fonction

de marquage de l'importance de l'information qui précède et le *chunking* (Venditti 2008). Ces similitudes donnent lieu à l'hypothèse que ces deux BPM sont deux variantes distinctes du même phénomène et que le BPM H% serait une variante « tronquée » du BPM HL% (Igarashi, 2015). Cette hypothèse est renforcée par la différence de durée entre les deux BPM (HL% étant plus long que H%).

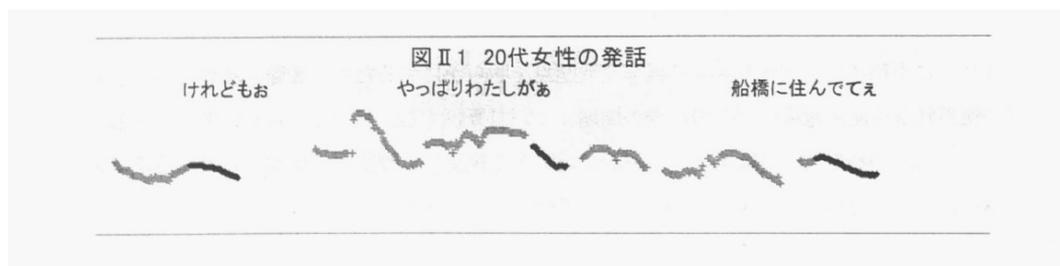


FIGURE 2.16 – BPM HL% – exemple issu de Sasaki-Hara 2004

- HLH% : ce BPM est rarement observé et il serait principalement utilisé dans le langage adressé à l'enfant. (Venditti et al., 2008)

Si les BPM montants sont fréquents aux frontières prosodiques du japonais en parole spontanée, la majorité de ces frontières sont cependant réalisées par un ton bas (L%). Ainsi, Koiso (2014) a noté la présence aux frontières prosodiques de : 74,59 % de tons bas (L%), 18,87 % de contours montants (L%H%), 6,31% de contours montant-descendants ((L%HL%), 0,22% de BPM L%HLH% et 0,22% de BPM L%LH%. Ces proportions peuvent cependant varier selon le style de parole, la force de frontière discursive (voir Yoneyama et al. 2003, détaillé plus loin dans ce chapitre) ou la situation de communication.

Nous avons vu que parmi les BPM présents dans le système japonais, les BPM H% et HL% (et H%>) avaient des fonctions similaires, permettant d'avancer l'hypothèse qu'il ne s'agisse que de deux variantes d'un même phénomène. Cependant, il faut noter que ces deux BPM se distinguent par différents points.

Tout d'abord, la présence de BPM HL% en japonais est relativement récente et cette façon de parler se serait dans un premier temps propagée chez les jeunes (notamment les jeunes femmes) avant de se généraliser (Inoue, 1997). Il subsiste cependant des différences inter-générationnelles puisque les locuteurs plus âgés produisent bien plus rarement ce

BPM que les jeunes locuteurs. Les personnes plus âgées ont tendance à utiliser des *fillers* (pauses remplies : « A- », « E- »...) à la place des HL% (Sasaki-Hara, 2004).

Dans son étude basée sur trois monologues du corpus CSJ (*Corpus of Spontaneous Japanese*), Maekawa (2006) met également en avant les liens existant entre le type de parole et l'utilisation de ces deux BPM. Elle montre un usage plus fréquent de BPM H% dans la parole jugée plus formelle et moins spontanée. À l'inverse, un type de discours jugé moins formel et une parole moins spontanée sont associés à une plus grande fréquence de BPM HL%.

Bien que l'on prête des fonctions pragmatiques semblables aux BPM H% et HL%, on peut aussi observer quelques différences. En effet, HL% est plus fortement associé à un rôle continuatif et explicatif que H%. Il est décrit comme ayant pour principales fonctions le maintien du tour de parole et la conservation de l'attention du locuteur (Inoue, 1997). L'étude de Koiso et al. (1998), analysant les 5 premières minutes de 8 dialogues issus du *Japanese Map Task Corpus* (Aono et al. 1994), confirme ces affirmations. En effet, les mouvements mélodiques montant-descendants (HL%) sont significativement plus souvent associés au maintien du tour de parole ainsi qu'à la présence de *backchannels*. Ces résultats peuvent être mis en parallèle avec ceux obtenus par Venditti et al. (1998), montrant que les BPM HL% sont jugés comme plus probables d'indiquer la continuation ou un discours explicatif que les BPM H% d'insistance. Le BPM H% est toutefois lui aussi associé à la continuation. En effet, Maekawa et al. (2002) observent un pourcentage plus important de H% dans le corpus d'enregistrements de présentations orales lors de conférences que dans le corpus de simulation de « speech » en public, et attribuent cette différence au caractère continuatif du BPM H%.

Yoneyama et al. (2003) ont conduit une étude sur les frontières discursives de 22 monologues issus du CSJ, et les indices prosodiques leur étant associés. Dans leur étude, elles ont défini trois différents niveaux de frontière discursive (*discourse boundary indices*, DBI) : un niveau DBI0 correspondant à la frontière entre deux phrases (*clauses*) faisant partie du même acte discursif (frontière faible), un niveau DBI1 correspondant à la frontière entre deux phrases d'actes discursifs différents mais fortement liées entre elles, ainsi qu'un niveau DBI2 de frontière entre deux phrases d'actes discursifs différents (frontière forte). Elles ont observé un pourcentage comparable de contours montants et de contours

montants-descendants (24,8 et 32,5%) aux frontières de discours jugées « faibles » (DBI0), c'est-à-dire ayant un rôle de continuation puisqu'ils n'indiquent pas la fin de l'acte discursif.

Enfin, il existe une dernière différence entre ces deux BPM : HL% est très rarement observé en parole lue (et, comme nous l'avons vu précédemment, apparaît dans des contextes plus spontanés), contrairement à H% (Venditti et al., 2008).

Koiso et al. (1998) notent également d'autres facteurs significatifs que le type de mouvement mélodique aux frontières prosodiques : la durée du dernier phonème (le plus souvent une voyelle en japonais, car seule la nasale peut apparaître en coda) est généralement plus longue en cas de maintien du tour de parole que de changement de locuteur (226 ms contre 142 ms), et une intensité décroissante est plus souvent associée à un changement de tour de parole. Ainsi, outre l'utilisation de contours montants ou montant-descendants pour indiquer une continuation ou le maintien du tour de parole, les japonophones produiraient également un allongement significatif du dernier phonème (Koiso et al., 1998) ou syllabe (Yoneyama et al., 2003).

Nous notons que les observations faites par Koiso et al. (1998) et Yoneyama et al. (2003) sont très proches, bien qu'elles portent sur un aspect différent du discours. En effet, Koiso et al. se sont intéressés aux tours de parole dans des dialogues et les indices prosodiques et syntaxiques étant associés à leur maintien ou changement, tandis que Yoneyama et al. se sont intéressées à la prosodie des frontières discursives des monologues. Dans les deux cas, nous pouvons dire qu'il s'agit de continuations ayant une valeur discursive, mais seules les continuations des dialogues auraient de surcroît une portée interactionnelle. La différence de résultat notable entre ces deux études est le rôle des contours montants. En effet, dans l'étude de Koiso et al., les contours montants H% ne sont pas associés au maintien du tour de parole et semblent au contraire plus présents lors des changements de locuteur, alors qu'ils sont plus souvent décrits aux frontières discursives faibles des monologues dans l'étude de Yoneyama et al., au même titre que les contours montant-descendants. Si, comme nous l'avons vu précédemment, les contours montants en japonais sont souvent considérés comme ayant une fonction de continuation, il est probable que lors des dialogues et en présence d'une situation interactionnelle ceux-ci n'aient pas une forte valeur continuative et soient ainsi délaissés au profit des contours montant-descendants.

## 2.4 Synthèse

Dans ce chapitre, nous nous sommes intéressés aux systèmes prosodiques du japonais et du français ainsi qu'à la place qu'occupent les contours mélodiques montants dans ces deux langues. En japonais, la montée de F0 aux frontières prosodiques non-finales ne peut avoir qu'un rôle discursif et n'apparaît donc principalement qu'en parole spontanée. La continuation du français a une fonction similaire à l'oral, comme nous l'ont montré les travaux de Portes et Bertrand (2005); en effet, les continuations majeures permettent le maintien du tour de parole. Cependant, en lecture de texte oralisée, les deux langues diffèrent grandement. En français, les frontières prosodiques sont généralement associées à des contours mélodiques montants et l'utilisation des indices prosodiques qui leurs sont associés (notamment la durée et la F0) varie selon la force de la frontière. Ils font donc partie intégrante du système phonologique français et sont également étroitement liés à la syntaxe. En japonais, par contre, les frontières prosodiques non-finales sont réalisées dans la plupart des cas par un ton bas (L%), ou tout au moins par un contour non-montant.

Bien que l'on observe des similarités entre le français et le japonais en parole spontanée avec la présence de contours montants aux frontières prosodiques, nous avons soulevé des différences notables. Tout d'abord, en japonais, ces contours montants sont facultatifs et bien moins fréquents qu'en français et, encore une fois, ce sont les tons bas (L%) qui prédominent en fin de constituant (74,59 % des cas du corpus CSJ Koiso 2014), à l'exception des frontières discursives faibles où leur présence est faiblement majoritaire (environ 60%, Yoneyama et al. 2003). Ensuite, le contour mélodique japonais ayant la plus forte valeur continue dialogique et interactionnelle est le contour montant-descendant (HL%), peu fréquent en français. Enfin, alors que les contours continuatifs montants du français sont associés à un allongement syllabique plus ou moins marqué selon le type de frontière prosodique (les continuations mineures et majeures de Delattre), les BPM montants du japonais ne sont pas nécessairement produits avec un allongement bien qu'une durée plus longue serait effectivement observée en frontière discursive forte ou lors du maintien du tour de parole.

Nous observons donc des différences très marquées entre les deux langues, qu'elles soient phonologiques ou phonétiques. De plus, les différents rôles que peuvent jouer la

continuation du français et la présence d'équivalent en situation discursive en japonais nous fait également anticiper de potentielles différences chez les apprenants entre leurs réalisations en parole lue et en parole spontanée.



# Chapitre 3

## Prosodie et contours montants en L2

Dans ce chapitre, nous proposons de faire la revue des études sur la prosodie L2. Rares sont les études s'intéressant au cas des contours continuatifs français ; c'est pour cela que nous allons faire une revue plus générale sur la prosodie en L2. Dans un premier temps, nous présentons des études sur les questions en L2 où l'on rencontre la même problématique des contours montants. Ensuite, nous présentons quelques études portant sur l'accentuation, la segmentation en unités prosodiques et le rythme en L2 – en français L2 mais également d'autres langues, notamment des études sur les apprenants francophones en L2. Les études sur l'accentuation du français L2 et les continuatifs – s'approchant ainsi particulièrement de la problématique de cette thèse – sont présentées en fin de chapitre et sont suivies par une synthèse récapitulative.

### 3.1 La prosodie des questions en L2

Un certain nombre d'études portent sur la prosodie des questions en L2 étant donné que les contours mélodiques ou l'organisation prosodique des différents types de questions peuvent poser des difficultés aux apprenants lorsque leur langue maternelle utilise des stratégies différentes.

Albin et Wang (2020) se sont intéressés à la production des questions en japonais L2 par des apprenants de différentes langues maternelles (russe, vietnamien, coréen, mandarin). Pour ce faire, ils ont fait lire à voix haute douze verbes à la forme conclusive et interrogative à 70 apprenants du japonais ainsi qu'à un groupe contrôle de 12 japonophones natifs de

Tokyo. En japonais, en l'absence de particule finale interrogative, la forme interrogative des verbes est produite par une montée de F0, tandis que la forme conclusive est descendante. Suivant l'hypothèse de transfert de la L1 vers la L2, leurs prédictions étaient les suivantes :

1. certains apprenants russophones produiraient des contours montant-descendants à la fin des formes interrogatives
2. certains apprenants mandarinophones et vietnamophones ne produiraient pas de montées finales et l'on observerait un registre mélodique plus élevé
3. la majorité des coréanophones produiraient les montées finales, comme les japonophones natifs

La moitié des verbes utilisés dans cette expérience de production étaient accentués sur la première, deuxième ou troisième more et l'autre moitié des verbes étaient inaccentués. Contre toute attente, les auteurs n'ont pas observé de transfert systématique de la L1 concernant les contours mélodiques ou l'empan de F0. Alors qu'ils s'attendaient à observer moins de contours montants chez les mandarinophones et les vietnamophones, ils n'ont noté l'absence de montées mélodiques qu'entre 3,8 à 10,3 % des cas selon le type de phrase chez les vietnamophones. Chez les mandarinophones, ces pourcentages s'élèvent entre 11,7 à 22,5% des cas selon le type de phrase, ce qu'ils considèrent comme « assez fréquent » mais qui n'est cependant pas la réalisation majoritaire. Si l'hypothèse de transfert de la L1 n'a pas vraiment pu être validée par leurs résultats, les apprenants n'avaient pas pour autant une prosodie native dans leur réalisation des questions. Les auteurs observent notamment dans tous les groupes d'apprenants des difficultés à produire la forte montée de F0 attendue en fin de phrase non-accentuée (« *tripling* »), ainsi que des difficultés à réaliser les courbes mélodiques appropriées.

Verdugo et Trillo (2005) se sont quant à eux intéressés aux « questions tag » de l'anglais produites par des apprenants hispanophones. Il s'agit par exemple (issu du corpus utilisé par Verdugo et Trillo) de « *It's cold today, isn't it?* », la question tag « *isn't it?* » pouvant se traduire par « n'est-ce-pas? » et pouvant être réalisée par un contour mélodique descendant lorsqu'il s'agit de demander la confirmation de l'interlocuteur, le locuteur s'attendant à ce que celui-ci partage son avis ou confirme ses propos. Un contour montant peut tou-

tefois être réalisé lorsque le locuteur n'est pas sûr de ce qu'il avance ou du fait que l'interlocuteur partage le même avis que lui. Ainsi, ces « questions tag » peuvent être réalisées par un contour montant ou descendant selon le contexte pragmatique. En espagnol, selon Verdugo et Trillo, les seules « questions tag » existantes ont toujours un rôle de confirmation et sont systématiquement réalisées par un contour montant. Afin d'étudier la gestion prosodique de ces deux types de questions par les hispanophones, ils ont enregistré 10 apprenants hispanophones et 10 anglophones natifs lisant par deux de courts dialogues. Leur analyse des contours produits à la fin de ces questions montre que les apprenants hispanophones réalisent systématiquement ces questions par une montée mélodique. Les auteurs attribuent ces résultats au transfert de leur L1 et à la surgénéralisation des contours montants en ne distinguant pas les deux contextes pragmatiques par l'intonation.

Dans leur étude sur les apprenants germanophones de l'anglais, Puga et al. (2017) ont étudié leur perception des patrons intonatifs : phrases déclaratives, continuations, questions totales « yes/no », « questions tag », sarcasme, etc. 20 apprenants germanophones et 25 anglophones natifs ont participé à une expérience de perception où il leur était demandé de sélectionner le patron intonatif approprié parmi cinq propositions (descendant, montant, montant-descendant, descendant-montant et plat) pour 28 phrases issues d'une courte histoire. Ils avaient la possibilité de réécouter les différents patrons mélodiques (prononcés par une locutrice anglophone native) et avaient accès à une version écrite du texte qu'ils avaient lue au préalable. L'analyse statistique des réponses obtenues montre que les apprenants ont globalement moins bien réussi l'expérience que les natifs, mais que cela dépendait également du type de phrase. En effet, les apprenants ont obtenu des résultats similaires aux natifs pour les *statement questions*, les questions WH, les questions totales « yes/no » et les questions d'écho, et même de meilleurs résultats que les natifs pour une des trois formes de continuation et pour une condition de sarcasme. Mais, comme pour les apprenants hispanophones de l'étude de Verdugo et Trillo, les apprenants germanophones ont montré des difficultés avec les « questions tag ». En effet, ils ont généralement choisi un contour montant pour les questions tag de confirmation demandant un contour descendant, et un contour descendant ou montant-descendant pour les questions tag normalement produites par une montée mélodique. Ici encore ces difficultés seraient liées à la L1 des apprenants, et au fait qu'il n'y ait pas d'équivalent aux questions tag de l'anglais

en allemand.

Santiago (Santiago 2014 ; Santiago et Delais-Roussarie 2012, 2015) s'est aussi intéressé à la réalisation des questions, chez les apprenants hispanophones du français. En français, il existe des questions totales et des questions partielles. Parmi les questions totales, Santiago en distingue trois catégories :

1. les questions totales déclaratives (exemple : « *Tu lis?* »)
2. les questions totales avec inversion du sujet (« *Lis-tu?* »)
3. les questions totales utilisant la locution « est-ce que » (« *Est-ce que tu lis?* »)

Les deux dernières catégories de questions totales – morphosyntaxiquement marquées – sont analysées ensemble. Pour les questions partielles, deux catégories sont retenues :

1. les questions partielles avec marqueur interrogatif antéposé (« *Que lis-tu?* »)
2. les questions partielles avec marqueur interrogatif *in situ* (« *Tu lis quoi?* »)

Il faut noter que les réalisations prosodiques des questions peuvent varier et que seules les questions totales déclaratives (« *Tu lis?* ») sont obligatoirement montantes (Di Cristo, 2016), et qu'elles se distinguent ainsi prosodiquement des questions totales morphosyntaxiquement marquées.

Santiago et Delais-Roussarie (2015) ont enregistré 35 locuteurs : 15 apprenants hispanophones du français, 10 francophones natifs et 10 hispanophones enregistrés dans leur L1 comme groupes contrôles. Le corpus issu du protocole COREIL (Delais-Roussarie & Yoo, 2011) comporte une tâche de lecture de textes et de dialogues ainsi qu'une tâche de jeu de rôle où les locuteurs doivent compléter un formulaire en posant des questions à leur interlocuteur. L'analyse des contours mélodiques produits à la fin des questions totales a montré que la quasi-totalité des contours produits par les locuteurs hispanophones étaient montants ou extra-montants, à proportions plus ou moins égales, en français L2 et en espagnol L1. Les francophones, eux, ont produit une proportion plus faible de contours extra-montants que les hispanophones et ont produit des contours plats ou descendants lorsque les questions totales étaient marquées morphosyntaxiquement, ce qui n'est pas le cas dans les productions des apprenants où ces contours sont quasi inexistantes. En ce qui concerne

les questions partielles, les francophones ont majoritairement utilisé des contours non-montants. En espagnol L1, c'est à peu près la moitié des questions partielles qui sont réalisées avec un contour non-montant et l'autre moitié avec un contour montant. Cependant, dans les réalisations des apprenants hispanophones en français, on observe une bien plus grande proportion de contours montants et extra-montants, avec en particulier un pourcentage de 40% chez les apprenants contre 4,5% chez les francophones et 25% chez les hispanophones L1 pour les questions partielles. Ce pourcentage s'élève à 60% pour les questions *in-situ* chez les apprenants contre 5% chez les francophones natifs. Ainsi, bien que les premiers résultats obtenus pour les questions totales puissent s'expliquer par un transfert de la L1 vers la L2, ce n'est pas le cas de ceux obtenus pour les questions partielles où l'on observe une grande proportion de contours non-montants en espagnol (comme en français L1) mais qui est réduite dans les productions en français L2. Ce faible pourcentage de contours non-montants se fait au profit de contours montants et notamment extra-montants, rarement produits chez les francophones.

Kamiyama et Sakamoto (2010) se sont intéressés à la production des questions totales en français par les apprenants japonophones. Ils ont analysé la production de 5 courtes questions par 8 apprenants japonophones et 3 locutrices francophones (uniquement 4 phrases pour les francophones). L'analyse des résultats montre que la plupart des apprenants produisent une montée finale moins ample que les francophones et qu'ils ne produisent pas la déclinaison de F0 souvent observée chez les natifs. Les auteurs ont ensuite voulu tester la perception de la prosodie des japonophones pour les questions totales. Ils ont utilisé des stimuli resynthétisés en transférant la durée et la F0 non-natives de 8 stimuli d'apprenants japonophones et de 2 francophones sur des stimuli produits par 2 locuteurs francophones. Les auditeurs (25 francophones natifs) ont bien perçu la différence question/continuation pour les stimuli resynthétisés à partir de la prosodie des francophones et de 3 apprenants japonophones. Les stimuli issus de la resynthèse d'un des apprenants sont jugés comme « continuations » (à un pourcentage comparable aux réponses « continuations » pour les continuations des francophones) et pour quatre autres apprenants les résultats se situent entre les deux.

### 3.2 Accentuation, organisation prosodique et rythme en L2

La première étude que nous présentons dans cette section est celle de Dupoux et al. (2008) portant sur la perception de l'accent chez les apprenants francophones de l'espagnol. Ils présentent deux expériences de perception menées sur 39 francophones apprenant l'espagnol, 20 francophones et 20 hispanophones natifs comme groupe contrôle. La première expérience utilise des séquences de quatre non-mots (de deux items différents) où les participants, après avoir appris à associer chaque item à un chiffre, doivent retranscrire la séquence entière (par exemple « 1211 » où 1 et 2 sont deux non-mots qui ne diffèrent que par l'emplacement de l'accent). Ils observent près de 3 fois plus d'erreurs chez les apprenants et les francophones natifs que les hispanophones du groupe contrôle, et n'observent pas d'amélioration selon le niveau de langue des apprenants. Dans la deuxième expérience, les mêmes participants étaient exposés à des stimuli correspondant à des non-mots ou à des mots existant réellement en espagnol. La consigne qui leur a été donnée est de juger le plus rapidement possible si chaque stimuli est un non-mot ou un mot réel. Des paires minimales non-mot / mot réel différant selon la position de l'accent et d'autres selon les phonèmes (comme contrôle) ont été intégrées dans les stimuli. Les résultats de cette expérience montrent que le taux d'erreur pour les stimuli différant selon la position accentuelle est de 24 % contre seulement 6% pour les stimuli de la condition contrôle. La fréquence des mots utilisés dans les deux conditions ne permet pas d'expliquer cette différence de résultat. Les auteurs affirment que les résultats obtenus dans ces deux études confirment les difficultés perceptives des francophones (qualifiés par les auteurs de « surdit  accentuelle ») et que les meilleurs r sultats obtenus dans certaines exp riences ant rieures s'expliquent par un protocole diff rent (exp riences de perception avec une charge cognitive moindre) : « *We conclude that the good performance of French listeners in some of the experiments in Dupoux et al. is due to the fact that in the experiments, subject could use an acoustic representation rather than an abstract phonological one* » (p.19).

Schwab et Llisterri (2012) se sont aussi int ress s   la perception de l'accent espagnol par les francophones. Le protocole de leur exp rience perceptive comporte deux phases :

- une premi re phase o  les participants doivent associer des triplets de non-mots (par exemple : **l **guido / luguido / luguid )   des formes visuelles. Un non-mot leur

est présenté auditivement et ils doivent choisir une forme parmi quatre présentées visuellement. Ensuite, un *feedback* leur est donné et les distracteurs (les trois autres formes) disparaissent de l'écran.

- une deuxième phase de test sur les mêmes non-mots qu'ils ont précédemment appris mais ayant subi des modifications acoustiques (durée, amplitude et F0). Les stimuli ont été modifiés de sorte que les indices acoustiques de chaque voyelle du stimulus équivalent mais dont l'accent est plus à droite soient transférés sur le stimulus de base. Ainsi, les valeurs de durée, d'amplitude et/ou la courbe mélodique des stimuli oxytons ont été transférées sur le stimulus de base paroxyton (P>O) et les valeurs de stimuli paroxytons sur les stimuli de base propoxyton (PP>P). Cela correspond à un déplacement accentuel vers la droite (potentiellement perçu ou non car tous les indices prosodiques ne sont pas modifiés selon les conditions ; par exemple dans la condition de durée, seule la durée est modifiée et non pas l'amplitude et la F0).

Le but de cette expérience est d'examiner quel(s) indice(s) acoustique(s) les francophones utilisent pour percevoir l'accent en espagnol, s'ils sont bien capables de le percevoir et de garder en mémoire les triplets de non-mots. 22 apprenants francophones de l'espagnol ainsi que 22 hispanophones natifs ont passé l'expérience. Les résultats montrent que les non-natifs sont capables de réaliser la tâche (73,94% de réponses correctes contre 90,5% pour le groupe contrôle d'hispanophones natifs). Les non-natifs montrent même de meilleurs résultats lorsque seule l'amplitude ou la durée ont été manipulées et lorsque la durée et l'amplitude ont été manipulées simultanément. Ils ont également plus souvent perçu le déplacement accentuel lorsque la F0 était manipulée dans les stimuli P>O (sur la dernière syllabe) tandis que les hispanophones natifs l'ont plus souvent perçu dans les stimuli PP>P (syllabe pénultième). De manière générale, les francophones ont mieux perçu les modifications portant sur la dernière syllabe. Les auteurs concluent que les francophones ne perçoivent pas de la même manière l'accent en espagnol que les locuteurs natifs ; ils semblent notamment être plus sensibles à la manipulation de l'amplitude et de la durée. Pour la durée, il pourrait s'agir d'un transfert de la L1 puisque c'est le principal indice prosodique de l'accent final français. De même, la sensibilité accrue aux stimuli P>O comportant des modifications sur la dernière syllabe peut être interprétée par un phénomène

de transfert positif. Cependant, la sensibilité à l'amplitude ne pourrait s'expliquer par un potentiel transfert de la L1 puisque ce n'est pas un indice robuste dans l'accentuation du français. Leur hypothèse pour expliquer ce résultat est que les hispanophones sont moins sensibles aux variations d'amplitude, naturellement plus présentes dans leur L1, que les locuteurs francophones.

Les travaux de Bartkova et al. (2012) portent sur la production des contours continuatifs chez des francophones en français L1 et en anglais L2. Un corpus lu de 40 courtes phrases a été utilisé pour le recueil des données, faisant partie des 8 différentes catégories syntaxiques suivantes :

1. « CAP » : configuration continuative à la fin de la première clause dans une phrase comportant deux clauses n'étant pas reliées par une conjonction de coordination (exemple : « Il dort chez Maria, il va finir tard. »)
2. « CAO » : configuration continuative à la fin de la première clause dans une phrase comportant deux clauses reliées par une conjonction de coordination (exemple : « Il dort chez Maria car il va finir tard. »)
3. « CIS » : configuration continuative à la fin d'un groupe nominal sujet (exemple : « Les agneaux ont vu leur mère. »)
4. « CIA » : configuration continuative à la fin d'un groupe nominal sujet dans la première clause d'une phrase comportant deux clauses. (exemple : « Nos amis aiment Nancy parce que c'est joli. »)
5. « QAS » : Question à la fin d'une clause (exemple : « Il dort chez Maria? »)
6. « QIS » : Construction interrogative avec un groupe nominal. (exemple : « Qui a appelé? Nos amis? »)
7. « DIS » : Phrase déclarative courte (exemple : « Nos amis. »)
8. « DAS » : Phrase déclarative plus longue (exemple : « Il dort chez Maria »)

Les mêmes phrases ont été traduites en anglais et 32 francophones natifs ont été enregistrés pour le corpus en français et en anglais L2, ainsi que 20 anglophones natifs dans leur

L1 en tant que groupe contrôle. Les résultats obtenus révèlent que les francophones utilisent plus de contours montants que les anglophones, en particulier dans les phrases CI où une continuation « mineure » est attendue. Les anglophones, quant à eux, produisent plus de contours descendants, en particulier dans les phrases CA (correspondant à une continuation « majeure »). De plus, les apprenants francophones gèrent différemment l'indice de durée en produisant en anglais L2 des voyelles plus longues que les anglophones natifs sur la syllabe accentuée dans les phrases CI, et plus courtes dans les phrases « QAS », « QIS », « DAS ». La durée des voyelles accentuées produites en L2 diffère néanmoins de celles produites en français L1 et elles tendent vers les valeurs des anglophones natifs.

Dans sa thèse, Horgues (2010) s'est intéressée à la prosodie des apprenants francophones de l'anglais. Elle a cherché dans un premier temps à confirmer l'importance de la prosodie dans la perception de l'accent français et a monté un test de perception de phrases resynthétisées, monotonisées (variations de F0 neutralisées) et filtrées par un filtre passe-bas. 18 anglophones natifs ont ainsi jugé les phrases de trois groupes de locuteurs (anglophones, francophones, et d'une autre L1) à l'aide d'une échelle comportant accent français « fort », « modéré », « léger » ainsi que « natif » ou « accent d'une autre langue ». Les résultats obtenus révèlent l'importance des indices prosodiques qui à eux seuls suffisent aux auditeurs pour percevoir l'accent français, en particulier les indices rythmiques. En effet, le taux de réponses correctes identifiant les stimuli français atteint 75% pour les stimuli filtrés (prosodie uniquement perçue), 78% pour les stimuli monotonisés (durée et segments non natifs mais F0 neutralisée) mais « seulement » 53% pour les stimuli uniquement resynthétisés (F0 non native, transférée sur des segments natifs). Une autre expérience, cette fois-ci de production (lecture de dialogue), montre que les locuteurs francophones ont un débit plus lent mais aussi une moindre isochronie que les anglophones, et ont des difficultés pour réduire la durée des syllabes inaccentuées par rapport aux syllabes accentuées. Au niveau mélodique, ils ont tendance à produire des montées mélodiques en frontière droite des unités intonatives non-finales, comme en français. Ces résultats peuvent donc être attribués à un transfert négatif de la L1 vers la L2.

Duběda (2009) a étudié les réalisations de l'accent initial du français par des apprenants tchécophones. En tchèque, l'accent est toujours sur la première syllabe des mots accentuables. En français il existe aussi un accent initial (cf. 2.1), mais facultatif. Tandis

que l'accent initial en français est principalement produit par une hausse mélodique (sans être associé à un allongement, contrairement à l'accent final), l'accent initial du tchèque n'est pas réalisé par une montée mélodique mais par des configurations L\*H ou HL\* où la syllabe accentuée est plus basse que la syllabe la précédant. La syllabe accentuée est possiblement suivie ou non d'une montée mélodique et n'est pas marquée par la durée, si ce n'est qu'elle a tendance à être plus courte que la syllabe précédente. L'accent initial étant différent phonologiquement et phonétiquement en français et en tchèque, Duběda a donc enregistré 12 étudiants tchèques de niveau B1 en français et 4 francophones natifs effectuant une tâche de lecture de texte afin d'étudier les réalisations des apprenants et de confirmer son intuition qu'ils ont tendance à réaliser cet accent « à la tchèque ». Il a ensuite séparé les apprenants en trois groupes de niveau de langue (a), (b), (c), du niveau le plus élevé avec la meilleure fluence au groupe ayant le plus de difficultés. Il a également annoté les réalisations des apprenants et fait confirmer cette annotation en confrontant l'annotation (de 16 phrases prises au hasard) à celle réalisée par deux étudiants tchèques en phonétique et en français. En moyenne, 37,2 accents initiaux chez les tchèques et 35,5 accents initiaux chez les francophones ont été relevés. Aucune différence de durée ni d'intensité n'a été observée entre les groupes d'apprenants et les francophones du groupe contrôle. L'analyse des accents initiaux produits dans les deux groupes montre que les tchécophones accentuent parfois certains mots qui ne sont jamais accentués par les francophones et ont tendance à suraccentuer les mots dissyllabiques. Les résultats montrent que les apprenants tchécophones ayant le plus bas niveau de langue (groupe (c)) utilisent significativement plus d'accents mélodiques typiques du tchèque (S\*, L\*H et HL\*) en position initiale que les francophones du groupe contrôle (29% contre 13% des cas). Duběda a ensuite réalisé une expérience de perception afin de confirmer entre autres le fait que ces contours mélodiques soient bien perçus comme non-natifs lorsqu'ils sont réalisés en français. Pour cela, il a transféré la prosodie d'énoncés non-natifs sur une base segmentale de locuteurs natifs. Un total de 12 phrases à prosodie non-native et de 8 phrases à prosodie native ont été créées. Les trois groupes de langue (a, b, c) sont représentés dans les phrases non-natives et 7 contours mélodiques « à la tchèque » sont présents. 16 francophones natifs étudiants en linguistique ont ensuite jugé l'intonation et le rythme de ces énoncés en soulignant les mots ayant une prosodie leur paraissant « peu naturelle », « non

française » ou « fausse ». Les résultats montrent que les mots porteurs d'accent initial ne sont pas plus souvent jugés non-natifs que les autres, mais que les accents mélodiques précédemment annotés et interprétés comme non-natifs et typiques du tchèque (S\*, L\*H et HL\*) ont été identifiés comme problématiques dans 18,7% des cas, contre 14,9% des accents initiaux dans les phrases à prosodie non-native. Ce pourcentage s'élève à 28,2% des cas pour les phrases utilisant la prosodie de locuteurs du groupe de niveau de langue (c). L'auteur conclut qu'il existe un phénomène de transfert négatif dans la distribution des accents initiaux en français L2 (tendance à suraccentuer les mots dissyllabiques) ainsi que dans leur réalisation (plus grande proportion d'accents mélodiques « à la tchèque », perçus comme non-natifs par les francophones). Nous notons également que la bonne gestion de la durée et de l'intensité peut être considérée à l'inverse comme un transfert positif du tchèque vers le français, étant donné que l'accent tchèque n'est pas non plus marqué par ces indices prosodiques.

Marchand et Le Gac (2020) se sont intéressés à la production des accents initiaux et finaux en français L2 par des apprenants turcophones. L'accent initial du français est un accent facultatif, qui est principalement produit par une hausse mélodique et n'est pas associé à un allongement (cf. 2.1). En turc, selon Marchand et le Gac, il existerait deux types d'accents primaires (A1) :

- un accent régulier, final et démarcatif
- un accent lexical, libre et distinctif

L'accent lexical du turc peut ainsi se trouver sur la première syllabe d'un mot prosodique, comme l'accent initial du français (A2). Il existe également un accent secondaire en turc (A2), facultatif et généralement en début de mot mais qui peut aussi être placé en fin de mot lorsqu'il est relativement éloigné de l'accent primaire. Les accents primaires et secondaires du turc sont uniquement produits par des variations de F0 (un accent mélodique H\* ou L\*), la durée n'étant pas un paramètre pertinent pour la réalisation de ces accents. L'accent secondaire turc est réalisé par une hauteur mélodique plus basse que l'accent primaire. De grandes différences phonologiques et phonétiques existent donc entre le turc et le français. Ainsi, les auteurs se sont intéressés à la gestion de l'accentuation du français par trois locutrices turcophones (deux locutrices niveau A1, une locutrice niveau B2/C1) lors d'une

tâche de répétition de 24 phrases déclaratives produites par un locuteur francophone. Les résultats obtenus sont un bon emploi des accents initiaux et finaux en français mais des réalisations phonétiques différentes avec des allongements plus marqués sur l'accent final que chez le locuteur francophone, des hausses de F0 moins élevées et un pic mélodique sur la syllabe accentuée plus tardif. Les auteurs notent en particulier des difficultés à gérer l'allongement et les variations de F0 selon le type de frontière prosodique.

Kaglik et Boula de Mareüil (2010) ont mené une étude sur la prosodie des apprenants polonais en français L2. Ils ont enregistré 5 francophones natifs, 5 monolingues polonais (en polonais), 3 apprenants polonais de niveau avancé ainsi que 27 bilingues vivant en France dans une tâche leur demandant de raconter un épisode de film muet. Des stimuli de synthèse vocale basés sur des extraits des phrases produites et transférant leur structure prosodique ont ensuite été créés, puis leur prosodie (ainsi que les stimuli originaux) a été jugée par 8 spécialistes d'une échelle allant de 0 à 4 (prosodie perçue comme native à non native). Les auteurs ont ensuite analysé la production des 6 apprenants ayant reçu le score le plus bas (les 3 apprenants et 3 des bilingues) et ils ont observé que ceux-ci avaient tendance à produire plus de périodes intonatives ne comportant qu'un seul groupe intonatif ainsi qu'un pourcentage plus élevé de groupes intonatifs segmentés à leur droite que les natifs. De plus, ces groupes intonatifs sont plus souvent marqués par un contour descendant et un allongement vocalique et fréquemment suivis d'une pause, ce qui est peu fréquent chez les natifs. Selon Kaglik et Boula de Mareüil, le transfert de la L1 ne permet pas d'expliquer ces résultats car le taux de segmentation est plus élevé que chez les polonais monolingues. Ils attribuent donc ce phénomène de sursegmentation aux difficultés cognitives dans la formulation des énoncés en L2.

Gabriel et al. (2014) ont étudié le rythme d'apprenants du français bilingues en allemand et en mandarin, et d'apprenants monolingues allemand. De par les similarités existantes au niveau rythmique entre le mandarin et le français, les auteurs prédisent une plus grande facilité pour les apprenants plurilingues à acquérir le rythme du français par rapport aux apprenants monolingues. Le protocole expérimental comporte une lecture de texte oralisée, 13 phrases ne comportant que des syllabes CV, ainsi que 10 non-mots insérés dans des phrases porteuses. 13 apprenants plurilingues ont été enregistrés en mandarin, français et allemand, 10 monolingues allemands ont été enregistrés en allemand et en fran-

çais, 10 monolingues mandarinophones enregistrés en français et en mandarin ainsi que 10 francophones natifs en français uniquement en tant que groupe contrôle. Les apprenants plurilingues ont également participé à différentes tâches extra-linguistiques : une interview, faire une comparaison commentée entre un enregistrement en L1 et en L2, ainsi que lire des phrases en français et commenter rétrospectivement leur prononciation. Les auteurs ont analysé les productions enregistrées en utilisant les mesures %V (Ramus et al., 1999) et VarcoV correspondant respectivement à la proportion des intervalles vocaliques dans l'énoncé et l'écart type de la durée des intervalles vocaliques normalisés pour prendre en compte les différences de débit de parole entre les locuteurs. Les résultats montrent que les mandarinophones monolingues obtiennent des valeurs plus proches de celles des francophones que ne le sont celles obtenues par les germanophones monolingues. Les résultats du groupe test (les apprenants bilingues allemand et mandarin) sont plus contrastés. En effet, si certains apprenants montrent bien des valeurs plus proches de celles des francophones que les apprenants germanophones monolingues, ce n'est pas le cas de tous les apprenants de ce groupe. L'hypothèse d'un transfert rythmique positif du mandarin sur le français chez les apprenants bilingues n'est donc pas entièrement validée. Cependant, selon les auteurs, la diversité des productions des apprenants bilingues peut s'expliquer par les facteurs extra-linguistiques, mis en évidence par les interviews et tâches annexes réalisées par les apprenants, comme leur intérêt pour l'allemand ou le français ou leur sensibilité à la phonologie des langues.

Pillot-Loiseau et Xie (2018) se sont intéressées au rythme des apprenants mandarinophones du français. Pour cela, elles ont dans un premier temps utilisé le même type de mesures que l'étude détaillée précédemment –  $V\%$ ,  $\Delta V$ ,  $\Delta C$  (Ramus et al., 1999),  $\text{Varco}\Delta V$  et  $\text{Varco}\Delta C$  (Dellwo 2006) ainsi que  $rPVI-C$  et  $nPVI-V$  (Low et al. 2000, Grabe et Low 2002) – afin d'analyser les différences rythmiques entre les productions des francophones et des apprenants. La mesure  $rPVI-C$  correspond à la variabilité brute des intervalles vocaliques et la mesure  $nPVI-V$  à la variabilité des intervalles vocaliques normalisée. Dix phrases lues par dix apprenants mandarinophones (de niveau de langue supérieur ou égal à B2) et cinq francophones natifs ont été enregistrées en mandarin pour les apprenants et en français pour les deux groupes. Les résultats de l'analyse rythmique vont à l'encontre de l'hypothèse de transfert ; en effet, aucune des valeurs extraites en français L2 ne se confond avec

celles obtenues pour le mandarin L1 et elles sont même plus proches des valeurs obtenues par le groupe contrôle de francophones. Seule la mesure %V du français L2 est plus proche du mandarin L1 que du français L1, tout en restant significativement différente de l'un et de l'autre. Après cette première étude acoustique, les auteures ont souhaité tester la perception du rythme des apprenants mandarinophones par un test de discrimination AXB où chacun des 20 triplets comporte une phrase délexicalisée du mandarin L1, français L2 et français L1. Les consonnes des stimuli ont été remplacées par des /s/ et les voyelles par des /a/ et les variations mélodiques neutralisées en ajustant la fréquence fondamentale à 150 Hz. 9 francophones ont participé à l'expérience. Les résultats montrent que les francophones ont plus souvent associé les stimuli en français L2 aux stimuli du mandarin L1 que du français L1 (dans 60,8 % des cas). Il semblerait donc qu'au niveau perceptif le rythme du français L2 serait plus proche de celui de la langue maternelle des apprenants que de celui de leur langue cible, ce qui est en contradiction avec les résultats de l'étude acoustique qui n'a pas montré de grandes différences rythmiques entre le français L2 et L1, ni de similitudes entre le mandarin L1 et le français L2. Face à ces résultats, les auteures remettent en doute la validité des mesures rythmiques pour les études de transfert en L2, mais évoquent également certains points du protocole (durée des /s/ conservée dans les stimuli qui peuvent être à l'origine de différences perçues au niveau segmental, différence de longueur des énoncés, niveau de langue relativement avancé des apprenants, interférence potentielle de l'anglais L2...) qui pourraient expliquer ces résultats contradictoires et/ou l'absence d'un transfert rythmique quantifiable.

Dans son étude, Schwab (2012) a cherché à savoir si les apprenants hispanophones transféraient les propriétés accentuelles de leur L1 sur le français L2. Pour ce faire, elle a créé une expérience de production avec des non-mots trisyllabiques français dans des phrases porteuses, en position nominale en frontière droite de AP (exemple : « un certain poutila piquait tous les touristes ») ou en position pré-nominale (exemple : « un poutila pic était très utile »). 18 non-mots CV (exemple : « poutila ») et 18 non-mots CVC finissant par /R/ ou /l/ (exemple : « poutila ») ont été intégrés dans un total de 72 phrases. L'auteure a enregistré la production des phrases de l'expérience ainsi que des non-mots isolés par 3 apprenants hispanophones de niveau avancé (B2/C1) et de 3 francophones natifs. La durée des voyelles, l'amplitude et la F0 ont été analysées. Les résultats révèlent que les ap-

prenants hispanophones sont capables de réaliser l'accent final en français en manipulant les indices de durée et de fréquence fondamentale, sans utiliser l'amplitude, comme les francophones natifs. Ils montrent également que les apprenants hispanophones tendent à produire la dernière syllabe des non-mots avec une plus longue durée (en particulier pour les mots CV) et une fréquence fondamentale plus élevée. Cela pourrait s'expliquer par le fait que les hispanophones suraccentuent la dernière syllabe des non-mots en position adjectivale (pas à la frontière droite d'un AP, par exemple « un poutila pic était très utile ») car ils n'auraient pas acquis le fait que le domaine d'accentuation du français est l'AP, et non pas le mot comme en espagnol. Malgré les différences existantes entre les productions des hispanophones et des francophones, un transfert direct de la L1 n'est pas observé. Selon Schwab, si les apprenants hispanophones avaient transféré leur L1, ils auraient produit différemment les stimuli CV des stimuli CVC et la seconde syllabe des non-mots trisyllabiques CV aurait été réalisée avec une plus longue durée. Les apprenants semblent avoir intégré dans leur interlangue la connaissance de l'accent final fixe en français bien qu'ils aient encore des difficultés à l'associer au domaine de l'AP.

Barquero-Armesto (2012) a aussi travaillé sur l'interlangue des apprenants hispanophones du français et certains points de son analyse rejoignent celle de Schwab. Elle a enregistré 4 apprenants hispanophones de niveau avancé (B1/C1) et 4 locuteurs francophones comme groupe contrôle lisant un extrait du texte « Le Premier Ministre ira-t-il à Beaulieu ? » (Durand et al., 2002). La position des accents, le nombre de groupes accentuels (AP) et phoniques (PG), la présence de pauses et leur durée ainsi que la durée des syllabes ont été analysés. Comme dans l'étude de Schwab, les apprenants hispanophones ne montrent pas de difficultés à produire l'accent du français sur la dernière syllabe. En effet, c'est le cas dans 96% des accents produits par les hispanophones contre 99,99% chez les francophones du groupe contrôle. L'auteure observe également plus de AP dans les groupes phoniques chez les apprenants hispanophones (77 en moyenne contre 67 chez les francophones) ainsi qu'un plus grand nombre de pauses (en moyenne 46,25 contre 22). La durée des pauses est toutefois comparable entre les deux groupes de locuteurs. Ainsi, les hispanophones placent correctement l'accent du français mais ont tendance à « sursegmenter » en produisant de courts *chunks*. Aussi, les apprenants hispanophones utilisent rarement la prééminence mélodique seule pour marquer les frontières prosodiques et ont

tendance à les réaliser par une pause. En effet, seules 45% des frontières prosodiques sont uniquement marquées par l'accent final chez les apprenants contre 74% chez les francophones du groupe contrôle. Nous notons cependant qu'aucune analyse de la F0 ni de la durée des segments en frontière prosodique n'a été effectuée et il est possible que d'autres difficultés au niveau prosodique puissent être relevées.

Lee (2014) a mené une étude sur la production des contours montants du français en fin d'AP chez les apprenants anglophones. Il a recueilli des données à l'aide d'une tâche d'élicitation d'énoncés de manière guidée (« *Tu as acheté de la glace à la vanille et à la noisette pour ta fête. Demande aux invités s'ils veulent de la glace à la vanille ou à la noisette* » -> réponse attendue : « Tu veux de la glace à la vanille ou à la noisette? ») et une tâche de narration d'un événement récent social auquel ils ont participé. 8 étudiants américains anglophones ont été enregistrés ainsi que 4 francophones natifs. Plusieurs points ont pu être observés dans les productions des apprenants :

1. les apprenants anglophones produisent plus de frontières de AP que les francophones natifs
2. certains apprenants anglophones réalisent le « coude » (ton L avant la montée mélodique) du contour montant plus tardivement que les francophones tandis que certains apprenants, à l'inverse, réalisent la montée mélodique et le ton L plus tôt que les francophones natifs.

Concernant le premier point, l'auteur propose deux explications possibles ; la première étant que les apprenants produisent des accents agrammaticaux supplémentaires et la deuxième étant que les apprenants ont tendance à parler plus lentement et avec moins d'aisance que les locuteurs du groupe contrôle. Le deuxième point observé pourrait lui aussi s'expliquer par deux hypothèses différentes ; la première étant que les apprenants percevraient les contours montants du français sans reconnaître l'alignement tonal et la réalisation précise de ces contours, la deuxième étant que les contours descendant-montants (L\*+H, *scooped accents*), qui existent en anglais et traduisent l'incertitude ou l'hésitation, seraient utilisés en français L2.

Grover et al. (1987) ont mené différentes expériences sur les continuations en français, anglais et allemand. Dans la première expérience, ils ont cherché à confirmer les différences

dans les pentes mélodiques (*slopes*) des continuations produites dans les trois langues (montantes en français et en allemand mais avec une montée initiale plus « abrupte » en français et descendantes en anglais) en enregistrant un locuteur natif pour chaque langue lisant à voix haute des phrases comportant deux *clauses*; par exemple « J'ai vu Anne mais pas Eric ». Leur analyse des pentes réalisées aux continuations (sur « Anne » dans l'exemple donné) confirment ces différences. Suite à cette première expérience, les auteurs ont enregistré de jeunes monolingues francophones et anglophones dans leur langue ainsi que de jeunes anglophones en immersion en France dans quatre groupes d'âge allant de 6 à 24 ans. La tâche utilisée pour cette deuxième expérience était une tâche de parole spontanée utilisant des images comme support, et seules les phrases de constructions similaires aux stimuli de la première expérience ont été analysées. Les auteurs ont observé que les anglophones en immersion en France réussissaient bien à produire les contours continuatifs avec des pentes mélodiques semblables à celles des francophones natifs, en particulier chez les locuteurs les moins âgés. A partir de 16 ans, les pentes mélodiques produites sont ainsi plus proches de celles des anglophones monolingues dans leur L1 que des francophones natifs, révélant ainsi la présence de transfert négatif de leur L1. Dans une troisième expérience, les auteurs ont cherché à tester la capacité des locuteurs anglophones et francophones à reconnaître les productions prosodiques non-natives des contours continuatifs. Pour ce faire, ils ont resynthétisé la F0 des stimuli de la première expérience afin de créer des stimuli « non-natifs » où la prosodie du continuatif de l'autre langue a été calquée sur un stimulus natif. Ils ont ensuite demandé aux mêmes participants de la deuxième expérience de choisir le stimulus perçu comme étant le plus naturel parmi une paire de stimuli resynthétisés ayant une prosodie native et non-native. Cependant, les résultats obtenus par les groupes contrôles ne dépassent pas le seuil de chance, bien que les locuteurs aient compris que la différence entre les stimuli était intonative. Cela révèle ainsi que les locuteurs anglophones et francophones ne perçoivent pas un stimulus comme natif ou non uniquement par la prosodie de son continuatif. Ils évoquent par ailleurs la difficulté de la tâche : « *it was extremely hard to prefer one version of a sentence to another, for they could imagine acceptable contexts for both.* » (p.290).

Xi et al. (2020), dans leur étude pilote, se sont intéressés à l'accentuation du français L2 par les apprenants mandarinophones, et en particulier des accents finaux en frontière

droite de AP. Le mandarin a un système accentuel très différent du français puisque seules les frontières de *prosodic phrase* (niveau intermédiaire) sont associées à un allongement important et que le rehaussement mélodique est un indice prosodique important permettant de distinguer les unités rythmiques. De par ces différences au niveau prosodique, les auteurs se sont intéressés à la gestion par les mandarinophones des proéminences accentuelles. Ils ont enregistré 7 mandarinophones apprenants du français (1 apprenant niveau B1, 3 niveau B2, 2 niveau C1 et 1 niveau C2) ainsi que 6 francophones lisant 6 phrases en français à voix haute. L'analyse de la durée montre que les apprenants mandarinophones ont produit un allongement syllabique en fin de phrase ainsi qu'en frontière de AP, pouvant être plus long que ceux des francophones. La durée d'allongement est d'autant plus importante que le niveau de langue des apprenants est bas. Concernant les variations mélodiques, les apprenants ayant un niveau de langue bas ont produit des contours descendants aux frontières prosodiques non-finales, par exemple à la fin de la première partie « Il a parlé du paysage » dans la phrase « Il a parlé du paysage, et de toute autre chose. », qui correspond à une frontière de IP selon nos critères théoriques. Cela n'a pas été observé chez les francophones du groupe contrôle, et si une telle réalisation reste possible en français elle paraît en effet ici peu naturelle. Enfin, l'analyse des pauses a montré la présence de pauses importunes dans les groupes de sens comme dans les exemples suivants : « parlé / du paysage », « le voyageur / égaré » ou « et vous étudiez / la littérature ». Ainsi, l'hypothèse de transfert de la L1 n'a été que partiellement validée. En effet, les apprenants ont tous été capables de produire les allongements contrairement aux prédictions des auteurs, mais leur production non-native de contours descendants non-finaux peut être attribuée à un transfert négatif du mandarin vers le français. D'autres difficultés dans la gestion de l'indice de durée (allongements trop importants chez certains apprenants, bien que les auteurs notent qu'ils pourraient être liés au débit lent des apprenants) et des pauses (présence de pauses dans les groupes de sens) sont toutefois notées et ne résulteraient pas d'un transfert direct de la L1.

Dans sa thèse sur la prosodie des apprenants hispanophones du français, Santiago (2014) s'est intéressé à leur production et perception des contours mélodiques montants à la fin des groupes prosodiques. Il a enregistré 15 apprenants hispanophones (6 apprenants de niveau A2 et 9 apprenants de niveau B1) ainsi que 10 francophones natifs et 10 hispa-

nophones (enregistrés en espagnol) en tant que groupes contrôles effectuant trois types de tâches du protocole COREIL (Delais-Roussarie & Yoo, 2011) :

1. deux productions orales interactives (une interview et un jeu de rôle)
2. deux productions orales monologuées (description d'un tableau et raconter une histoire à partir d'une image)
3. une lecture oralisée de courts textes et de dialogues

A partir de ces données, il a analysé les questions et les frontières prosodiques produites par les apprenants. Les résultats concernant les questions ont été présentés au début de ce chapitre (3.1). Les résultats de l'analyse des frontières d'IP montrent, comme pour les questions, la présence de contours extra-montants (HH%), qui ne sont cependant cette fois-ci pas attribuables à un transfert direct de la L1 puisque les hispanophones produisent des contours montants (H%) à ces positions prosodiques dans leur L1. Les contours extra-montants sont également plus souvent associés à une pause. De plus, la proportion de contours extra-montants diminue chez les apprenants de niveau B1. En ce qui concerne la gestion de l'indice de durée chez les apprenants, ils ont tendance à produire des allongements moins importants aux frontières des unités prosodiques. Santiago propose trois hypothèses qui pourraient expliquer ce phénomène : la première est que les apprenants hispanophones utilisent des contours extra-montants pour compenser l'allongement syllabique nécessaire en français mais qu'ils ont plus de difficultés à produire, la deuxième est qu'il s'agit d'un contour « par défaut » dans l'interlangue des apprenants et la troisième hypothèse est que ces contours expriment l'insécurité linguistique des apprenants, hypothèse notamment appuyée par le fait que ces contours soient souvent associés à des pauses.

Il existe quelques rares études sur la prosodie et la gestion de l'accentuation du français par des apprenants japonophones – quatre à notre connaissance – que nous allons présenter ici.

Kamiyama (2005) a mené une expérience de production avec un corpus de 8 phrases sous la forme de « l'économie de + N » – par exemple « L'économie du Canada. » lues par 11 apprenants japonophones du français et 4 francophones natifs. Hormis 2 apprenants, les

japonophones se différencient des francophones en réalisant une descente mélodique dès le début de la phrase ou à partir de la deuxième syllabe. Si certains apprenants produisent bien une montée mélodique sur la syllabe accentuée, elle est précédée d'une forte pente. Ainsi, alors que les francophones produisent habituellement une montée continuative sur « L'économie », les apprenants japonophones ont tendance à produire des descentes mélodiques à cet endroit qui peuvent être attribuées à un phénomène de transfert de la L1 et à l'organisation accentuelle du japonais (cf. Chapitre 2).

Anan (1981) a enregistré les productions de 4 apprenants japonophones et de 4 francophones dans une tâche de lecture de phrases. Les phrases utilisées comme corpus comportent différentes structures syntaxiques :

1. des phrases à SN avec des éléments d'adjonction

- SN multiples (2 ou 3) dans deux positions (exemples : « Monique et André viendront », « Je lui offrirai une jupe, un pull et un sac. »)
- Apposition nominale avant et après le nom (exemples : « Paris, capitale de la France, a longtemps été la capitale culturelle de l'Europe. », « Capitale de la France, Paris a longtemps été la capitale culturelle de l'Europe. »)
- Incises à l'intérieur et en fin d'énoncé (exemples : « Il est parti, je crois. », « Il est, je crois, parti. »)

2. des phrases sans verbe (exemples : « Ces enfants, insupportables. », « Ce grand monument, une déception. »)

3. Projection d'un des actants (= dislocation à gauche, exemple : « Pierre, il est parti. »).

L'auteure note différentes réalisations typiques chez les apprenants japonophones, et détaille certains points qui sont visibles dans l'exemple donné en Figure 3.1.

1. Les japonophones produisent les phrases avec un départ plus haut que la fréquence fondamentale moyenne de l'énoncé
2. Les japonophones ont tendance à produire un pic mélodique suivi d'un plateau ainsi qu'un schéma mélodique montant-descendant à travers le groupe prosodique, s'apparentant aux schémas accentuels du japonais

3. L'écart mélodique entre les deux premiers groupes prosodiques est plus réduit chez les apprenants japonophones que chez les francophones
4. Les japonophones produisent une montée mélodique insuffisante à la fin du noyau (« en couleur » dans l'exemple), même s'ils réalisent bien une pause à sa suite
5. Concernant la fin de la phrase, les japonophones produisent la première syllabe du dernier groupe prosodique avec une mélodie plus haute (« une » dans l'exemple) suivie de la descente finale, tandis que dans les réalisations des francophones c'est sur la première syllabe du nom que l'on observe une montée (« trouvaille ») avant la descente finale

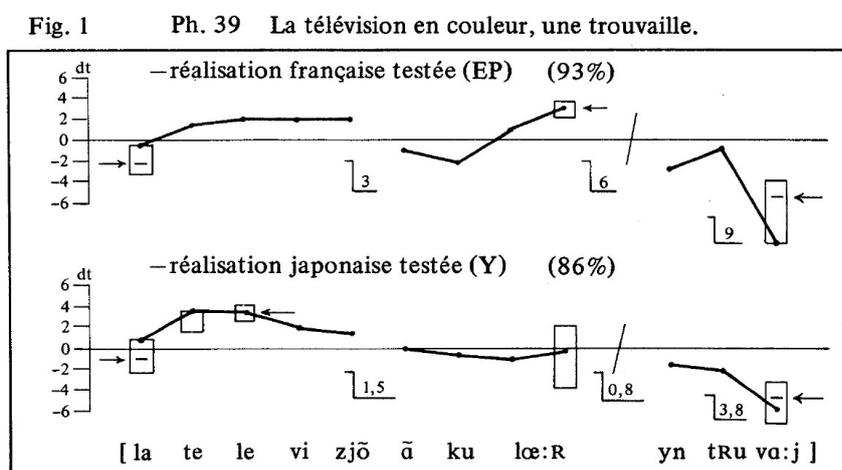


FIGURE 3.1 – Exemple issu de Anan 1981 : réalisation de la phrase « La télévision en couleur, une trouvaille » par un locuteur français et un locuteur japonais

L'auteure note également que les apprenants japonophones ont tendance à produire des sommets mélodiques graduellement descendants, des pauses plus longues et un empan mélodique plus réduit que les francophones.

Lepetit (1989) a mené une étude sur les contours prosodiques réalisés en français par des apprenants canadiens anglophones et japonophones. Il a enregistré 45 apprenants canadiens et 30 apprenants japonophones effectuant une tâche de lecture de phrases, en utilisant un corpus de 25 phrases pour les anglophones et de 22 phrases pour les japonophones. Pour l'analyse des productions obtenues, Lepetit utilise le cadre théorique de Philippe Martin (voir 2.2.1 pour plus de détails). Les résultats montrent que les apprenants

canadiens produisent des contours descendants sur des contours C1 (qui correspondent à une continuation majeure) dans 35,4% des contrastes C1 C0 et 53,9% des contrastes C1 C1 C0, qu'il attribue à un potentiel transfert de la L1. Il observe également 26,1% de schémas prosodiques incorrects pour les contrastes C3 C0, pourtant similaires dans les deux langues. Concernant les apprenants japonophones, ils ont également tendance à produire des contours descendants sur les contours C1 dans 11,94% des contrastes C1 C0 et 15,54% des contrastes C1 C1 C0. Nous notons cependant que ces pourcentages ne sont pas aussi élevés que pour les apprenants canadiens et donc que ce phénomène est loin d'être aussi fréquent. Le contraste ayant posé le plus de difficultés aux apprenants japonophones n'est pas C1 (C1) C0 mais C2 C1 – par exemple « la **soeur** (C2) de **Louis**, (C1) [...] » – où les apprenants ont réalisé deux contours montants à la place d'un contraste de pente descendant sur la syllabe accentuée de la première partie (C2) et montant sur celle de la deuxième partie (C1).

Dans son étude sur la réalisation des segments et supra-segments français chez les apprenants japonophones, Kojima (1978) fait des observations sur l'accentuation des fins de groupes prosodiques. Il a enregistré 90 apprenants japonophones de niveau débutant en français, 58 japonophones de niveau intermédiaire ainsi qu'une francophone native, et la partie intonative de son analyse se base principalement sur la lecture d'un court texte d'une soixantaine de mots. Dans un premier temps, il note la présence de contours plats dans les productions des japonophones à la fin d'un groupe prosodique qui correspond selon nous à un syntagme intonatif (« [...]les rois de France s'établirent à **Paris**,[...] ») et qui est réalisé par une montée chez les francophones, qui sont toutefois bien associés à un allongement. Lorsque les apprenants produisent des montées en fin de groupes rythmiques, l'auteur remarque une tendance non-native à produire des contours montants « abrupts », et cette fois-ci caractérisés par un manque d'allongement final. Chez certains apprenants, il note également la production d'une longue descente avant la montée finale, aussi présente sur la syllabe accentuée dans la production de la locutrice francophone du groupe contrôle, mais bien plus courte que celles produites par les japonophones. En dehors de l'accentuation finale, il note également une tendance chez les japonophones à fortement accentuer le début des mots (accent initial).

Ainsi, les quatre études que nous avons présentées sur la prosodie des apprenants ja-

ponophones montrent une tendance à produire des descentes mélodiques ou des contours plats due à l'interférence de la L1. Lorsqu'ils réalisent bien la montée continuative, Kamiyama (2005) et Kojima (1978) notent que celle-ci est précédée d'une forte descente et Anan (1981) montre que cette montée mélodique est insuffisante, en comparaison avec les francophones du groupe contrôle. Lepetit (1989) note également des difficultés à produire certains contrastes de pente en produisant cette fois-ci une montée sur la première syllabe accentuée à la place d'une descente. Les limites de ces études sont qu'elles sont toutes basées sur des corpus de phrases lues (ou de texte dans Kojima 1978) assez réduits, parfois avec un faible nombre de participants. Par ailleurs, la gestion des continuatifs par les apprenants japonophones en parole spontanée reste encore à notre connaissance un sujet inexploré.

### 3.3 Synthèse

Nous avons présenté dans ce chapitre différents travaux sur la prosodie des apprenants en L2 qui nous permettent une réflexion plus avancée concernant notre recherche. Sans surprise, la L1 des apprenants semble bien interférer sur la prosodie L2 et la grande majorité des études présentées ici ont montré la présence de **transfert négatif**. Cependant, parmi ces mêmes études, un certain nombre d'entre elles constatent que ce phénomène de transfert de la L1 n'est pas toujours clair et direct, ou ne permet pas d'expliquer les résultats obtenus dans leur totalité. Il semblerait que le transfert de la L1 vers la L2 soit moins systématique en prosodie L2 que dans les études segmentales. Nous notons également différents types de transferts négatifs – le transfert des **représentations phonologiques** de la L1 et celui des **propriétés phonétiques**. Le premier type de transfert peut correspondre à la production en L2 dans un contexte donné d'un contour mélodique habituellement produit en L1 dans ce même contexte. Par exemple, certains travaux que nous avons cités montrent que des apprenants en français L2 produisent plus de contours descendants que les francophones aux frontières prosodiques non finales en raison d'un phénomène de transfert de leur L1 (Anan, 1981 ; Bartkova et al., 2012 ; Grover et al., 1987 ; Kamiyama, 2005 ; Kojima, 1978 ; Lepetit, 1989 ; Xi et al., 2020), et les francophones ont tendance à produire à l'inverse plus de contours montants en fin d'unité intonative en anglais L2 (Bartkova et al., 2012 ; Horgues, 2010). D'autres travaux, en revanche, montrent que des contours habituellement

réalisés comme non-montants en L1 sont toutefois correctement produits comme montants en L2 (Albin & Wang, 2020 ; Lee, 2014). Cependant, les auteurs observent des difficultés **phonétiques** que l'on peut attribuer au deuxième type de transfert négatif. Dans Albin et Wang (2020) il s'agit notamment de difficultés à produire la forte montée de F0 présente dans les questions en japonais, et dans Lee (2014) les difficultés observées concernent l'alignement tonal des contours montants en fin de AP en français. D'autres travaux notent également la production de montées mélodiques réduites pouvant être expliquées par un transfert négatif (Kamiyama & Sakamoto, 2010).

La gestion de la durée peut également être problématique avec la réalisation d'allongements moins importants (Santiago, 2014) ou trop longs (Marchand & Le Gac, 2020 ; Xi et al., 2020) bien que pour ces deux dernières études ils ne soient pas attribuables à un transfert de la L1. En effet, certaines difficultés phonétiques ou phonologiques ne sont pas la conséquence d'un phénomène de transfert direct de la L1. Par exemple, en l'absence de phénomène similaire dans leur L1, les apprenants hispanophones et germanophones ont des difficultés à distinguer les types de « questions tag » en anglais (Puga et al., 2017 ; Verdugo & Trillo, 2005) et à y associer le contour mélodique approprié ; dans l'étude de Verdugo et Trillo les hispanophones surgénéralisent les contours montants à toutes les « questions tag ». Dans la thèse de Santiago (2014), les apprenants hispanophones produisent des contours extra-montants en français qu'ils surgénéralisent aux frontières de IP et à la fin des questions ; il s'agirait ici d'une stratégie liée à la construction de l'interlangue des apprenants. Santiago propose comme hypothèses explicatives le fait que les apprenants compensent leur production d'allongements finaux insuffisants, que ces contours extra-montants soient présents par défaut dans l'interlangue des apprenants, ou qu'ils manifestent ainsi leur insécurité linguistique. D'autres phénomènes sont observés dans le parler des apprenants tels que la segmentation en courts *chunks* (Barquero-Armesto 2012 ; Lee 2014, phénomène de « sursegmentation » relevé par Kaglik et Boula de Mareüil 2010) liés à une moindre fluence chez les apprenants, ou des pauses importunes au milieu des groupes de sens (Xi et al., 2020).

Dans certains cas, les apprenants intègrent une partie du système prosodique L2 dans leur interlangue sans pour autant être capables de le maîtriser ; par exemple Marchand et Le Gac (2020) observent un bon emploi des accents en français par les turcophones mais

des difficultés à gérer l'allongement et les variations mélodiques selon le type de frontière prosodique. Schwab (2012) a également obtenu des résultats montrant que les hispanophones ont intégré dans leur interlangue la connaissance de l'accent final en français mais qu'ils ne l'ont pas associé au niveau prosodique AP. La **gestion des différents niveaux prosodiques** du français et des indices prosodiques qui leur sont associés semble donc problématique chez des apprenants de différentes L1.

Il faut aussi noter que si les interférences issues d'un transfert négatif sont plus souvent mises en avant que celles résultant d'un **transfert positif** du fait de leur intérêt didactique (il est plus pertinent de pointer les erreurs et différences prosodiques dans le parler des apprenants que leurs réussites), il n'en reste pas moins que le transfert positif peut permettre d'anticiper des facilités ou expliquer certains résultats chez les apprenants. Gabriel et al. (2014) notent par exemple que les apprenants mandarinophones et plurilingues mandarin et allemand ont plus de facilité à produire le français avec un rythme proche de celui des natifs que les germanophones monolingues en raison de la proximité rythmique du mandarin avec le français. D'autres études observent ce qui s'apparente à un transfert positif de la L1 sur la L2 dans la gestion des indices prosodiques : la durée (Duběda, 2009 ; Schwab & Llisterri, 2012) et l'intensité (Duběda, 2009).

Enfin, parmi les études que nous avons présentées, il n'est pas rare que les résultats obtenus varient grandement entre les apprenants (Duběda, 2009 ; Gabriel et al., 2014 ; Grover et al., 1987 ; Kamiyama & Sakamoto, 2010 ; Kojima, 1978 ; Xi et al., 2020). Ces résultats différents sont parfois attribués au niveau de langue (Duběda, 2009 ; Xi et al., 2020), bien qu'un niveau de langue plus élevé ne soit pas systématiquement corrélé à de meilleurs résultats, comme dans Dupoux et al. (2008), à l'âge (Grover et al., 1987), ou à des facteurs extralinguistiques tels que l'intérêt pour la L2 ou la sensibilité phonétique (Gabriel et al., 2014). Il nous paraît donc primordial d'observer attentivement et de présenter les **différences individuelles** dans notre recherche.

Nous terminons cette synthèse par quelques **remarques méthodologiques**. En effet, il apparaît que les travaux en prosodie L2 utilisent des méthodologies très différentes, des corpus très différents (le plus souvent les tâches de production sont des lectures de phrases ou de textes, mais parfois des tâches plus spontanées comme des descriptions d'images, parole élicitée par un contexte donné, parole libre sur un sujet en particulier,

etc.). Ensuite, les groupes d'apprenants comportent un nombre de locuteurs très variable pouvant aller de quelques personnes jusqu'à une soixantaine d'apprenants, et ils ne sont pas nécessairement séparés en groupes de niveau de langue. Enfin, les données analysées diffèrent selon les expériences : les propriétés rythmiques, les contours mélodiques produits (annotation ToBI ou non), les données prosodiques brutes (F0, durée des voyelles ou des syllabes, intensité) ; la présence de pauses, etc. Les études perceptives sont également plus rares et servent souvent d'appui à des études basées sur la production, que ce soit pour vérifier les résultats obtenus en production et/ou pour analyser comment les productions des apprenants peuvent être perçues par les locuteurs natifs. Enfin, concernant la problématique de l'**erreur** en prosodie L2, nous pouvons constater que tous les travaux que nous avons présentés comportent un groupe contrôle de natifs qui permet une comparaison directe des paramètres analysés entre les apprenants L2 et les locuteurs du groupe contrôle servant de « norme » prosodique et/ou ayant pour rôle de confirmer la base théorique mise en avant dans les hypothèses de recherche. Si certaines études parlent d'« erreurs » prosodiques ou émettent un jugement sur les réalisations des apprenants en fonction des écarts constatés par rapport aux locuteurs natifs (par exemple, des montées mélodiques « insuffisantes », des allongements vocaliques « trop importants », etc.), d'autres cherchent à vérifier la portée de ces écarts dans des tests de jugement effectués par des francophones natifs, souvent mis au point avec l'aide de la synthèse vocale ou de modifications des stimuli dans des logiciels. Si cela semble être une bonne solution pour mettre en avant les différences prosodiques problématiques dans le parler des apprenants L2, ces études ont cependant aussi leurs limites. En effet, les tâches de jugement basées sur la prosodie peuvent être difficiles, comme les participants de l'étude de Grover et al. (1987) le rapportent. Ils évoquent notamment la possibilité d'imaginer un contexte possible pour chaque réalisation du continuatif, et c'est précisément cela qui rend difficile l'utilisation du terme d'erreurs prosodiques – et par conséquent les conclusions que l'on peut tirer. Les « erreurs » prosodiques peuvent être phonologiques ou phonétiques ; ainsi un contour habituellement produit montant dans un certain contexte remplacé par un contour descendant sera considéré comme une « erreur » phonologique tandis qu'une « erreur » phonétique sera, par exemple, la production d'une montée mélodique moins importante que chez les natifs. En perception, Dupoux et al. (2008) mentionnent également l'importance

du type de tâche concernant cette différence entre erreurs phonologiques et phonétiques ; les apprenants francophones réussiraient mieux les tâches ne demandant pas une longue mémorisation des stimuli car celles-ci se baseraient sur les différences phonétiques tandis que les tâches plus complexes demandant une mémorisation plus longue ou liées à l'intégration du lexique L2 seraient plus difficiles pour eux, car évaluant leurs compétences phonologiques.



# Chapitre 4

## Objectifs et questions de recherche

Les contours continuatifs, généralement de mélodie montante, sont présents dans la quasi-totalité des énoncés du français. En effet, ils jouent un rôle primordial dans la hiérarchie prosodique du français en étant associés aux frontières prosodiques, et, si peu qu'un énoncé ne soit pas excessivement court, il intégrera nécessairement un ou plusieurs de ces contours. Ainsi, toutes les descriptions prosodiques du français intègrent ces contours prosodiques, sous des noms et représentations différents que nous avons détaillés en 2.2.1. S'il n'existe pas de consensus sur le nombre de niveaux prosodiques du français, il est généralement admis qu'il existe au minimum deux niveaux auxquels sont associés des réalisations prosodiques différentes à leurs frontières. Dans notre thèse, nous choisissons de suivre la théorie auto-segmentale métrique et de retenir pour nos analyses trois niveaux : *Accentual Phrase* (à présent AP, syntagme accentuel), *intermediate phrase* (à présent ip, syntagme intermédiaire) et *Intonational Phrase* (à présent IP, syntagme intonatif). Dans notre analyse des tâches de production, nous allons principalement nous intéresser aux contours réalisés aux frontières de syntagmes accentuels et de syntagmes intonatifs. Ces contours se différencient par la gestion de la F0 et de la durée : les premiers sont associés à des « continuations mineures » (Delattre, 1966b) tandis que les seconds à des « continuations majeures », produites par un allongement final et une montée mélodique plus importants. Les contours continuatifs jouent également un rôle dans l'organisation des tours de parole : les contours continuatifs majeurs, lorsqu'ils sont en position focale, sont associés au maintien du tour et permettent de « suspendre et de reporter le moment de

l'interprétation » des énoncés (Bertrand et al., 2007). Ainsi, que ce soit en lecture oralisée ou en parole spontanée, les contours continuatifs ont un rôle central dans la prosodie du français et c'est pour cela que nous nous intéressons à leur production et perception par les apprenants japonophones.

Nos questions de recherche sont les suivantes :

- **Comment les apprenants japonophones produisent-ils et perçoivent-ils la continuation du français ?**
- **Sont-ils capables de gérer les différents niveaux prosodiques et d'y associer des continuations plus ou moins fortes selon le type de frontière prosodique ? Plus précisément, sont-ils capables de produire – et percevoir – des continuations mineures aux frontières des syntagmes accentuels et des continuations majeures (réalisées par une montée mélodique et un allongement final plus importants) à celles des syntagmes intonatifs ?**
- **Est-ce que leurs potentielles difficultés (ou facilités) peuvent s'expliquer par le transfert de leur langue première, ou par des stratégies spécifiques aux apprenants, telles que la surgénéralisation des contours prosodiques ?**
- **Est-ce que les apprenants sont capables d'intégrer la continuation du français à leur interlangue prosodique ? En d'autres termes, les apprenants de niveau plus avancé produisent et perçoivent-ils mieux la continuation du français que les apprenants de moins bon niveau ?**

Nous avons vu dans le premier chapitre de notre état de l'art (cf. Chapitre 1) que l'acquisition L2 et les potentielles difficultés ou facilités observées chez les apprenants pouvaient s'expliquer par différents phénomènes : le transfert de la L1, des stratégies d'apprentissages, l'âge de début d'apprentissage, la qualité et quantité de l'input auquel les apprenants sont exposés, la motivation de l'apprenant pour l'apprentissage de la langue et plus particulièrement de son aspect phonétique, ainsi que ses capacités individuelles. Parmi eux, le transfert de la L1 est souvent considéré comme ayant l'influence la plus importante, en particulier en ce qui concerne les compétences phonétiques et phonologiques. Ainsi, dans le deuxième chapitre (cf.2) nous avons comparé les systèmes prosodiques du

---

français et du japonais afin d'évaluer de quelle manière la production et perception des contours continuatifs du français pourraient être influencées par la L1 des apprenants japonais. La comparaison des systèmes prosodiques du français et du japonais a mis en évidence l'absence d'un phénomène similaire à la continuation du français en japonais, où les frontières prosodiques sont généralement associées à des tons bas (L%). Cependant, des contours mélodiques montants facultatifs ont été décrits à ces frontières, généralement en parole spontanée, ayant un rôle discursif et interactionnel. C'est la raison pour laquelle nous avons choisi d'analyser plusieurs types de parole dans nos expériences. Nous présentons ainsi dans cette thèse deux expériences de production : une tâche de lecture de textes oralisée (parole lue) et une tâche de description d'images (parole semi-spontanée). Nous avons vu en 1.5 que les compétences perceptives étaient souvent liées à la production, bien qu'il puisse exister un décalage entre ces deux compétences. Cela justifie l'intégration de deux expériences faisant appel aux compétences perceptives des apprenants afin d'avoir un point de vue plus complet sur le phénomène d'acquisition des continuatifs chez les japonais : une tâche de reproduction d'énoncés modifiés et une tâche de complétion de phrases. Ces quatre expériences nous permettront ainsi de répondre à nos questions de recherche.



**Deuxième partie**

**Etudes expérimentales**



---

Dans cette partie expérimentale, nous présentons quatre expériences sur les contours continuatifs du français chez des apprenants japonophones. Dans le chapitre 5, nous nous intéressons à comment les apprenants produisent les continuatifs, et pour cela nous présentons nos analyses 1. d'une tâche de lecture de textes oralisée (parole lue) et 2. d'une tâche de description d'images (parole semi-spontanée). Ensuite, dans le chapitre 6 nous nous intéressons à la perception des contours continuatifs et présentons les résultats de 3. une tâche de reproduction d'énoncés modifiés testant leur intégration des contours continuatifs dans leur représentation phonologique du français et 4. une tâche de complétion de phrases visant à tester la perception des frontières de AP et de ip par les japonophones.



# Chapitre 5

## Expériences de production des contours continuatifs en parole lue et semi-spontanée

### 5.1 Introduction

La production des apprenants d'une langue étrangère est toujours analysée à travers le prisme du système phonétique et phonologique de leur L1. Nous avons vu dans le Chapitre 1 que la L1 pouvait influencer les productions des apprenants; un phénomène dit de « transfert » pouvant être positif comme négatif. Cependant, ce phénomène de transfert, bien que difficilement contestable, ne parvient pas à expliquer dans son entièreté les productions des apprenants L2 (ou interlangue). En effet, certaines autres stratégies telles que la surgénéralisation ou l'hypercorrection ont été observées dans différentes études (Pešková, 2019; Ramírez Verdugo, 2006; Rasier & Hiligsmann, 2007; Santiago, 2014).

Dans ce premier chapitre expérimental, nous nous intéressons à la production des contours continuatifs du français par les apprenants japonophones. La continuation en français a un rôle essentiel en étant très fortement liée à la hiérarchie prosodique. Bien que syntaxe et prosodie ne soient pas entièrement congruentes, certaines positions syntaxiques sont associées à un type de continuation en particulier; c'est le cas par exemple des ajouts à gauche qui sont réalisés avec un contour continuatif majeur (Delais-Roussarie

et al., 2020). En parole spontanée, nous avons également vu que les contours continuatifs avaient un rôle supplémentaire : un rôle discursif et interactionnel. En effet, les continuations majeures sont associées au maintien du tour de parole et organisent ainsi les interactions dialogiques (Portes & Bertrand, 2005). Plus précisément, elles permettent généralement de relier différents énoncés devant être interprétés ensembles, et indiquent donc à l'interlocuteur la nécessité de ne pas interrompre le tour de parole. On comprend donc l'intérêt d'une bonne production des contours continuatifs du français chez les apprenants L2. En effet, réaliser le bon type de continuation aux positions appropriées permet une parole plus fluide, un discours plus facilement compréhensible ainsi qu'une interaction naturelle entre l'apprenant et ses interlocuteurs ; en cas contraire, les interlocuteurs pourraient par exemple avoir des difficultés à comprendre si l'apprenant souhaite ou non garder le tour de parole, ce qui peut avoir des conséquences néfastes sur le déroulé de la discussion.

Le chapitre 2 nous a permis de constater que les systèmes phonologiques du français et du japonais étaient très éloignés, notamment au niveau des frontières prosodiques droites. Est-ce que les apprenants japonophones sont capables de produire les montées continuatives du français réalisées à ces positions prosodiques ? Il s'agit donc dans un premier temps de s'intéresser à l'intégration *phonologique* des continuations du français. Est-ce qu'ils sont également capables de différencier les frontières prosodiques y associant un contour plus ou moins important selon la force de la frontière ? Si des difficultés existent, est-ce que nous pouvons les expliquer par le transfert de la L1 et est-ce qu'il y a des différences observées entre la production des indices prosodiques (durée et F0) ? Enfin, compte tenu des différences existantes en L1 entre la parole lue et spontanée (cf. 2), est-ce que le style de parole joue un rôle dans les productions des apprenants ?

Nos hypothèses concernant les expériences de production sont les suivantes :

**Hypothèse 1** – les productions des contours continuatifs par les apprenants japonophones pourront en partie être expliquées par un phénomène de transfert de leur langue maternelle.

**Hypothèse 2** – le transfert à lui seul ne permet pas d'expliquer les productions des apprenants et d'autres stratégies comme la surgénéralisation des faits prosodiques peuvent

être mises en place par les apprenants.

**Hypothèse 3** – les apprenants sont capables d'apprendre le système prosodique d'une langue cible, aussi éloigné qu'il puisse être de leur langue maternelle.

Dans le cas de nos apprenants japonophones, nous prédisons des difficultés à produire les contours montants aux continuations du français, avec cependant des différences liées au style de parole. En effet, les montées mélodiques n'ont pas de rôle phonologique en japonais et ne permettent pas de segmenter les énoncés en unités prosodiques comme c'est le cas en français où les frontières prosodiques (AP, ip, IP) sont associées à des montées mélodiques. En revanche, elles peuvent avoir un rôle interactionnel en parole spontanée et indiquer la continuation de l'acte de parole, bien que même dans ce cas elles restent facultatives. Ces différences entre les deux types de parole nous laisse donc prédire que les montées mélodiques seront plus facilement réalisées par les apprenants en parole spontanée qu'en parole lue.

Nous prédisons également des difficultés à produire différents types de continuation suivant la frontière prosodique et à gérer les indices de durée et de F0 de manière appropriée. En japonais, la longueur vocalique ayant une fonction distinctive (voyelles courtes VS longues), l'allongement vocalique du français poserait d'autant plus problème aux apprenants. Encore une fois le style de parole pourrait jouer un rôle important puisque les mouvements prosodiques montants en parole spontanée (facultatifs) indiquant la continuation sont parfois associés à un allongement vocalique, notamment le contour montant-descendant, HL%. De plus, une durée plus longue du dernier phonème (Koiso et al., 1998) ou syllabe (Yoneyama et al., 2003) serait observée en frontière discursive forte ou lors du maintien du tour de parole, à des positions discursives qui correspondent à celles des continuations majeures du français. Cela permettrait donc un potentiel transfert positif uniquement en parole spontanée.

Toutes les prédictions énoncées ci-dessus sont associées à l'**hypothèse 1** du transfert prosodique de la L1 sur la L2.

L'**hypothèse 2**, quant à elle, prédit la présence de phénomènes non imputables à la langue maternelle des apprenants. En particulier, nous nous attendons à observer une sur-généralisation des contours montants, et en particulier de contours extra-montants aux frontières de IP, comme il a pu être observé dans les études de Santiago (Santiago, 2014).

Si cette hypothèse semble être en concurrence avec l'hypothèse de transfert, qui prédit au contraire des taux réduits de contours montants, il est toutefois possible que différents aspects des résultats puissent être expliqués par l'une ou l'autre des hypothèses.

Enfin, nous prédisons que les apprenants de niveau plus avancé auront des productions plus proches de celles des natifs, car un niveau de langue plus élevé est souvent associé à une meilleure production (Duběda 2009 ; Xi et al. 2020 entre autres) : ils produiront ainsi des taux de contours montants équivalents à ceux des natifs et seront mieux capables de produire les différentes forces de continuation et de gérer les indices de durée et de F0 (hypothèse 3).

## 5.2 Protocole expérimental

### 5.2.1 Corpus

Pour l'étude de la production des contours continuatifs, nous avons besoin d'un corpus comportant différents types de productions permettant à la fois une analyse de la parole lue et de parole plus spontanée. En effet, nos hypothèses de recherche concernant la production des apprenants supposent un traitement différent des contours continuatifs en parole lue et spontanée, dû à des différences existantes en langue maternelle. De plus, il nous paraît important d'un point de vue didactique d'analyser des productions variées, plus représentatives des enjeux réels de l'apprentissage du français. Nous considérons que la parole lue ne pourrait être représentative de tous les types de parole et ne serait donc pas suffisante à l'analyse des contours continuatifs du français, notamment car ceux-ci ont diverses fonctions dont certaines directement liées à l'organisation du discours et à l'interaction. Nous avons choisi d'utiliser un corpus déjà existant et spécifiquement créé pour la recherche en prosodie L2 : le corpus COREIL (Delais-Roussarie & Yoo, 2011). Le principal avantage de ce corpus est qu'il valide nos premiers critères en comportant une grande variété de tâches (de lecture à parole spontanée). L'autre atout de ce corpus est le fait qu'il a été spécifiquement conçu pour les apprenants et que le vocabulaire employé leur est donc accessible. Le corpus COREIL comporte ainsi différentes tâches :

1. lecture de textes (quatre textes courts d'environ 50 mots chacun, ainsi qu'un texte long issu du *Petit Prince* de Saint-Exupéry) et de dialogues

2. tâche de description d'images (8 images)
3. jeu de rôle où les apprenants doivent poser des questions à l'examineur
4. interview guidée sur différents sujets de la vie quotidienne où l'examineur s'aide d'une liste de questions (« Où est-ce que vous vivez? », « Où êtes-vous allé en vacances récemment? »...)

Après des enregistrements pilotes du corpus, nous avons choisi de ne pas enregistrer la lecture du texte long en raison des contraintes de temps que cela nous imposait, préférant privilégier une tâche de parole spontanée (4) plus longue. Cela nous a aussi permis d'avoir une durée maximale de 1 heure pour les enregistrements des différentes tâches et les procédures de consentement ainsi que le remplissage du questionnaire. Parmi les différentes tâches, certaines sont plutôt axées sur la production des questions et c'est particulièrement le cas pour le jeu de rôle, mais aussi des dialogues. Bien que nous les ayons enregistrées, nous n'avons pas utilisé ces données dans la présente recherche. Toutefois, elles sont des données précieuses et pourront être exploitées par la suite. Il serait par exemple intéressant de savoir comment les apprenants produisent les montées des questions et si leur réalisation est proche de celle des contours continuatifs ou non.

### 5.2.2 Participants

Nous avons enregistré des apprenants japonophones étudiant le FLE à Tokyo. La grande majorité des apprenants étaient étudiants à l'université Sophia, lieu où s'est déroulée l'expérience. Nous avons uniquement sélectionné les apprenants n'ayant jamais vécu dans un pays francophone pendant plus d'un mois et n'ayant pas vécu à l'extérieur du Japon durant leur enfance. 17 apprenants correspondaient à ces critères et leurs données ont donc pu être analysées. La plupart des apprenants ne suivaient pas un cursus spécialisé en français, mais simplement des cours de FLE en option (LV2). Nous avons directement recruté les apprenants dans leurs cours par l'intermédiaire des professeurs de l'université, à l'exception de deux apprenants recrutés par d'autres intermédiaires (au cours d'un atelier de conversation organisé par l'université et par des connaissances communes).

Nous avons classé les apprenants dans trois groupes de niveau : A1-A2, A2-B1 et B1-B2.

Participant	Age	Sexe	Classe	Durée d'apprentissage	Niveau	Séjour en pays francophone
JP2	21	H	LV2 avancé	3 ans	A2-B1	0
JP3	23	H	LV2 avancé	2 ans	A2-B1	0
JP4	21	F	LV2 avancé	4 ans	A2-B1	2 semaines (France)
JP5	22	F	LV2 avancé	3 ans	A2-B1	1 mois (France)
JP7	20	F	LV2 avancé	3 ans	A2-B1	1 semaine (France)
JP9	19	H	LV2 avancé	2,5 ans	A2-B1	2 semaines (France)
JP10	19	F	LV2 avancé	4 ans (dont 3 lycée)	A2-B1	0
JP11	19	F	LV2 débutant	3 ans (lycée)	A2-B1	3 semaines (France)
JP15	21	F	Spécialiste	2 ans	B1-B2	0
JP16	19	H	LV2 intermédiaire	2 ans	A1-A2	0
JP18	19	F	LV2 débutant	7 mois	A1-A2	0
JP19	21	F	LV2 avancé	4 ans	A2-B1	0
JP20	21	F	Spécialiste	2 ans	B1-B2	0
JP24	20	F	Spécialiste	2 ans	B1-B2	0
JP25	19	F	Autre	3 ans (lycée)	A2-B1	0
JP26	21	F	Spécialiste	3 ans	B1-B2	0
JP39	19	F	Spécialiste	3 ans	B1-B2	0

TABLEAU 5.1 – Données sur les apprenants japonophones ayant participé à l'expérience de production

Tous les apprenants suivant le cours de LV2 avancé ont été classés de niveau A2-B1, ceux suivant les cours de LV2 débutant et intermédiaire comme A1-A2. Les apprenants suivant un cursus avec le français en tant que spécialité ont été classés au niveau B1-B2. Certains apprenants avaient également étudié le français au lycée soit pour une durée de 3 ans. Pour certains d'entre eux en première année de licence, c'était même leur seule expérience en français, et nous les avons donc classés comme de niveau A2-B1 qui correspond aux objectifs linguistiques généralement visés par ces enseignements. En effet, les cours de français au lycée au Japon visent généralement la préparation des concours d'entrée à l'université en français LV2 qui demandent un niveau équivalent à fin A2 début B1 sur l'échelle du CECRL. Ainsi, 2 apprenants ont été classés dans le groupe A1-A2, 10 dans le groupe A2-B1 et 5 dans le groupe B1-B2. Le faible nombre de participants dans le groupe de débutants (A1-A2) s'explique par le fait que les apprenants ayant un niveau de langue plus bas aient été plus réticents à participer à une expérience de production enregistrée. Le groupe A2-B1 comporte le plus d'apprenants en raison du plus faible nombre d'étudiants spécialistes, qui, de surcroît, devaient correspondre à nos critères (notamment celui de n'avoir jamais vécu en pays francophone).

### 5.2.3 Protocole

Les apprenants ont été enregistrés dans la chambre sourde du laboratoire de phonétique de l'université de Sophia. Nous avons utilisé la carte son Quad Capture Roland ainsi qu'un micro casque SHURE WH20XLR. La configuration de la chambre sourde ne permettant pas de placer la carte son et l'ordinateur sur lequel elle était branchée (Asus Zenbook) à l'extérieur, ils ont été éloignés du microphone autant que possible afin d'obtenir une meilleure qualité sonore. Les enregistrements ont été réalisés avec le logiciel Praat (Boersma, 2001) avec un échantillonnage de 44100 Hz. Les participants du groupe contrôle que nous avons enregistrés à Paris pour la tâche de description d'images ont été enregistrés dans les mêmes conditions que les apprenants, à l'exception de la carte son et ordinateur qui, cette fois-ci, ont pu être positionnés à l'extérieur de la pièce. Pour chaque participant nous avons réservé un créneau d'une heure. Les participants étaient d'abord invités à compléter le formulaire de consentement ainsi que le questionnaire linguistique. Ensuite, nous avons enregistré les lectures de textes et de dialogues, les descriptions d'images, le jeu de rôle puis l'interview. Le support écrit et les images pour les deux premières tâches ont été imprimés sur papier et stabilisés par un porte-bloc pour éviter les bruits parasites du papier. Pour la tâche de lecture, nous avons incité les participants à préparer les textes en les lisant une ou deux fois seuls avant d'être enregistrés. Pour la description d'images, la consigne était de décrire les images le plus longuement que possible. La dernière tâche étant l'interview, la longueur de celle-ci dépendait du temps restant dans le créneau horaire et varie donc suivant les participants, allant de 11 à 30 minutes, avec une moyenne de 18 minutes. Dans tous les cas, la durée totale de l'expérience de production était d'une heure et les apprenants étaient rémunérés 1000 yen (équivalent à environ 7 euros) pour leur participation.

Nous allons maintenant présenter les résultats des deux expériences de production que nous avons choisi d'analyser. Nous allons dans un premier temps présenter les résultats de la tâche de lecture de textes (parole lue) et dans un deuxième temps la tâche de description d'images (parole semi-spontanée).

## 5.3 Expérience 1 : Lecture de textes

### 5.3.1 Introduction

La première tâche de notre corpus est une tâche de lecture de textes. Notre principal objectif est de savoir comment les apprenants produisent les contours continuatifs en parole lue, où ceux-ci sont très fortement liés à l'organisation phrastique. En effet, comme nous l'avons vu dans le Chapitre 2, différentes forces de contours continuatifs sont associées aux différentes positions prosodiques, intrinsèquement liées aux niveaux syntaxiques et sémantiques. Ainsi, par exemple, nous retrouvons des contours majeurs à la fin de la première partie d'une subordonnée (ex : « Si je dois partir avant midi de Marseille ») qui se manifestent par un allongement plus important de la voyelle accentuée et une forte montée mélodique, par rapport aux contours mineurs que l'on retrouve généralement en fin de syntagme accentuel. En japonais, en revanche, il apparaît que les frontières prosodiques ne sont généralement pas associées à des contours montants, les mouvements prosodiques ayant lieu à l'intérieur des AP (que celui-ci soit accentué ou non, LH- ou LH-H\*+L) tandis que chaque frontière est produite avec un ton bas (L%). Nous avons différentes prédictions pour cette étude, issues de nos principales hypothèses de recherche. Tout d'abord, nous nous attendons à observer moins de montées mélodiques aux continuations du français chez les apprenants que chez les natifs, en raison d'un transfert négatif de la L1 ([hypothèse 1](#)). Nous nous attendons aussi à observer des difficultés chez les apprenants à utiliser les indices prosodiques de façon appropriée, notamment celui de durée, et à distinguer les frontières prosodiques en produisant différents types de continuation (allongement vocalique et montée de F0 plus importants en frontière de IP que de AP par exemple).

Si notre première hypothèse prévoit donc un certain nombre de difficultés liées au transfert négatif, l'[hypothèse 2](#) prédit la présence de stratégies non explicables par le phénomène de transfert et qui seraient représentatives de l'interlangue des apprenants L2 de façon plus générale, par exemple des stratégies de surgénéralisation ou de hypercorrection. En particulier, en nous appuyant sur les études précédentes, nous prédisons la présence de contours extra-montants en fin de IP, observés par Santiago (2014) dans ses études sur les hispanophones. D'autres stratégies pourraient également être observées sans pouvoir être expliquées par la structure prosodique de la L1.

Enfin, nous nous attendons à ce que les résultats soient différents suivant le niveau d'apprentissage (*hypothèse 3*). Ainsi, nous prédisons que les apprenants de niveau plus avancé auront moins de difficultés pour produire les montées mélodiques et distinguer les différentes frontières prosodiques.

#### 5.3.2 Méthodologie

Les données des quatre courts textes (voir Annexe C) ont été recueillies selon le protocole décrit en 5.2.3. Nous avons donné un temps de préparation pour les apprenants avant la lecture de chaque texte, afin d'éviter au maximum les erreurs et hésitations. Ce temps de préparation était libre et les apprenants pouvaient commencer la lecture à voix haute dès qu'ils le désiraient.

Comme groupe contrôle, nous avons utilisé des enregistrements pré-existants de 5 francophones natifs sur les mêmes textes du corpus COREIL (Delais-Roussarie & Yoo, 2011).

Avant l'analyse des données, nous avons annoté manuellement toutes les frontières prosodiques supposées dans les quatre textes. Ainsi, nous avons annoté cinq niveaux prosodiques :

- Un niveau correspondant à l'absence de frontière prosodique ("mot prosodique", MP ; par exemple entre un clitique et un nom ; de mon quartier, je me régale) : 0
- Un niveau correspondant à la frontière d'un mot lexical ( « quelque chose », « je suis allé » : 1
- Un niveau correspondant à la frontière de syntagme accentuel (AP, « Est ce que vous pourriez me donner la liste des restaurants de mon quartier ? ») : 2
- Un niveau correspondant à la frontière de syntagme intermédiaire (ip ; « un wagon de jeux », « elle refuse absolument ») : 3
- Un niveau correspondant à la frontière de syntagme intonatif (IP) : 4

« Hier, je suis allé me promener dans les calanques », « Il fait un temps magnifique et la

mer est très bonne ».

Pour l'annotation des frontières de AP, nous avons choisi la définition classique de ceux-ci c'est-à-dire de considérer comme AP une tête lexicale avec les items qui en dépendent sur leur versant non récursif (gauche). Les critères pris en compte pour la définition des syntagmes intermédiaires sont ceux de Michelas (2011) : la frontière prosodique d'un syntagme comportant un minimum de 2 AP (MIN-BIN) coïncidant à une frontière de XP (ALIGN-XP,R, ip, R) est considérée comme frontière de ip. Pour l'annotation des IP non finaux, nous avons pris des critères semblables à ceux de Delais-Roussarie et Post (2008) dans sa définition des groupes intonatifs guidés par la syntaxe (GI-Synt). Nous avons considéré comme frontière de IP les frontières droites de subordinées antéposées, d'ajouts et dislocations, de phrases coordonnées ainsi que de XP incidents. Une différence subsiste dans notre traitement des XP incidents puisque nous avons considéré ceux-ci comme IP ayant une frontière à droite et à gauche là où seule la frontière droite est retenue par Delais-Roussarie.

Nous avons ainsi défini un total de 115 frontières prosodiques (15 occurrences de niveau 0, 22 occurrences en frontière de mot prosodique, 53 en frontière de AP, 13 en frontière de ip et 12 en frontière de IP) que nous avons nommées « items ».

Le choix de cette annotation principalement basée sur des règles phono-syntaxiques et non sur les réalisations réelles des francophones pourrait être contesté. Il est d'autant plus vrai que les critères et même l'existence de certains niveaux prosodiques ne fait actuellement pas consensus parmi les chercheurs. Cependant, nous considérons que cette méthode permet une comparaison directe et objective entre les productions des francophones et des apprenants. De plus, l'intégration des items dans les facteurs aléatoires de nos modèles statistiques permet de prendre en compte la part de variation expliquée par ceux-ci.

Les données ont été segmentées automatiquement avec le logiciel SPPAS (Bigi, 2015), en adaptant certains mots selon leur prononciation pour faciliter cette segmentation automatique. Une fois la segmentation automatique réalisée, nous l'avons manuellement vérifiée avec Praat (Boersma, 2001). Ensuite, nous avons noté sur une première tire chaque frontière prosodique et si elle était associée à un contour montant ou non (noté 1 ou 0).

Les critères de décision sont les suivants : la montée mélodique doit être audible et l'écart entre la voyelle accentuée et la voyelle précédente doit dépasser un demi-ton. Nous avons également annoté dans une autre tire les hésitations et la voix craquée afin de les exclure respectivement de l'analyse de la durée et de la F0. La voyelle accentuée est celle de la dernière syllabe pleine, et ne peut donc théoriquement pas être un schwa. Cependant, les japonophones ont tendance à produire des épenthèses vocaliques. Nous avons choisi comme voyelle accentuée le schwa final lorsque celui-ci semblait accentué par l'apprenant, notamment par une hauteur mélodique plus élevée que la précédente voyelle normalement accentuée. Nous avons ensuite créé et utilisé un script Praat afin d'extraire les données de durée de chaque voyelle ainsi que de F0 à différentes positions de la voyelle (25%, 50% et 75% de la voyelle). Nous avons ensuite calculé la différence de F0 entre la voyelle accentuée aux frontières annotée et celle qui précède en utilisant les valeurs à 75% de la voyelle. Enfin, ces données ont été importées dans R (R Core Team, 2018) pour les analyses statistiques.

Un exemple de segmentation Praat est présenté en Figure 5.1. Nous pouvons y voir un segment de voix craquée sur le premier mot, noté « cr » ainsi que des hésitations notées « + ». Malgré la présence de voix craquée sur la syllabe précédente, nous pouvons clairement percevoir une montée mélodique à cette frontière de IP qui est d'autant plus visible sur la syllabe accentuée. Cette manière de procéder nous a permis de conserver certaines informations (en particulier la présence ou non de montée mélodique) même lorsque le signal était partiellement illisible (ici par la présence de voix craquée) et que les données de F0 étaient quant à elles exclues de l'analyse prosodique.

#### 5.3.3 Résultats

##### **Observations générales**

Nous présentons ici (Tableau 5.2) différentes informations sur les enregistrements analysés dans notre étude : le débit de parole, le pourcentage moyen de pauses produites en frontière de AP et de IP ainsi que leur durée et le taux d'hésitations produites par niveau de langue.

Nous pouvons faire les observations suivantes :

- Le débit de parole augmente avec le niveau de langue et est bien plus élevé chez les

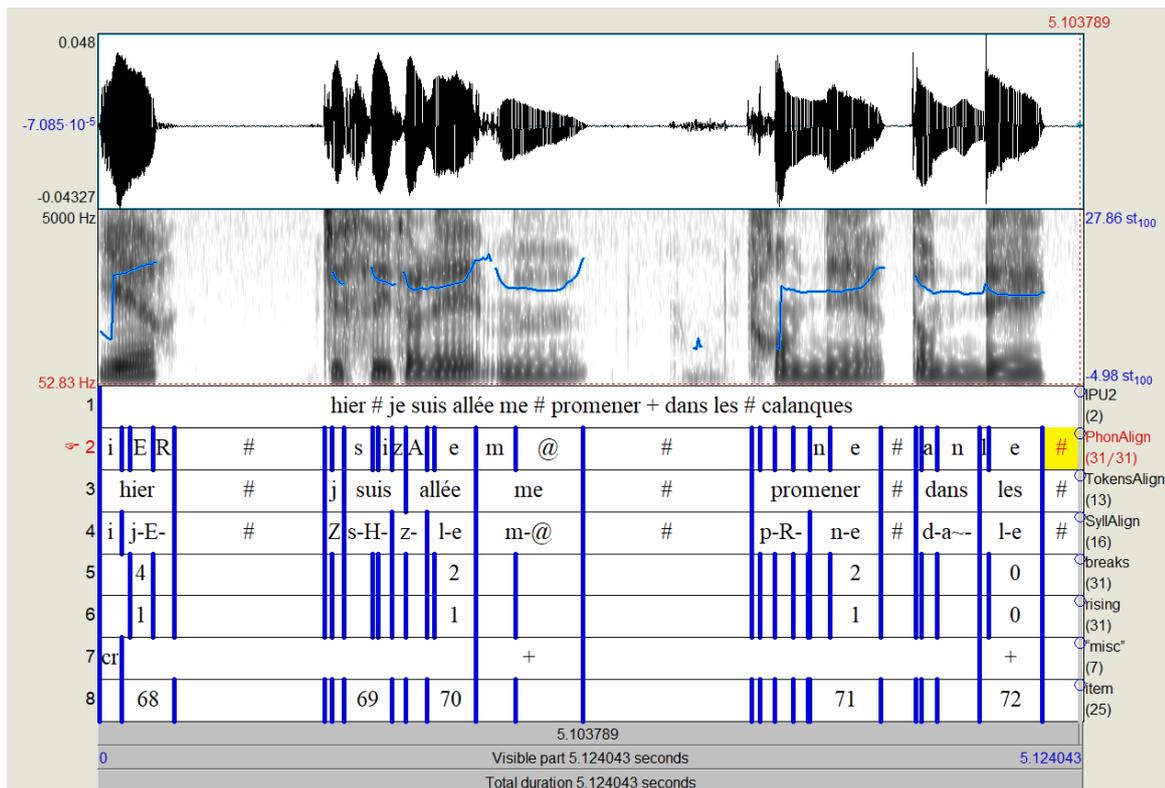


FIGURE 5.1 – Exemple d’annotation Praat

	Niveau de langue			
	A1-A2	A2-B1	B1-B2	natifs
débit de parole (syllabes / s)	3.7 (ET = 0.6)	4.1 (ET = 0.5)	4,2 (ET = 0.3)	6,1 (ET = 0.7)
% pauses frontière AP	48 % (ET = 21)	35 % (ET = 15)	20 % (ET = 13)	0,4 % (ET = 0,8)
durée moyenne pauses frontière AP	622 ms (ET = 367)	571 ms (ET = 377)	441 ms (ET = 203)	600 ms (ET = NA)
% pauses frontière IP	67 % (ET = 0)	70 % (ET = 9)	69 % (ET = 17)	11 % (ET = 13)
durée moyenne pauses frontière IP	638 ms (ET = 400)	490 ms (ET = 223)	561 ms (ET = 256)	336 ms (ET = 125)
% hésitations	3.1 % (ET = 3)	6.5 % (ET = 4)	2.5% (ET = 2,5)	0 %

TABLEAU 5.2 – Informations globales sur les données de lecture de textes

francophones que chez les apprenants. Ce résultat n’est pas surprenant puisqu’une meilleure fluence implique un débit de parole plus élevé.

- Les francophones ne produisent généralement pas de pauses aux frontières de AP (0,4%) tandis qu’elles sont fréquentes dans les productions des apprenants (20-48%).

Ce taux diminue légèrement avec le niveau de langue.

- La durée des pauses produites par les apprenants aux frontières de AP diminue également lorsque le niveau de langue augmente
- Les apprenants produisent six fois plus de pauses aux frontières de IP (67 à 70%) que les francophones natifs (11%).
- On observe dans tous les groupes de langue un pourcentage plus élevé de pauses aux frontières de IP que de AP.

Ainsi, nous constatons que les productions des apprenants diffèrent de celles des francophones du groupe contrôle. Le pourcentage très élevé de pauses aux frontières AP et IP chez les apprenants nous indique qu'ils ont tendance à produire de courts *chunks* délimités par des pauses lors de la lecture, ce qui correspond aux observations faites dans la littérature L2 (Barquero-Armesto 2012; Lee 2014, entre autres). Il est aussi intéressant de noter que les apprenants semblent faire une distinction entre les frontières de AP et de IP, notamment dans les groupes de niveau les plus élevés (A2-B1 et B1-B2) puisqu'ils produisent nettement plus de pauses aux frontières IP que AP.

#### **Résultats : Les montées mélodiques**

Alors que nous nous attendions à observer moins de contours montants chez les apprenants que chez les natifs, nos résultats montrent que seuls les apprenants débutants (groupe A1-A2) semblent produire moins de contours montants (voir Figure 5.2). Les apprenants des autres groupes d'apprentissage, quant à eux, semblent capables de produire autant de contours montants que les natifs, et surtout semblent capables de différencier les frontières prosodiques en produisant plus de contours aux frontières fortes (AP, ip et IP) que faibles (niveaux 0, pas de frontière et 1, mot prosodique).

Afin de vérifier ces observations, nous avons construit un modèle linéaire généralisé mixte avec la réalisation du contour prosodique montant (1) ou non (0) comme variable dépendante, une interaction entre le niveau de langue et le type de frontière prosodique comme effets fixes, et item et participant comme effets aléatoires. L'interaction entre le

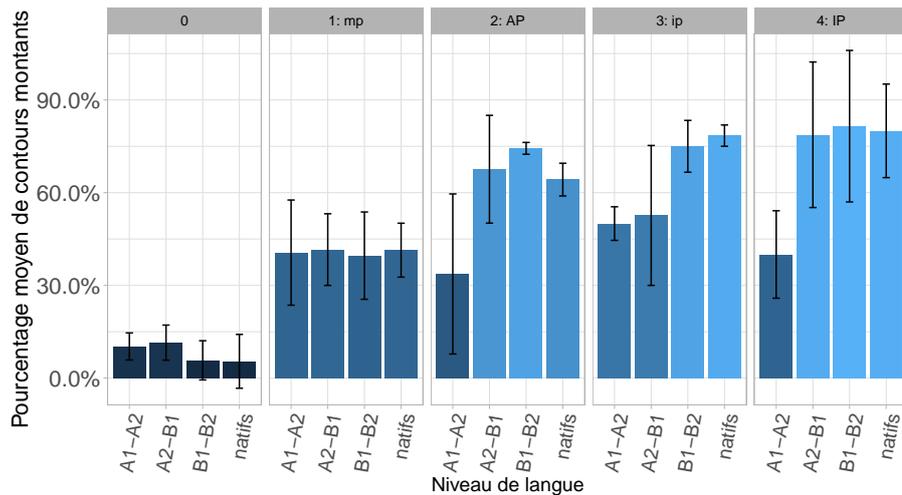


FIGURE 5.2 – Pourcentages moyens de contours montants produits à différents frontières prosodiques selon le niveau de langue. Les barres d’erreur représentent les écarts-types

niveau de langue et le type de frontière est significative ( $\chi^2(12) = 47.71, p < 0.001^{***}$ ). Nous avons ensuite utilisé le package *emmeans* pour une analyse des contrastes de l’interaction. Les résultats statistiques nous confirment les tendances observées : les apprenants des groupes A2-B1 et B1-B2 ont produit autant de contours montants que les natifs, excepté aux frontières de ip où les apprenants du groupe A2-B1 ont produit significativement moins de contours montants ( $p = 0.019^*$ ). Concernant les apprenants débutants (A1-A2), ils produisent significativement moins de contours montants que le groupe contrôle aux frontières de AP et de IP (respectivement  $p = 0.009^{**}$  et  $p = 0.027^*$ ). Le groupe A1-A2 se distingue également des deux autres groupes d’apprenants à ces positions prosodiques. Néanmoins, ce résultat reste incertain du fait du nombre restreint d’apprenants dans ce groupe (2 apprenants) et, de ce fait, du manque de puissance de l’analyse statistique pour ce cas précis.

Nous avons défini les frontières prosodiques manuellement et il est évident que cette annotation, bien que nécessaire, ne peut prendre en compte tous les phrasés prosodiques possibles pour un natif ni les différences contextuelles pouvant donner lieu à des réalisations différentes des contours. Nous pouvons voir ces différences dans notre groupe contrôle où l’on observe des taux de contours montants avoisinant les 75 % aux frontières prosodiques fortes (AP, ip et IP), mais sous-entendant qu’un quart des contours produits à ces frontières n’étaient pas montants. Or, cette proportion de contours montants et non-montants n’est pas répartie uniformément à travers tous les items. Ainsi, certains items

sont toujours associés à un contour montant par les sujets du groupe contrôle, tandis que d'autres ne le sont que rarement. C'est entre autres la raison pour laquelle nous avons intégré l'item comme effet aléatoire dans notre modèle statistique. Mais, de ce fait, même si les apprenants produisent généralement autant de contours montants que les francophones natifs, nous ne savions pas s'ils produisaient les mêmes proportions de contours montants que les natifs selon les items. C'est pour cela que nous avons classé les items selon leur pourcentage de contours montants produits par les sujets du groupe contrôle en 5 catégories : 0% (jamais montants), 20-35%, 75-80% et 100% (toujours montants). Nous observons que les proportions de contours montants produits chez les apprenants ne varient pas selon le type d'item. Cela s'observe notamment pour les items de frontières AP, ip et IP (voir Figure 5.3) où aucune corrélation entre le taux de contours montants chez les natifs et celui chez les apprenants n'est visible. Les apprenants généralisent donc les contours montants aux frontières prosodiques fortes.

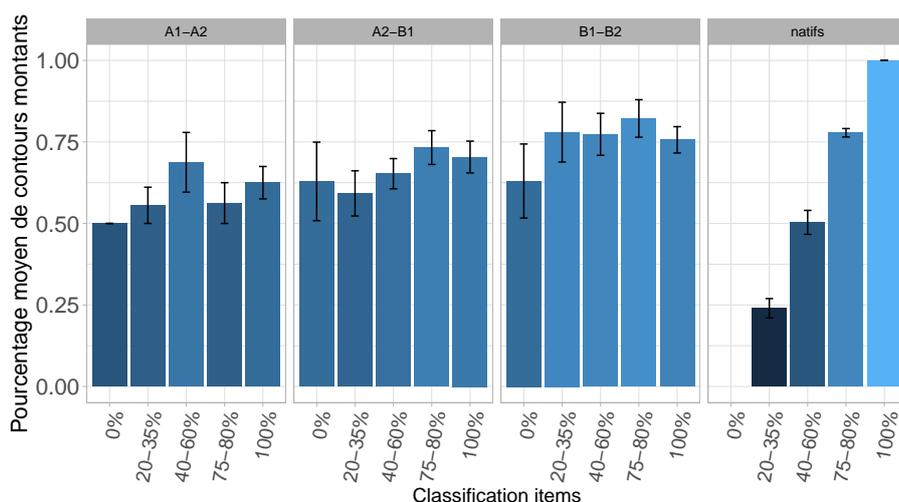


FIGURE 5.3 – Pourcentage de contours montants produits aux frontières fortes (AP, ip, IP) selon le taux de montées réalisées par les natifs (items à 0% de montées mélodiques, 20-35%, 75-80% et 100%)

Ainsi, les productions des apprenants sont rarement caractérisées par une absence de contour montant là où les francophones produiraient une montée. Toutefois, sur certains items nous observons un taux de contours descendants plus important chez les apprenants que les natifs et une absence de montée mélodique. C'est le cas notamment des exemples que nous présentons ici. Ces exemples ont très généralement été réalisés avec une montée mélodique par les francophones du groupe contrôle mais pas par les apprenants.

Par exemple, à la fin de la construction verbale « j’aimerais savoir » (voir Figure 5.4) les apprenants ne produisent que 30% de montées mélodiques alors que tous les francophones du groupe contrôle ont produit une montée. On observe en revanche une tendance à produire une montée sur la première partie de la construction « j’aimerais » alors que celle-ci est facultative chez les natifs.

On retrouve le même phénomène dans les exemples présentés ci-dessous où l’on observe l’absence de montée mélodique à la fin du syntagme nominal « un wagon de **jeux** » (Figure 5.6) et du syntagme adjectival « une peur **bleue** » (Figure 5.5), ainsi qu’une montée à la frontière droite du premier constituant (« wagon » et « peur »).

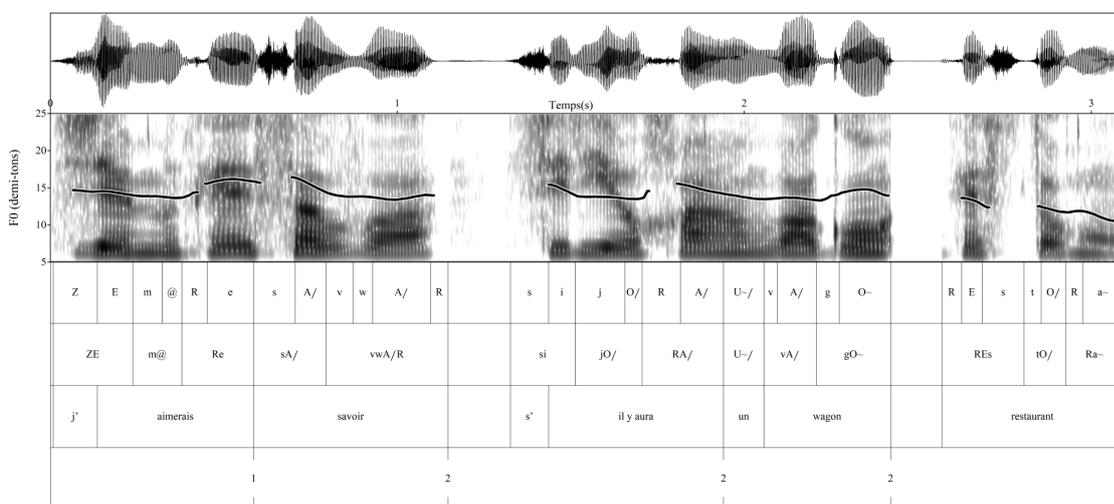


FIGURE 5.4 – Réalisation de l’énoncé « j’aimerais savoir s’il y aura un wagon restaurant » par JP5

Les frontières de syntagme nominal sujet (et ne comportant qu’un AP) n’ont pas posé de problème aux apprenants qui les ont généralement bien réalisées par une montée de F0, comme nous pouvons l’observer sur les mêmes exemples en Figure 5.5 (« ma **soeur** ») et 5.6 (« la SNCF »).

Concernant la réalisation des ajouts sur la gauche, ils ont globalement bien été réalisés bien que l’un d’entre eux affiche un taux moins élevé de montées avec un quart des productions où celle-ci est absente comme présenté ici en Figure 5.7 où l’ajout « hier » est réalisé par un contour descendant.

Nous observons également de la variation entre les sujets dans les différents groupes de

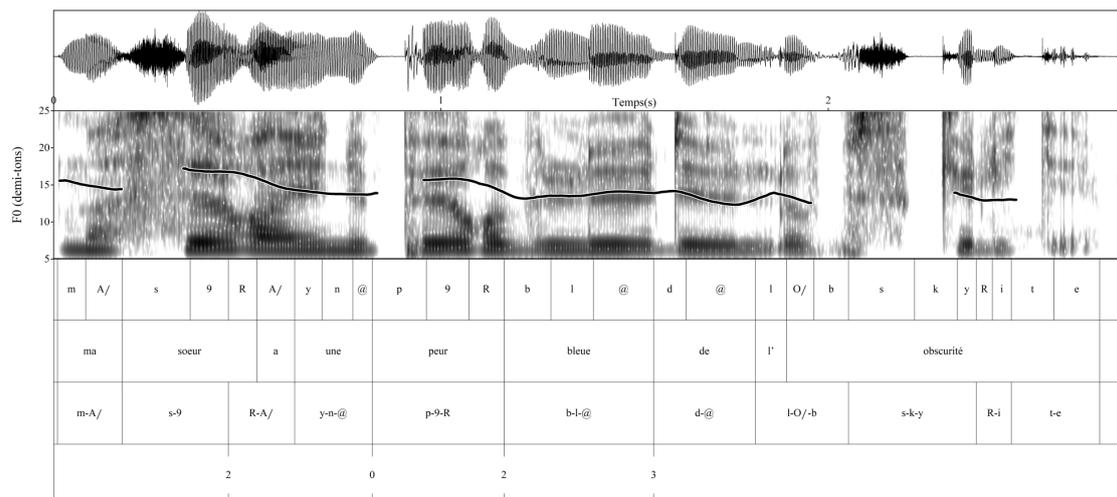


FIGURE 5.5 – Réalisation de l'énoncé « Ma soeur a une peur bleue de l'obscurité » par JP39

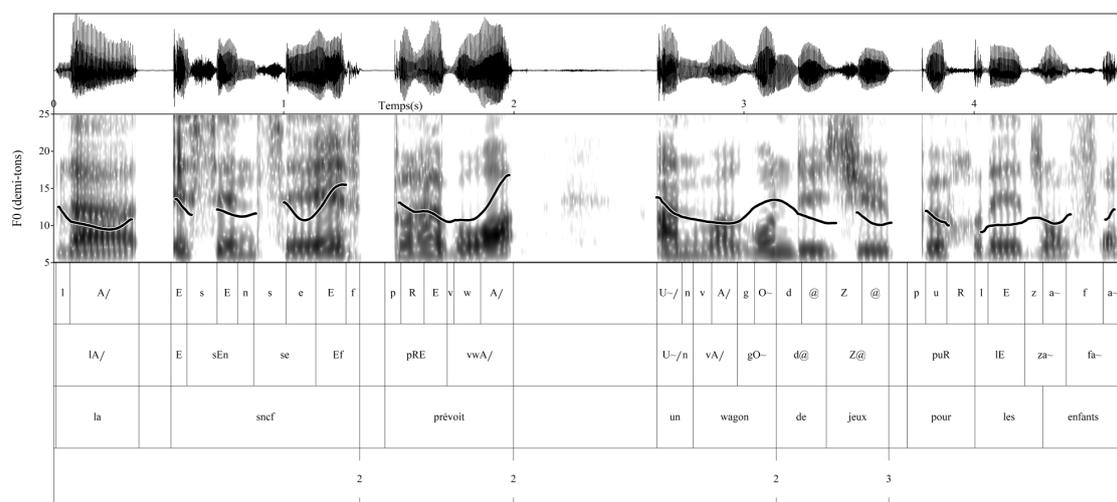


FIGURE 5.6 – Réalisation de l'énoncé « La SNCF prévoit un wagon de jeux pour les enfants » par JP26

niveau (voir Figure 5.8). Les francophones natifs produisent généralement plus de contours montants aux frontières de IP que de AP bien que ce ne soit pas le cas pour le sujet FR05. Ils produisent aussi une majorité de contours montants à ces positions prosodiques (56 à 69% pour les AP et 58 à 100% pour les IP). Chez les apprenants, en revanche, nous pouvons observer différentes stratégies :

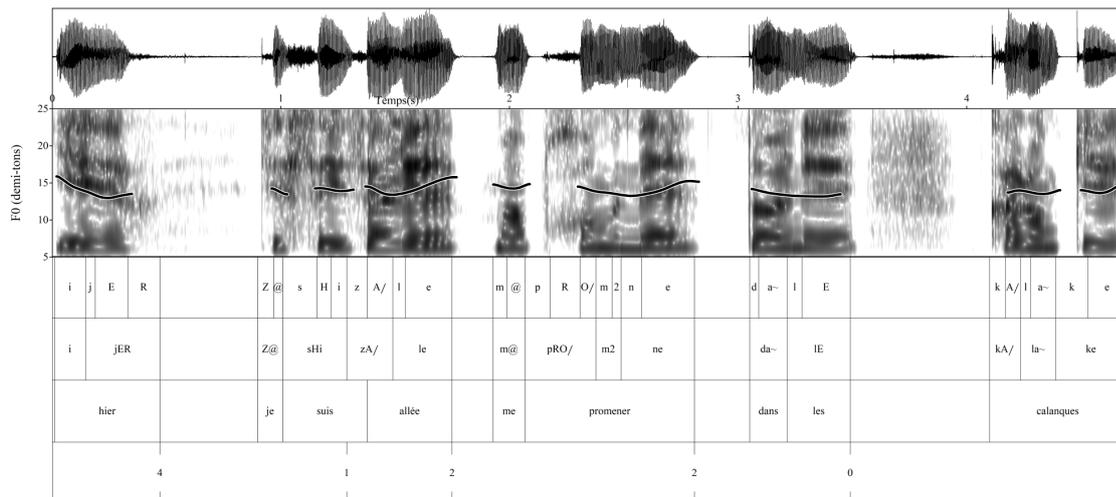


FIGURE 5.7 – Réalisation de l'énoncé « Hier, je suis allé me promener dans les calanques » par JP20

1. deux apprenants ont produit moins de 30% de contours montants à ces deux positions prosodiques (exemple sur Fig. 5.8 : JP9).
2. deux apprenants ont produit de 50 à 55% de contours montants à ces deux positions prosodiques (exemple sur Fig. 5.8 : JP16)
3. un apprenant a réalisé dans moins de la moitié des cas une montée mélodique en frontière de IP (exemple sur Fig.5.8 : JP20)
4. le reste des apprenants (12 apprenants, répartis dans les groupes A2-B1 et B1-B2) ont produit un pourcentage élevé de contours montants avec intervalle encore plus haut que celui des natifs (71 à 90% pour les AP and 75 à 100% pour les IP)

Les apprenants utilisant les stratégies 1 et 2 étaient soit dans le groupe de débutants A1-A2 soit dans le groupe A2-B1.

Un exemple de production du locuteur JP9 est présenté ci-dessus en Figure 5.9. Nous pouvons observer qu'à chaque frontière prosodique, AP et ip dans l'exemple, il ne produit pas de contour montant et produit soit un contour plat (fin de l'AP « prévoit ») soit un contour descendant (fin de l'AP « wagon »). Dans notre étude, nous n'avons pas différencié

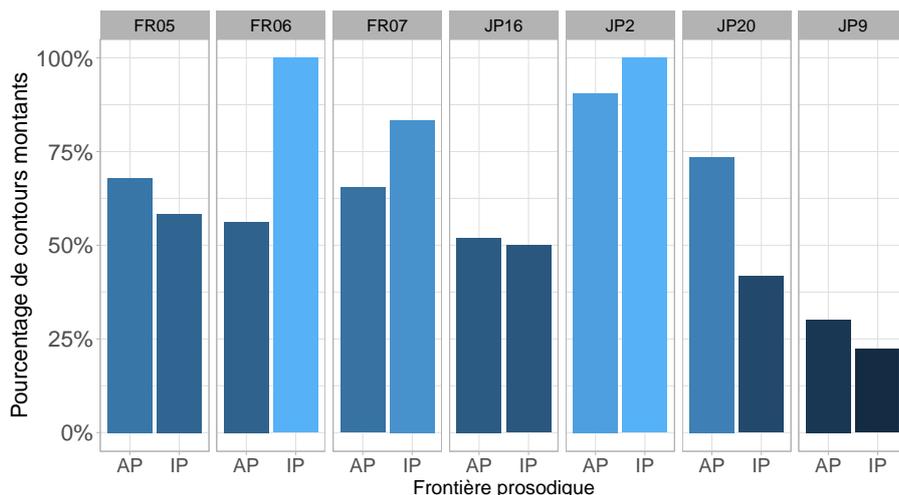


FIGURE 5.8 – Exemples de productions de contours montants de 4 apprenants et 3 francophones

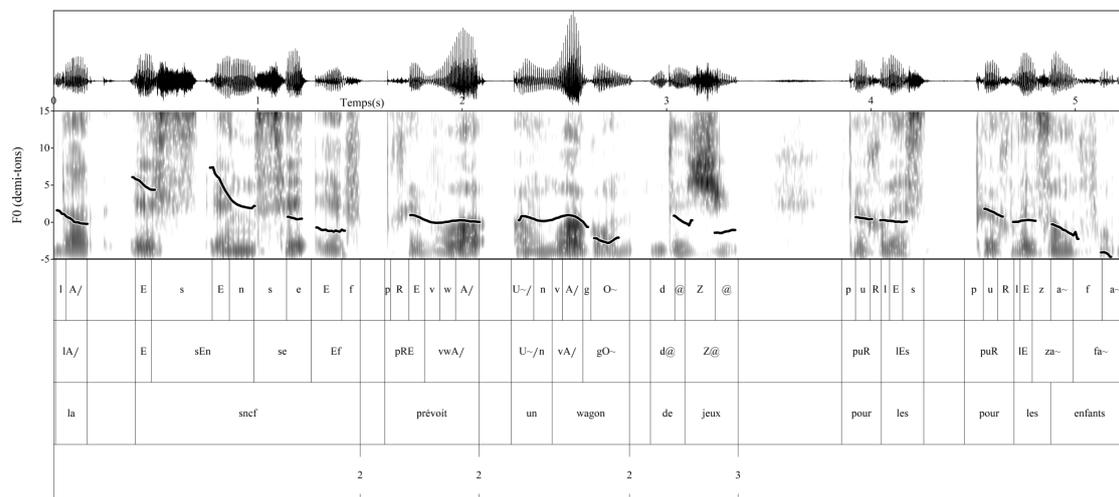


FIGURE 5.9 – Exemple de production du locuteur JP9

ces types de contours non-montants mais nous constatons que parmi les deux locuteurs produisant peu de contours montants (JP9 et JP18), le locuteur JP9 tend à produire plus de contours descendants, tandis que le locuteur JP18 produit une majorité de contours plats.

### Analyse des données de F0

Nous avons vu que les apprenants étaient capables de produire des contours montants aux frontières prosodiques fortes et que seuls certains sujets montraient des taux

de contours montants plus bas que ceux du groupe contrôle. Cependant, il est nécessaire d'analyser plus précisément la réalisation phonétique de ces contours afin d'observer si les apprenants sont capables de distinguer les différentes frontières prosodiques et d'y associer une montée mélodique plus ou moins importante. Nous prédisons également la présence de contours extra-montants aux frontières comme il a été observé dans des études antérieures. Nous allons maintenant analyser les données de F0 des contours montants réalisés avec un écart d'au minimum 1 demi-ton entre la voyelle accentuée et celle qui précède. Les résultats de F0 sont présentés dans la Figure 5.10. Comme attendu, on observe que les francophones natifs ont tendance à produire un écart de F0 plus important aux frontières des niveaux prosodiques plus élevés. La différence est très nette entre les frontières de AP et de IP où les natifs produisent des montées mélodiques plus importantes aux frontières de IP que de AP. Les apprenants, quant à eux, semblent ne pas faire cette distinction et produisent des montées similaires aux trois frontières prosodiques fortes. Seule une faible tendance à produire des montées plus importantes en IP est visible chez les apprenants débutants.

Nous avons construit un modèle linéaire mixte avec l'écart de F0 comme variable dépendante, une interaction entre le niveau de langue et le type de frontière prosodique comme effets fixes ainsi que item et sujet comme effets aléatoires. L'interaction entre le niveau de langue et le type de frontière prosodique était significative ( $\chi^2(12)=34.09$ ,  $p<.001^{***}$ ). A nouveau, nous avons utilisé le package *emmeans* pour analyser les différents contrastes de l'interaction. Les résultats nous ont confirmé qu'il n'y avait pas de différence significative entre les montées de F0 réalisées en frontière de AP, ip et IP chez les apprenants alors que les francophones natifs distinguent les contours en frontière de IP des AP et ( $p<.001^{***}$ ) et ip ( $p = 0.015^*$ ). Cependant, nous n'avons pas observé de différence significative entre les frontières de AP et de ip.

Des exemples des écarts de F0 produits par quelques sujets aux frontières de AP et de IP sont présentés dans la Figure 5.11. Nous pouvons y observer que tous les francophones natifs (en haut de la figure) distinguent clairement les deux niveaux prosodiques en produisant une montée de F0 plus importante en frontière de IP que de AP. Cette distinction est la moins claire chez le sujet FR05 puisque l'on voit que les boxplots se chevauchent. Parmi les apprenants, seuls quatre d'entre deux montrent un schéma semblable aux natifs

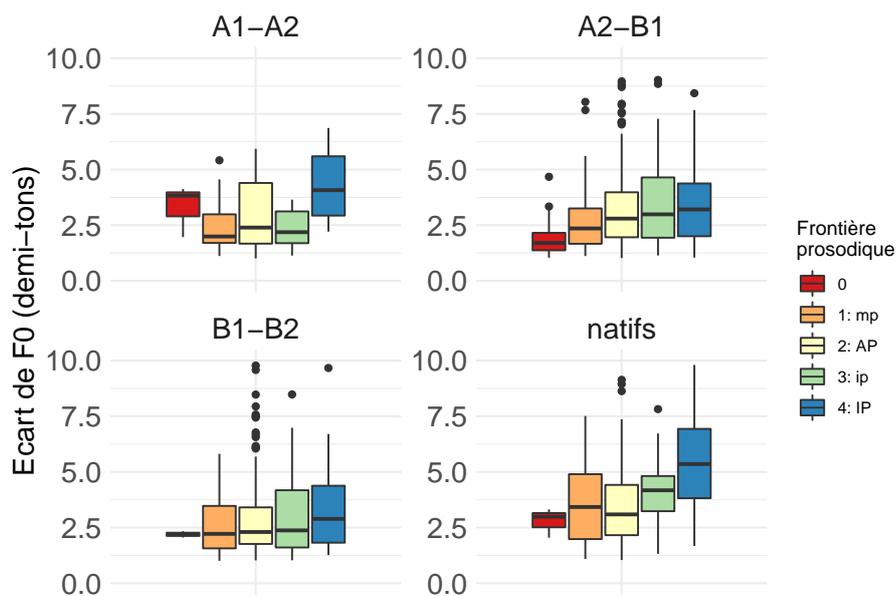


FIGURE 5.10 – Ecart de F0 suivant la frontière prosodique et le niveau de langue (en demi-tons)

(exemples dans la Figure 5.11 : JP16 et JP26) et le reste d’entre eux ont produit un contour semblable aux deux positions prosodiques (exemple dans la Figure 5.11 : JP24)

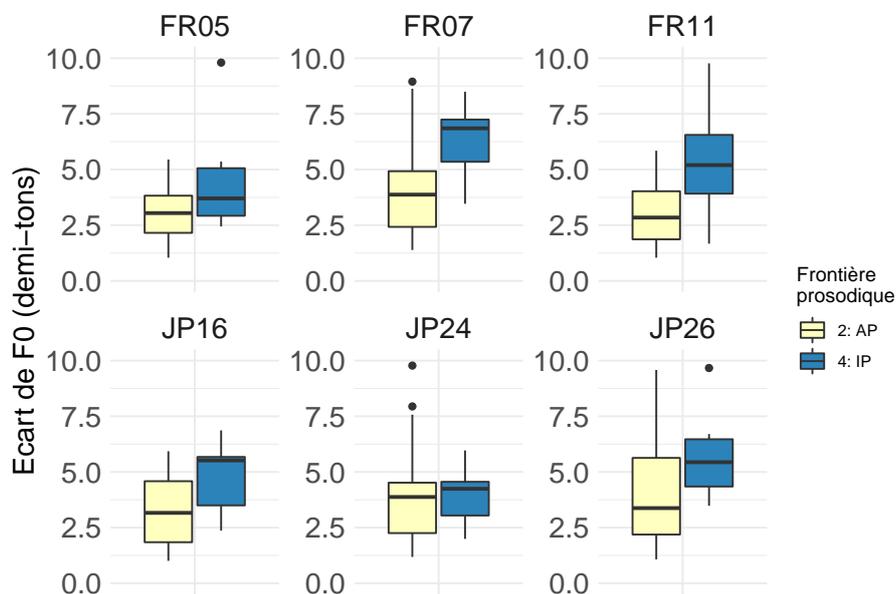


FIGURE 5.11 – Ecart de F0 produit par 6 sujets aux frontières prosodiques de AP et IP (en demi-tons)

Il apparaît donc que les apprenants ne distinguent pas les frontières prosodiques dans

leur production de la F0 et produisent une montée équivalente aux frontières de AP et de IP. Nous pouvons observer cela dans les exemples présentés en Figure 5.12 et 5.13 montrant la lecture de la phrase « Il fait un temps magnifique et la mer est très bonne » par un japonophone et un francophone. Nous pouvons observer que l'apprenant produit une faible montée mélodique en IP sur « magnifique » comparable aux montées produites en frontière de AP (fait ici produit **faut** et **temps**) là où le francophone du groupe contrôle produit une montée plus importante en frontière de IP. Ici, on observe également une différence au niveau de la durée puisque l'allongement vocalique produit par le japonophone en IP est moins prononcé que celui du francophone.

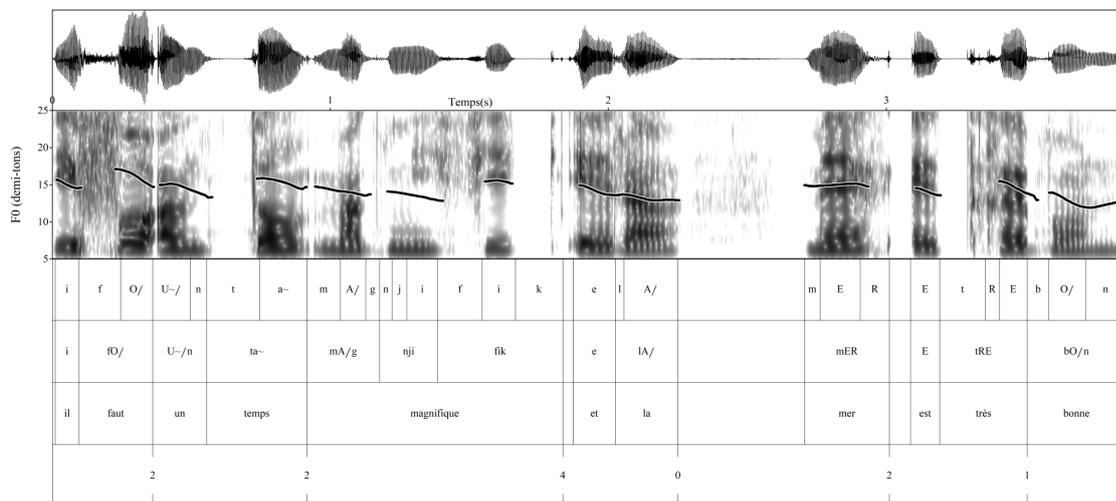


FIGURE 5.12 – Réalisation par un apprenant (JP5) de la phrase « Il fait un temps magnifique et la mer est très bonne »

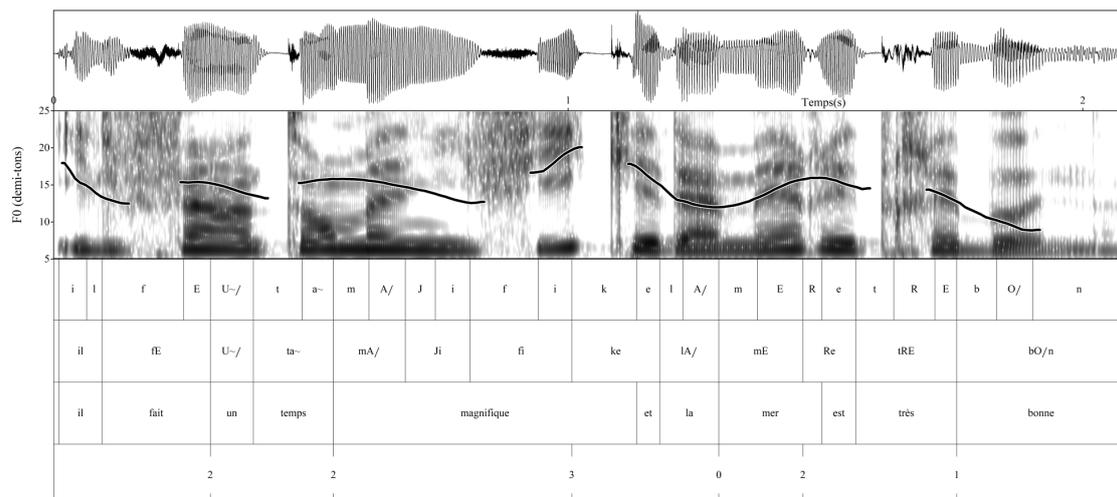


FIGURE 5.13 – Réalisation par un francophone (FR11) de la phrase « Il fait un temps magnifique et la mer est très bonne »

Notre [seconde hypothèse](#) prédisait la présence de contours extra-montants, notamment aux frontières de IP. Cependant, nous n'avons pas pu observer cette stratégie dans nos données. Au contraire, même, nous avons pu observer que les apprenants avaient tendance à produire des montées moins marquées que celles des francophones (voir Figure 5.10), en particulier aux frontières de IP. Les francophones ont produit des montées mélodiques à ces frontières de IP de 5.5 demi-tons en moyenne, tandis que la moyenne des apprenants est de 4.3 demi-tons pour le groupe A1-A2, 3.4 demi-tons pour les niveaux A2-B1 et B1-B2. Cette différence entre les apprenants et les natifs est significative pour le groupe A2-B1 ( $p=.001^{**}$ ) et le groupe B1-B2 ( $p=.006^{**}$ ), mais ne l'est pas pour le groupe de débutants A1-A2. Cependant, cela peut s'expliquer encore une fois par le nombre réduit d'apprenants dans ce groupe. En plus de cette tendance à produire des montées moins importantes de façon générale, le nombre de contours extra-montants (+ de 9 demi-tons) était très réduit ; 17 occurrences uniquement, soit 1,4% des contours conservés pour l'analyse de la F0, sachant que 7 d'entre elles ont été produites par des francophones du groupe contrôle. Il est toutefois vrai que l'on peut parfois observer des stratégies semblables à celles décrites dans les études antérieures en prosodie L2, soit la segmentation en courts *chunks* suivis de pauses et la présence de fortes montées mélodiques (bien qu'étant rarement supérieures à 9 demi-tons). L'exemple présenté en Figure 5.14 illustre bien ce type de stratégie, qui par

ailleurs semble être plus fréquente lorsque les énoncés sont longs et complexes. Cependant, cette stratégie bien que réelle ne peut être généralisée et elle semble à la fois liée à la complexité de l'énoncé et à l'apprenant. Chez certains apprenants (en particulier JP18 et JP39) nous observons même une toute autre stratégie : des mouvements prosodiques très réduits et par conséquent une faible étendue mélodique (voir Figure 5.15). Encore une fois cela ne semble pas lié à leur niveau de langue puisque l'un d'entre eux est débutant et a de réelles difficultés à s'exprimer (JP18) tandis que l'autre est de niveau B1-B2 et s'exprime de façon fluide en français.

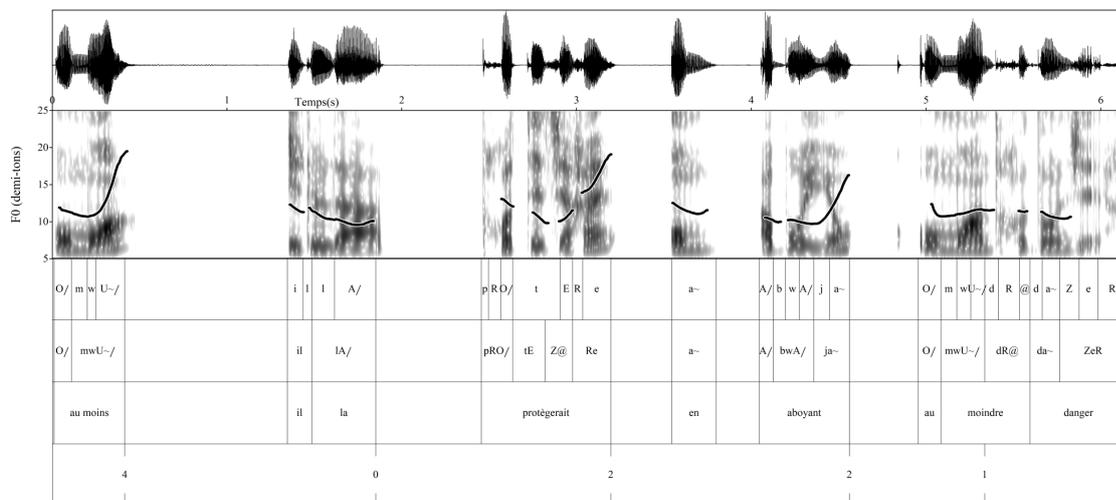


FIGURE 5.14 – Exemple de réalisation de découpage en chunks suivis de pauses avec montées mélodiques importantes par l'apprenant JP26 phrase « Au moins, il la protégerait en aboyant au moindre danger »

Les apprenants semblent donc avoir des difficultés à gérer la F0 de la même façon que les francophones natifs, et en ce qui concerne les contours continuatifs, à produire une montée mélodique suffisante en frontière de IP. Dans la section suivante, nous allons nous intéresser à leur production de l'indice de durée qui joue également un rôle essentiel dans la production des contours continuatifs.

### Analyse de l'indice de durée

Nous allons maintenant analyser les données de durée vocalique afin d'observer si les apprenants japonophones distinguent les différentes forces de frontière en produisant un

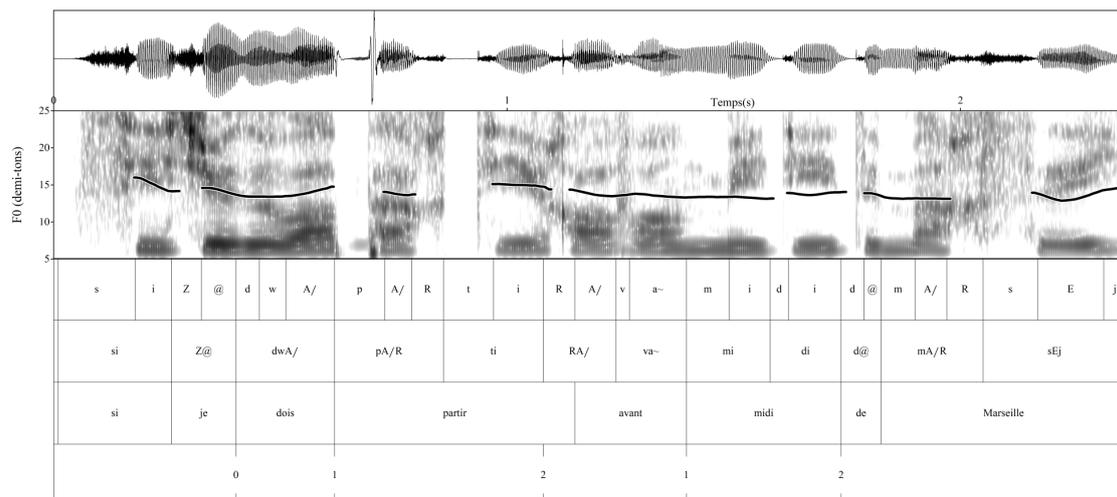


FIGURE 5.15 – Exemple de production avec faible étendue mélodique, produit par l'apprenant JP39, phrase « Si je dois partir avant midi de Marseille »

allongement vocalique plus ou moins important. En effet, en français, les contours continuatifs sont produits avec un allongement de la syllabe accentuée dont la durée est plus ou moins longue selon le type de frontière prosodique. En particulier, l'allongement produit en frontière de IP est généralement plus long que celui réalisé en frontière de AP.

Nous pouvons voir sur la Figure 5.16 que des apprenants semblent produire des allongements vocaliques plus importants aux frontières prosodiques fortes (AP, ip et IP) que faibles (pas de frontière prosodique ou MP) et donc distinguer dans une certaine mesure les types de frontières prosodiques. En revanche, de façon très semblable à l'analyse de la fréquence fondamentale, les apprenants produisent des allongements équivalents de la voyelle accentuée à toutes les frontières prosodiques fortes, à l'inverse des natifs qui les distinguent et produisent notamment des allongements plus importants en frontières de IP et ip que de AP. Afin de confirmer ces observations, nous avons construit un modèle linéaire mixte avec la durée de la voyelle accentuée comme variable dépendante, le débit de parole par phrase et une interaction entre le niveau de langue et le type de frontière prosodique comme effets fixes ainsi que item, voyelle et sujet comme effets aléatoires. L'interaction entre le niveau de langue et la frontière prosodique est significatif ( $\chi^2(12) = 63.31$ ,  $p < .001^{***}$ ). Nous avons approfondi l'analyse de l'interaction avec le package *emmeans* qui a confirmé nos observations : les francophones natifs ont produit des voyelles significati-

vement plus longues aux frontières de ip ( $p < .004^{**}$ ) et de IP ( $p = .019^*$ ) que de AP, tandis que l'on n'observe aucune différence significative chez les apprenants.

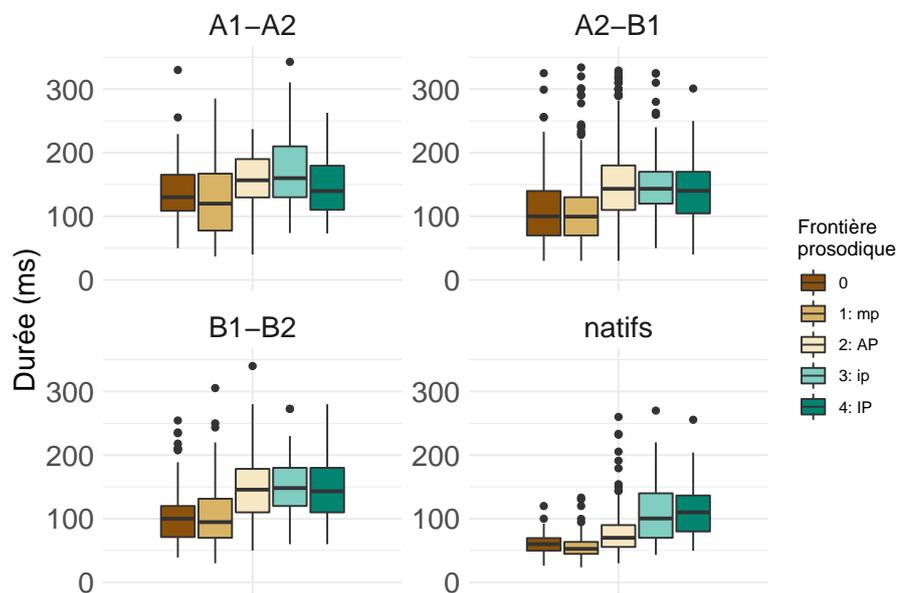


FIGURE 5.16 – Durée des voyelles accentuées selon la frontière prosodique et le niveau de langue (en millisecondes)

Sur la Figure 5.17, nous présentons la durée de voyelles accentuées en frontières AP et IP produites par certains locuteurs. Nous pouvons observer que, tout comme nous l'avons précédemment vu avec les données de F0, tous les francophones natifs distinguent les deux niveaux. A nouveau, cette distinction est moins claire chez le sujet FR05. Chez les apprenants nous observons plusieurs schémas :

- deux apprenants montrent un schéma comparable à ceux produits par les natifs avec une voyelle plus longue en IP qu'en AP (exemple dans la Fig. 5.17 : JP39)
- cinq apprenants ont produit un schéma inverse en produisant des voyelles plus longues en frontière de AP que de IP (exemple sur la Fig 5.17 : JP18)
- le reste des apprenants ont produit un allongement similaire aux deux positions prosodiques (exemple dans la Fig 5.17 : JP5).

Le niveau de langue ne semble pas avoir d'influence sur le type de schéma produit.

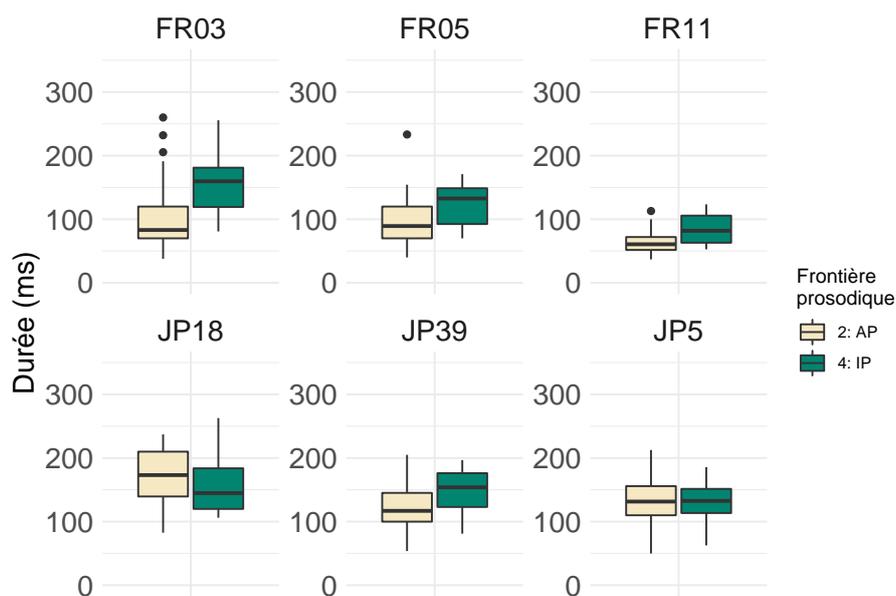


FIGURE 5.17 – Durée des voyelles accentuées de 6 sujets aux frontières prosodiques de AP et IP

### Allongement vocalique et ratio

Lors de l'analyse de la durée des voyelles accentuées, nous avons observé que les apprenants produisaient des allongements similaires aux frontières de AP, ip et IP, là où les francophones produisaient un allongement plus important en IP qu'en AP. Cependant, nous ne pouvons pas comparer les allongements des apprenants à ceux des natifs tels quels étant donné que le débit de parole diffère grandement. En effet, les apprenants ont un débit de parole bien plus réduit que celui des natifs (de 3,7 à 4,2 syllabes/seconde pour les groupes d'apprenants, contre 6,1 syllabes/seconde pour les francophones). Nous avons donc choisi d'observer le ratio des voyelles « accentuées » en frontières 1, 2, 3, 4 sur les voyelles non-accentuées aux frontières 0, afin de permettre une meilleure comparaison des allongements vocaliques suivant les groupes de langue. Ces résultats sont présentés dans la Figure 5.18.

Nous pouvons observer que les apprenants produisent des allongements vocaliques aussi importants que les natifs en frontière de AP, mais clairement plus réduits que chez les natifs aux frontières des niveaux prosodiques supérieurs ip et IP. En 5.3.3, nous avons pu observer que les voyelles réalisées aux frontières de AP, ip et IP n'étaient pas signifi-

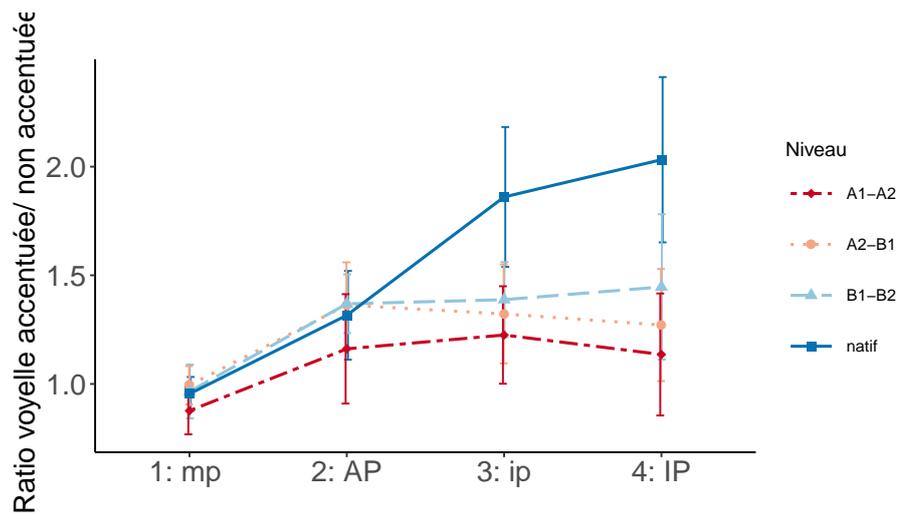


FIGURE 5.18 – Ratio voyelles accentuées/non accentuées selon le niveau de langue

cativement différentes chez les apprenants, et ici nous pouvons voir que le ratio de durée vocalique produit à toutes ces frontières correspond à un ratio en frontière de AP chez les natifs. Cela signifie que les apprenants produisent un allongement vocalique comparable à celui des natifs en AP mais qu'ils ne produisent pas assez d'allongement aux frontières de ip et IP. Il faut tout de même noter que les apprenants, comme nous l'avons vu précédemment, avaient un débit de parole bien plus réduit que celui des natifs et nous ne savons pas en quoi cela pourrait influencer les résultats et si l'on observerait les mêmes ratios chez les francophones ayant un débit de parole plus lent. En effet, il est probable que les allongements des syllabes accentuées en frontière de AP soient plus importants avec un débit de parole plus lent puisque nous avons choisi d'analyser toutes les frontières prosodiques où un contour continuatif est *attendu* mais n'est pas obligatoirement réalisé par les participants. En effet, en français, lorsque le débit de parole est plus lent les locuteurs ont tendance à produire plus de AP (Martin, 2018). Ainsi, le ratio obtenu pour la position AP chez les francophones serait de ce fait plus élevé. Cependant, cela ne changerait pas nos principales observations et seule une différence entre les apprenants et les francophones apparaîtrait aux frontières de AP.

Les Figures 5.19 et 5.20 sont des exemples représentatifs des différences que l'on observe entre productions natives et non-natives. On y observe un allongement bien moins prononcé sur la frontière de IP (**Marseille**) chez l'apprenante que chez la francophone. Pour donner un ordre d'idée, le mot accentué « Marseille » correspond à 23% de l'énoncé

chez l'apprenante contre 34% chez la locutrice francophone.

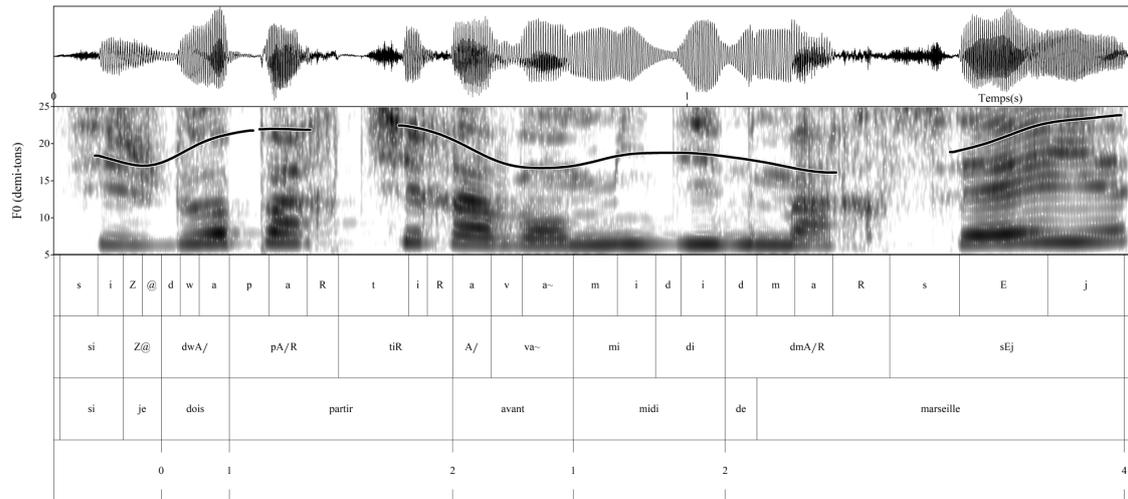


FIGURE 5.19 – Réalisation par une francophone (FR07) de l'énoncé « Si je dois partir avant midi de Marseille »

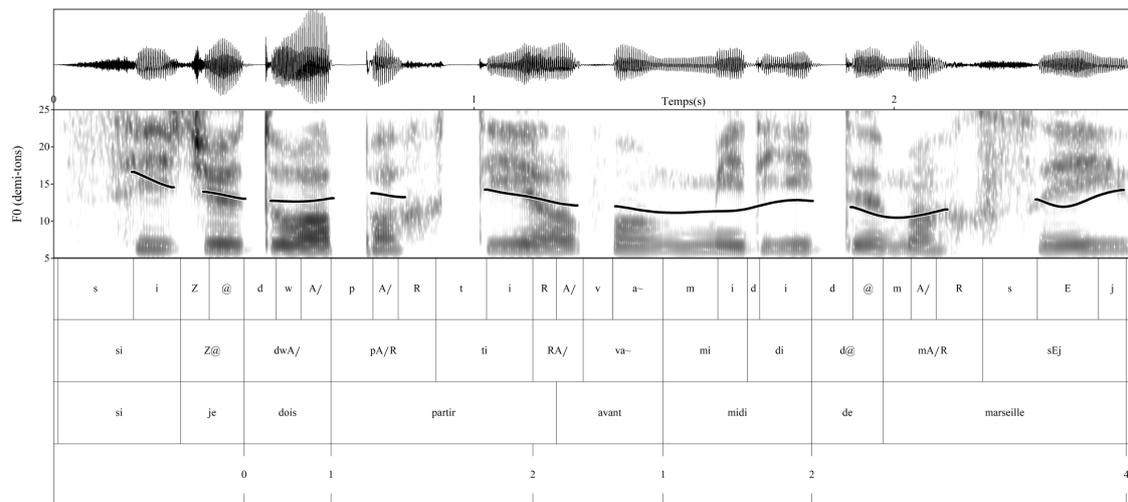


FIGURE 5.20 – Réalisation par une apprenante japonophone (JP15) de l'énoncé « Si je dois partir avant midi de Marseille »

### 5.3.4 Synthèse

Dans cette expérience de lecture de textes, nous avons analysé les réalisations des contours continuatifs chez des apprenants japonophones et des francophones; en parti-

culier les pourcentages de contours montants produits et leur gestion de la durée et de la F0. Nous présentons ici nos principaux résultats. En premier, les résultats sur les montées mélodiques :

- Les apprenants japonophones n'ont, de façon générale, pas de difficultés à produire des montées aux frontières prosodiques fortes (AP, ip et IP).
- Les apprenants généralisent ces montées de la F0 à toutes les positions prosodiques fortes là où les productions des francophones varient selon les items
- Nous avons cependant observé différentes stratégies chez les apprenants et certains d'entre eux produisent peu de montées mélodiques aux frontières fortes. Ces apprenants font partie des deux premiers groupes de niveau (A1-A2 et B1-B2).

L'analyse de la F0 nous a montré les résultats suivants :

- Les apprenants japonophones produisent des montées de F0 similaires aux frontières de AP, ip et IP tandis que les francophones distinguent les frontières en produisant des montées plus importantes aux frontières de IP qu'aux frontières de AP et de ip.
- Nous n'avons pas observé plus de contours extra-montants chez les apprenants que chez les natifs, et ceux-ci étaient très rares. A l'inverse, les apprenants ont produit des montées mélodiques moins marquées que celles des francophones aux frontières de IP.

Enfin, concernant l'indice de durée de la voyelle accentuée aux frontières prosodiques, nous avons observé que :

- Les apprenants produisent des voyelles de durée similaire aux frontières de AP, ip et IP, tandis que les francophones produisent des voyelles accentuées plus longues aux frontières de ip et de IP qu'aux frontières de AP. Un certain nombre d'apprenants tendent même à produire des voyelles plus courtes en frontière de IP que de AP.
- L'observation du ratio voyelle accentuée / voyelle non accentuée nous indique que l'allongement vocalique produit par les apprenants japonophones aux frontières de

ip et de IP est insuffisant si on le compare aux allongements produits par les francophones.

Ainsi, notre [première hypothèse](#) concernant un transfert négatif de la L1 sur les productions des japonophones est partiellement confirmée. En effet, nous avons bien pu observer des difficultés à distinguer les frontières prosodiques et à produire différentes forces de continuation par la gestion de la durée et de la F0. Cependant, nous n'avons pas observé de difficultés à produire des montées mélodiques, à l'exception de quelques apprenants, alors que nous nous attendions à observer moins de contours montants à ces positions prosodiques chez les apprenants que chez les francophones.

Notre [seconde hypothèse](#) concernant les stratégies L2 indépendantes de la L1 n'a que partiellement été confirmée. En effet, nous n'avons pas observé une propension des apprenants à produire des contours extra-montants aux frontières de IP, comme il a pu être observé dans de précédentes études. Au contraire, les montées produites par les japonophones étaient même inférieures à celles des natifs à ces frontières. Cependant, nous avons pu observer une variété de stratégies individuelles, dont certaines semblent plus relever de l'interlangue que d'un simple calque de la L1. Nous avons entre autres observé des découpages de la parole en *chunks* associés à des montées mélodiques importantes (bien que souvent inférieures à 9 demi-tons) et suivis de pauses, qui peuvent être interprétés comme une certaine insécurité linguistique. Les différentes stratégies observées ne sont cependant pas généralisées et dépendent de la variation individuelle et des énoncés.

Notre [troisième hypothèse](#) sur le rôle du niveau de langue sur les résultats est plus difficile à vérifier. En effet, nous avons observé des taux bas de contours montants aux frontières AP et IP dans le groupe débutant A1-A2. Il semblerait donc qu'il existe un transfert négatif au niveau des montées mélodiques aux premières étapes de l'apprentissage. Cependant, nous n'avons que deux apprenants dans le groupe de débutants et il est donc difficile de généraliser ces résultats, bien que ceux-ci aillent dans le sens de notre hypothèse. Enfin, nous n'avons pas pu observer de lien entre le niveau de langue et la gestion de la F0 et de la durée. Ainsi, s'il existe une amélioration de la production des continuateurs français avec le niveau de langue, elle ne concernerait que la capacité à produire des montées mélodiques en excluant la gestion fine des indices de durée et de F0.



## 5.4 Expérience 2 : Description d'images

### 5.4.1 Introduction

La tâche de lecture de textes, que nous avons analysée précédemment, reste un exercice de style particulier demandant une interprétation d'énoncés déjà construits, et ne correspond pas forcément aux productions des apprenants dans la vie quotidienne. Ainsi, il nous paraît nécessaire d'analyser, en parallèle des résultats décrits précédemment, une parole plus spontanée, qui demanderait des capacités langagières différentes de celles utilisées en lecture de texte ; il s'agirait pour les apprenants non plus de mettre en parole des énoncés mais de les construire eux-mêmes et ainsi de gérer le rythme et l'intonation sans autre contrainte que leurs compétences linguistiques. Dans cette partie, nous allons analyser les résultats de l'expérience de description d'images, qui correspond à une tâche de parole semi-spontanée. Cette tâche a le mérite de faciliter la parole des apprenants en les guidant par la présence de l'image comme support, tout en leur permettant une certaine liberté d'expression. Il ne s'agit cependant pas d'un échange réel et cette tâche reste toutefois assez artificielle. Ainsi, nous ne sommes pas en mesure de qualifier ce type de parole comme « spontané », d'où l'utilisation du terme « semi-spontané ». Cependant, nous considérons que ce type de parole est fondamentalement différent de la lecture de textes, et que si cette tâche ne permet pas une interaction réelle, elle met toutefois en avant les composantes discursives que l'on peut retrouver en parole spontanée. Par exemple, les montées mélodiques en fin de IP n'ont plus une simple fonction phonologique, mais permettent surtout d'indiquer la continuité, ou le non-achèvement, de l'acte d'énonciation. De plus, la description d'images étant un exercice que l'on retrouve fréquemment dans les cours de Français Langue Étrangère, l'analyse de ces résultats a également un intérêt didactique indéniable.

Bien que nous suivions les mêmes hypothèses que pour l'expérience de lecture de textes, nous n'avancions pas exactement les mêmes prédictions en raison des différences entre les types de parole. Nos prédictions issues de l'[hypothèse 1](#) (transfert de la L2) sont les suivantes :

- Les apprenants japonophones vont produire moins de contours montants aux frontières prosodiques que les francophones natifs en raison de leur caractère facultatif en japonais spontané.

- Les apprenants auront des difficultés à distinguer les différents niveaux prosodiques dans leur production de la durée ; ils produiront moins d’allongements marqués en frontière de IP que les francophones. Concernant la gestion de la F0, l’absence de données phonétiques des BPM non-finaux en japonais, à notre connaissance, ne nous permet pas de formuler de prédictions sur ce point. Cependant, en ce qui concerne la durée, nous prédisons que les apprenants japonophones auront moins de difficultés à produire un allongement plus conséquent en frontière de IP puisqu’en japonais une durée plus longue de la dernière more ou du dernier phone est reliée à la continuation de la parole (Koiso et al., 1998 ; Yoneyama et al., 2003).

Nos prédictions sont identiques à celles de lecture de texte pour l’hypothèse 2 soit le fait que d’autres stratégies non imputables au transfert comme la surgénéralisation des faits prosodiques soient mises en place par les apprenants. Nous nous attendons donc à observer des stratégies communes aux productions des apprenants L2, en particulier la production de contours extra-montants aux frontières de IP.

Enfin, nous prédisons une meilleure production des continuatifs chez les apprenants de niveau plus avancé (hypothèse 3) c’est-à-dire une plus grande quantité de contours montants et une meilleure gestion des indices prosodiques.

#### 5.4.2 Méthodologie

Les données des descriptions d’images ont été recueillies selon le protocole décrit en 5.2.3. Les instructions données aux participants étaient de décrire le plus longuement possible et avec leurs propres mots les huit images présentées une par une, sous format papier. Pour notre analyse, nous avons sélectionné quatre images (dessins ou photos) représentant des scènes de vie quotidienne (une scène dans une épicerie, dans une rue passante, et un goûter d’anniversaire) ainsi qu’un tableau de Van Gogh (une chambre) (voir Annexe C). Pour le groupe contrôle, nous avons cette fois-ci enregistré nous-même 4 francophones natifs, 2 femmes et 2 hommes. Ce groupe contrôle est donc différent de celui de la tâche de lecture de textes. Nous avons ensuite procédé à une analyse similaire à celle faite précédemment pour la lecture de textes. Puisque nous ne pouvions pas décider en amont d’une segmentation phonologique du fait qu’il s’agit cette fois-ci de parole (semi-)spontanée, nous avons décidé d’annoter manuellement différentes frontières prosodiques, indépen-

damment de leur réalisation phonétique. Ainsi, nous avons différencié les frontières prosodiques non-finales des frontières finales.

Pour la segmentation, nous avons utilisé l'annotation semi-automatique de SPPAS (Bigi, 2015) et nous l'avons ensuite corrigée avec Praat (Boersma, 2001). Nous avons ensuite rajouté deux tires ; la première contient une annotation ToBI des énoncés et la deuxième contient un découpage des différents niveaux prosodiques. Nous nous sommes basés sur les critères d'identification des contours montants donnés dans Portes et al. (2007). Nous avons donc distingué les syntagmes intonatifs (IP) terminaux et non-terminaux ainsi que les énumérations. Pour la présente analyse, nous avons choisi de regrouper deux niveaux : le niveau 2 correspondant aux frontières de AP (frontières mineures), et le niveau 4 correspondant aux frontières de IP en position finale d'énoncé (terminaux) et d'énumération. Les voyelles restantes non porteuses d'un accent H ont automatiquement été notées comme de niveau 0 soit non accentuées. Les mesures acoustiques de la durée des voyelles ainsi que de F0 à différents pourcentages de la voyelle ont été extraites automatiquement à l'aide d'un script Praat. La différence de F0 entre la voyelle accentuée et la voyelle de la syllabe précédente a également été calculée (valeurs à 75% de la voyelle) et les contours ont été classifiés comme montants lorsqu'ils étaient audibles et que cet écart était supérieur à 1 demi-ton. Les hésitations produites par les locuteurs ont également été annotées ainsi que la présence de voix craquée afin de les exclure respectivement de l'analyse de la durée et de la F0. Les derniers contours de chaque description d'image ont également été écartés de l'analyse. En effet, même si ces contours correspondent à des frontières de IP, il n'ont pas le même rôle continuatif que les contours précédents puisqu'ils permettent ici de conclure la description de l'image, et sont généralement réalisés comme descendants. Tout comme nous l'avons fait pour la précédente analyse, avons procédé à une annotation phonologique avant nos analyses et défini les frontières prosodiques *attendues* et non nécessairement réalisées telles quelles par les locuteurs afin de comparer au mieux les productions des apprenants avec celles des francophones natifs. Enfin, les données recueillies ont été traitées avec le logiciel R R Core Team, 2018 pour l'analyse statistique.

Un exemple de segmentation Praat est présenté en Figure 5.21. Nous pouvons y voir deux types de frontières IP : à la fin de l'ajout à gauche « au premier plan de l'image » et tout à la fin de l'énoncé. Le premier type de frontière n'a donc pas été pris en compte dans

la présente analyse puisque nous l’avons classé comme « non terminal » et que seuls les contours terminaux ayant une fonction discursive continue (ici celui en fin d’énoncé) ont été retenus. Nous remarquons aussi que les frontières annotées ne sont pas toujours associées à la réalisation phonétique attendue. En effet, comme expliqué ci-dessus, il s’agit d’annotations phonologiques et non phonétiques. Cela est bien sûr à prendre en compte dans l’interprétation des résultats, notamment pour les résultats concernant les frontières de AP.

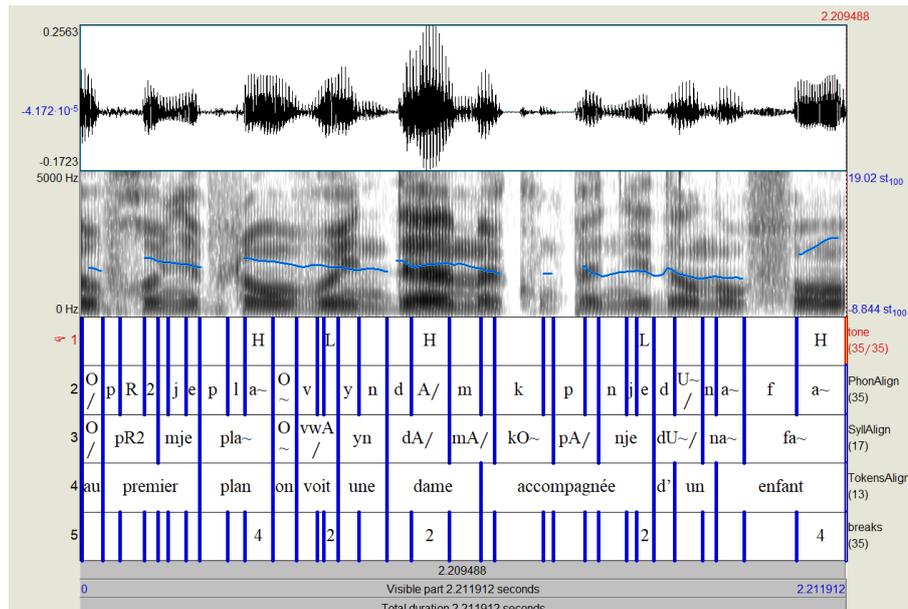


FIGURE 5.21 – Exemple de segmentation Praat pour la tâche de description d’images

### 5.4.3 Résultats

#### Observations générales

Nous présentons ici (Tableau 5.3) des informations générales sur le débit de parole, le nombre de mots et les pauses produites dans la tâche de description d’images selon le niveau de langue.

	Niveau de langue			
	A1-A2	A2-B1	B1-B2	natifs
débit de parole (syllabes / s)	3 (ET = 0.4)	3.4 (ET = 0.3)	3.4 (ET = 0.3)	4.6 (ET = 0.5)
% pauses frontière AP	50 % (ET = 48)	54 % (ET = 12)	42 % (ET = 20)	7 % (ET = 2.1)
% pauses frontière IP	89 % (ET = 2.6)	90 % (ET = 11)	95 % (ET = 6.8)	53 % (ET = 2.1)
% hésitations	23 % (ET = 6)	24 % (ET = 8)	16 % (ET = 4)	12 % (ET = 6)

TABLEAU 5.3 – Informations générales sur les productions de la tâche de description d'images

Nous pouvons observer que :

- Le débit de parole moyen est plus élevé chez les francophones que chez les apprenants, mais cette différence est moins importante que celle observée en lecture de textes entre les apprenants et le premier groupe contrôle (cf. 5.3.3).
- Les francophones produisent peu de pauses aux frontières de AP (7%) tandis que les apprenants associent ces frontières à une pause dans la moitié des cas (42-54%).
- Les apprenants japonophones produisent presque systématiquement une pause aux frontières de IP (89-95%) alors que l'on n'en observe que dans la moitié des cas chez les locuteurs du groupe contrôle (53%).
- Les francophones produisent légèrement moins d'hésitations que les apprenants et le groupe d'apprenants de niveau le plus élevé (B1-B2) produit en moyenne moins d'hésitations que les autres groupes d'apprenants.

Tout comme pour la tâche de lecture de textes (cf 5.3.3) on observe plus de pauses chez les apprenants, montrant leur tendance à découper leur production en courts *chunks* suivis de pauses. La principale différence entre les deux tâches est la différence moins importante entre le débit de parole des apprenants et du groupe contrôle en description d'images. Cette différence s'explique en particulier par un débit de parole moins rapide chez les locuteurs francophones de cette tâche que chez le groupe contrôle de la tâche de lecture de textes (-1.5 syllabes / seconde ; contre -0.7 chez les apprenants). Puisqu'il s'agit

de deux groupes contrôles différents, il est toutefois difficile de savoir s'il ne s'agit pas de différences individuelles. Dans tous les cas, le débit de parole moyen est plus lent en description d'images que ce soit chez les apprenants que chez les francophones.

### Résultats : les montées mélodiques

Les pourcentages moyens de montées mélodiques réalisées par les participants durant les descriptions d'images sont présentés dans la Figure 5.22. Nous pouvons observer que ces pourcentages sont très proches entre les groupes d'apprenants et le groupe contrôle, à l'exception des deux locuteurs du groupe de débutants (A1-A2) qui semblent produire moins de montées en frontière de AP. De manière générale et pour tous les groupes de langue nous observons des pourcentages de montées produites plus importants en frontière de IP que de AP. Nous avons construit un modèle linéaire généralisé mixte (GLMM) avec comme variable dépendante le contour prosodique observé (montant ou non, codé 1 ou 0), une interaction entre le niveau de langue et la frontière prosodique comme effets fixes et le locuteur comme effet aléatoire. L'interaction entre le niveau de langue et la frontière prosodique s'est révélée significative lors de notre test ANOVA ( $\chi^2(3) = 10.10, p = 0.018^*$ ). Nous avons donc mené une analyse plus poussée de cette interaction à l'aide du package *emmeans*. Cette analyse nous a confirmé que tous les groupes de langue distinguent bien les deux frontières en produisant significativement plus de montées mélodiques aux frontières de IP que AP ( $p < .001^{***}$  pour le groupe contrôle ainsi que les groupes A1-A2 et A2-B1, et  $p = 0.005^{**}$  pour le groupe B1-B2). Nous n'avons pas non plus observé de différences significatives entre le groupe contrôle et tous les groupes d'apprenants aux deux frontières prosodiques. Les seules différences observées sont en frontière de AP entre les apprenants du groupe A1-A2 et ceux du groupe A2-B1 et B1-B2 ( $p = 0.024^*$  et  $p = 0.013^*$ ).

Toutefois, il est important de noter que l'on observe une grande variabilité entre les locuteurs à la fois chez les apprenants et les francophones, notamment aux frontières de IP (voir Figure 5.23). En effet, même si l'on observe chez trois des francophones un pourcentage de contours montants élevé allant de 89 à 92%, la quatrième locutrice (FR2) a une stratégie très différente puisqu'elle produit seulement 37,5% de montées mélodiques. Si tous les apprenants produisent eux aussi une majorité de montées en frontières de IP, sur un intervalle plus étendu allant de 63 à 100%, l'un d'entre eux (JP9) ne produit que très peu

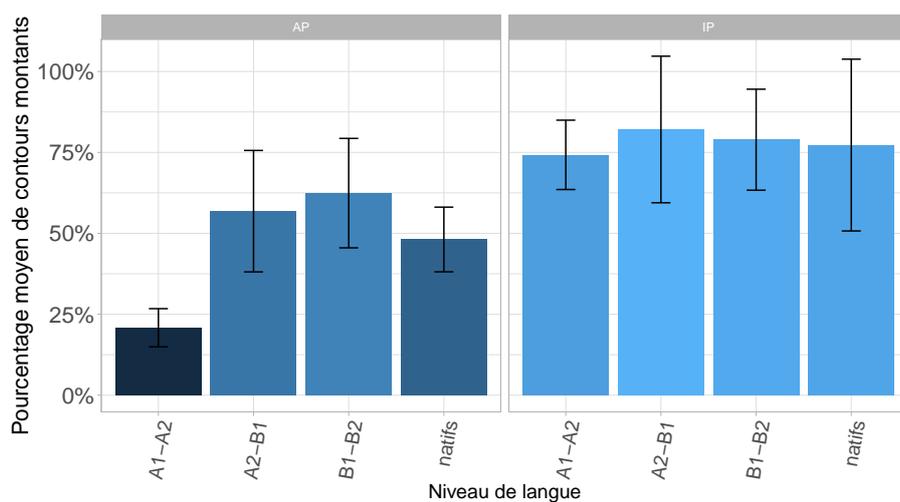


FIGURE 5.22 – Pourcentage de contours montants produits aux frontières de AP et IP selon le niveau de langue

de contours montants (25%).

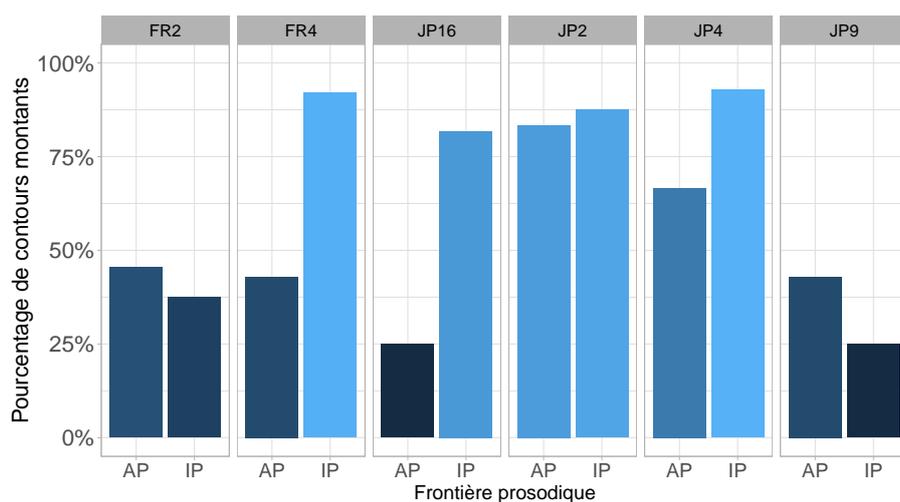


FIGURE 5.23 – Pourcentage de contours montants produits aux frontières de AP et IP : exemple de 6 locuteurs

On observe également chez certains apprenants des pourcentages de montées mélodiques très importants en frontière de AP comme de IP (exemple dans la Figure 5.23 : JP2). Ce phénomène peut ainsi s'apparenter à une surgénéralisation des contours montants aux frontières prosodiques du français

### Analyse des données de F0

Dans la section précédente, nous avons vu que les apprenants japonophones produisaient, de façon générale, autant de montées mélodiques aux frontières de AP et de IP que les francophones natifs du groupe contrôle. Dans cette section, nous nous intéressons à la réalisation phonétique de ces montées en analysant l'écart mélodique entre la voyelle accentuée et celle qui la précède (point situé à 75% de la voyelle). Les résultats de l'analyse de la F0 des contours montants (voir Figure 5.24) montrent que les frontières de IP sont associées à un écart de F0 plus important qu'aux frontières de AP, que ce soit chez les apprenants ou les natifs. Afin de vérifier ces observations, nous avons construit un modèle linéaire mixte avec comme variable dépendante l'écart de F0 en demi-ton, une interaction entre niveau de langue et la frontière prosodique comme effets fixes et le locuteur ainsi que la nature de la voyelle accentuée comme effets aléatoires. L'interaction entre le niveau de langue et la frontière s'est révélée significative lors de notre test ANOVA ( $\chi^2(3) = 27.38$ ,  $p < .001^{***}$ ). L'analyse des contrastes de cette interaction avec le package *emmeans* nous a montré que la différence d'écart de F0 entre les frontières AP et IP était significative pour tous les groupes de langue, à l'exception du groupe A1-A2 ( $p = 0.55$  pour A1-A2  $p < .001^{***}$  pour tous les autres groupes.)

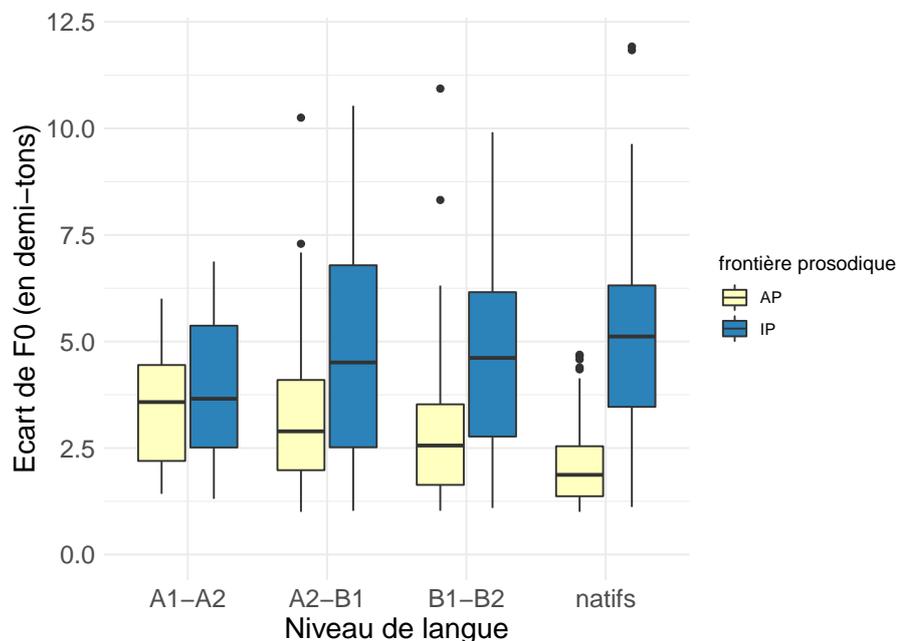


FIGURE 5.24 – Ecart de F0 suivant la frontière prosodique et le niveau de langue (en demi-tons)

Les apprenants sont ainsi, de façon générale, capables de produire des montées mélodiques plus importantes aux frontières de IP quel que soit leur niveau de langue. De plus, cette distinction est généralisée et nous n'observons que peu de variation entre les différents locuteurs : seulement 3 locuteurs montrent une tendance inverse avec des montées mélodiques légèrement moins importantes en frontière de IP (deux d'entre eux sont présentés dans la Figure 5.25 : JP19 et JP9).

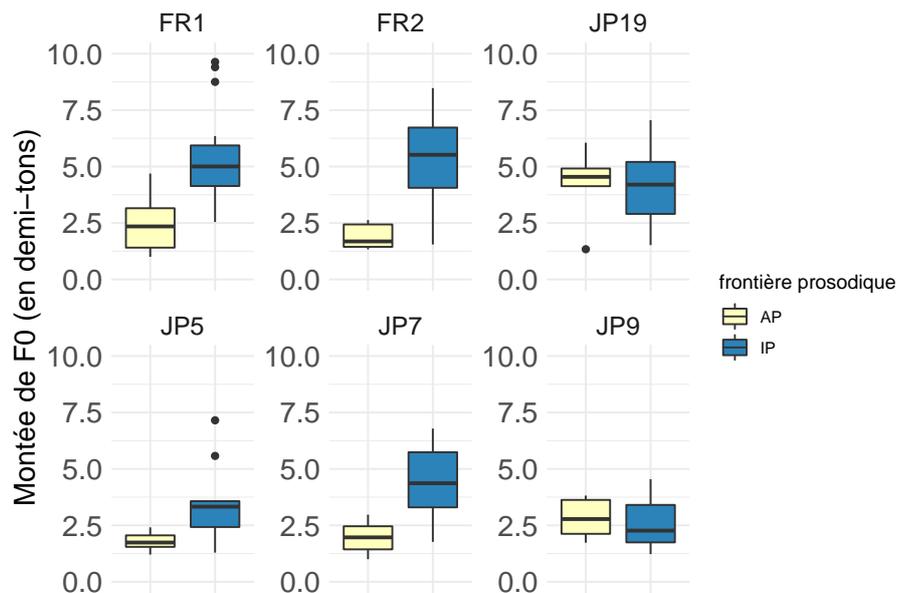


FIGURE 5.25 – Ecart de F0 produit par 6 sujets aux frontières prosodiques de AP et IP (en demi-tons)

Le nombre de contours extra-montants (+ de 9 demi-tons) observés est très faible : 13 occurrences uniquement, soit 2,3% des contours conservés pour l'analyse de la F0, sachant que 5 d'entre elles ont été produites par des francophones du groupe contrôle. Il n'y a pas non plus de différence significative entre les montées mélodiques des apprenants et des francophones aux continuations majeures ; les apprenants ne produisent donc pas de montées plus ou moins importantes que le groupe contrôle.

### Analyse de l'indice de durée

Si les résultats de l'analyse du F0 ont montré que les apprenants distinguent les frontières de AP des frontières de IP, il semblerait que cette distinction ne se retrouve pas dans les résultats de durée de la voyelle accentuée (voir Figure 5.26). En effet, les francophones

produisent des voyelles nettement plus longues en frontière de IP tandis que l'on n'observe qu'une faible tendance chez les apprenants. Pour confirmer ces observations, nous avons construit un modèle linéaire mixte avec comme variable dépendante la durée de la voyelle en milliseconde, une interaction entre le niveau de langue et la frontière prosodique comme effets fixes et le locuteur ainsi que la nature de la voyelle accentuée comme effets aléatoires. L'interaction entre le niveau de langue et le type de frontière prosodique est significative ( $\chi^2(6) = 190.67, p < .001^{***}$ ). L'analyse des contrastes de l'interaction va dans le sens des observations faites : les francophones natifs produisent des voyelles significativement plus longues en frontière de IP que de AP ( $p < .001^{***}$ ) mais ce n'est pas le cas des deux groupes intermédiaires (A2-B1 et B1-B2). Cette distinction significative est aussi présente dans le groupe de débutants ( $p = .001^{**}$ ) mais il faut noter que ce groupe étant seulement composé de deux locuteurs il est donc difficile d'interpréter ce résultat.

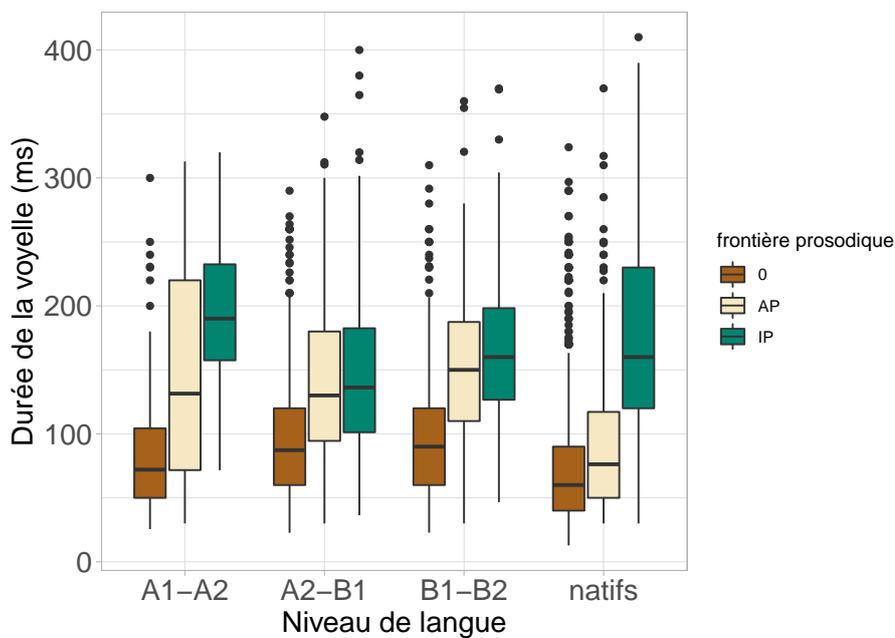


FIGURE 5.26 – Durée des voyelles accentuées selon la frontière prosodique et le niveau de langue (en millisecondes)

Différentes stratégies peuvent être observées chez les locuteurs (Figure 5.27). Les natifs produisent tous des voyelles plus longues en frontière de IP que de AP mais cette distinction est plus marquée chez certains locuteurs (FR2) où l'écart important entre les deux frontières est également associé à la suppression de la distinction voyelle non-accentuée – frontière de AP. Les quelques apprenants montrant un schéma similaire aux natifs (ici JP5

comme exemple) se rapprochent plus de la production de FR1.

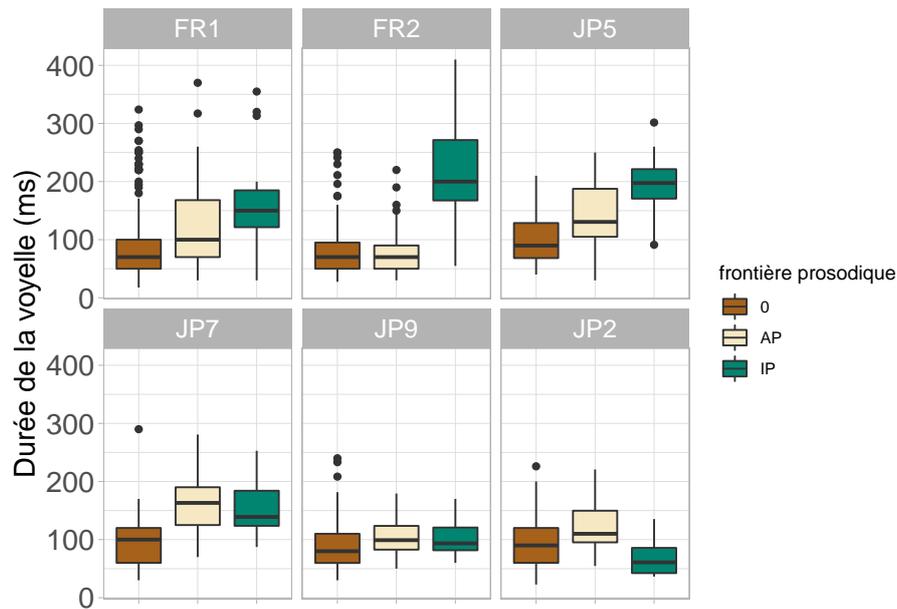


FIGURE 5.27 – Durée des voyelles accentuées de 6 locuteurs aux frontières prosodiques de AP et IP

Mais seulement 4 apprenants observent un tel schéma, le reste d'entre eux ne distinguent pas les frontières de AP de celles de IP et trois autres schémas peuvent être observés :

1. 9 apprenants produisent des voyelles accentuées plus longues que celle non-accentuées mais ne distinguent pas les frontières de AP de celles de IP (exemple : JP7)
2. 1 apprenant produit des voyelles de durées équivalentes qu'elles soient accentuées ou non (JP9)
3. 3 apprenants produisent des voyelles plus courtes en frontière de IP que de AP (exemple : JP2)

La locutrice FR2 est également celle qui produisait un faible pourcentage de contours montants aux frontières de IP parmi les natifs et il est intéressant de noter qu'en contrepartie elle y associe un allongement très important de la voyelle. Ceci est illustré par l'exemple présenté en Figure 5.28. Ce n'est pas le cas de l'apprenant JP9, qui produit également peu de contours montants à cette position prosodique, mais dont la durée des voyelles accentuées

n'est pas allongée. Un exemple est présenté en Figure 5.29 où l'on peut voir qu'aux frontières de IP, l'apprenant produit un allongement vocalique relativement peu important, que ce soit associé à un contour montant (« **lit** ») ou non (« **tableaux** », « **portes** »). Cela montre que si les deux stratégies peuvent sembler similaires (contours non-montants), elles diffèrent dans leur réalisation phonétique. Cette différence dans la gestion des indices prosodiques pourrait s'interpréter par un transfert négatif de la L1 chez l'apprenant japonophone. En effet, bien que nous ayons vu que l'allongement final semble également associé à la continuation de la parole en japonais, nous pensons que cet usage reste moins systématique qu'en français. Un contour descendant en français associé à une voyelle courte sera très certainement interprété comme exprimant la finalité, tandis qu'associé à un allongement final il indique la continuation, notamment lors des énumérations (cf. 2.3 et en particulier la Figure 2.14 où ce contour spécifique est schématisé par Di Cristo en *f*).

Quant au niveau de langue des apprenants, il ne semble pas être lié aux différents schémas observés qui semblent plutôt relever de la variation individuelle.

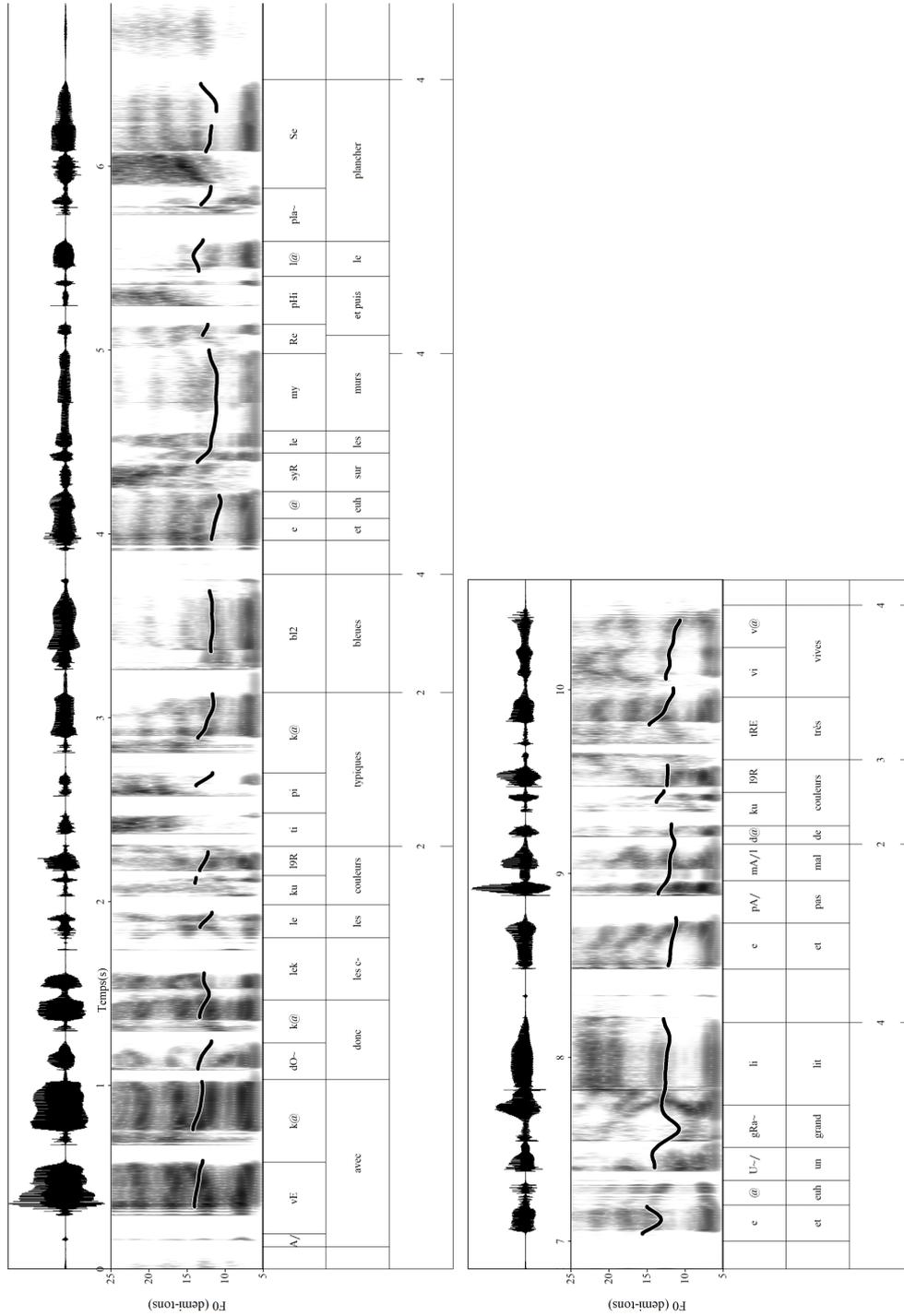


FIGURE 5.28 – Exemple de production du sujet FR2, description de l'image 2. « Avec les couleurs typiques bleues et euh sur les murs et puis le plancher + et euh un grand lit + et pas mal de couleurs vives »



### Allongement vocalique et ratio

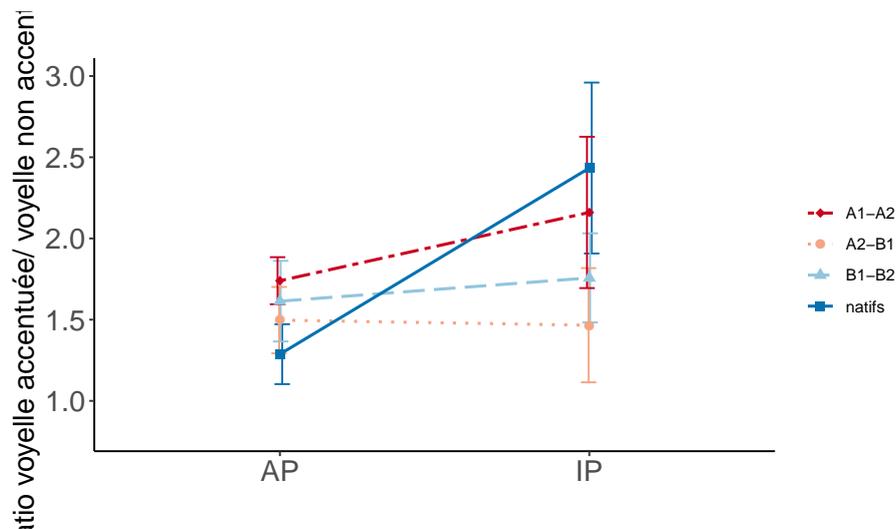


FIGURE 5.30 – Ratio voyelles accentuées/non accentuées selon le niveau de langue

Les analyses présentées en 5.4.3 nous ont montré que les apprenants japonophones ne différenciaient pas les frontières prosodiques par l'indice de durée et ne produisaient donc pas des voyelles plus longues en frontière de IP que de AP. La Figure 5.30 présente le ratio voyelle accentuée / voyelle non-accentuée aux frontières de AP et de IP, et nous pouvons observer que ce ratio tend à être plus élevé chez les natifs en frontière de IP que pour les apprenants, en particulier les groupes A2-B1 et B1-B2. En frontière de AP, il est plus réduit chez les natifs mais cette différence semble moins importante. Encore une fois, nous observons également beaucoup de variation dans les groupes de locuteurs comme nous le suggère l'écart-type important à l'intérieur des groupes. De manière générale, et en prenant pour référence le groupe de francophones natifs et par rapport à la durée des voyelles non accentuées, l'allongement vocalique des apprenants japonophones serait donc légèrement trop important en frontière de AP tandis qu'il serait insuffisant en frontière de IP.

Cette différence se retrouve dans les exemples présentés en Figures 5.32 et 5.31. En effet, nous pouvons y observer des allongements vocaliques plus marqués chez le francophone en frontière de IP que de AP (Figure 5.31), ce que l'on n'observe pas chez l'apprenante (Figure 5.32). Il est aussi intéressant de noter que dans la production de la locutrice japonophone, on observe de courts *chunks* réguliers, caractéristiques du parler des appre-

nants L2, mais aussi des allongements vocaliques liés à des hésitations (« il y a », « [autour de] la »...). Dans cette étude, les productions des apprenants japonophones semblent donc caractérisées par des allongements vocaliques importants à l'intérieur des IP mais des allongements réduits à leurs frontières droites. Bien que nous ayons exclu ces hésitations de l'analyse de la durée, il est intéressant de noter que cela confère ainsi un rythme très particulier au parler des apprenants. En effet, si nous avons parfois aussi pu observer des réalisations similaires chez les francophones, elles sont loin d'être aussi fréquentes que chez les apprenants.

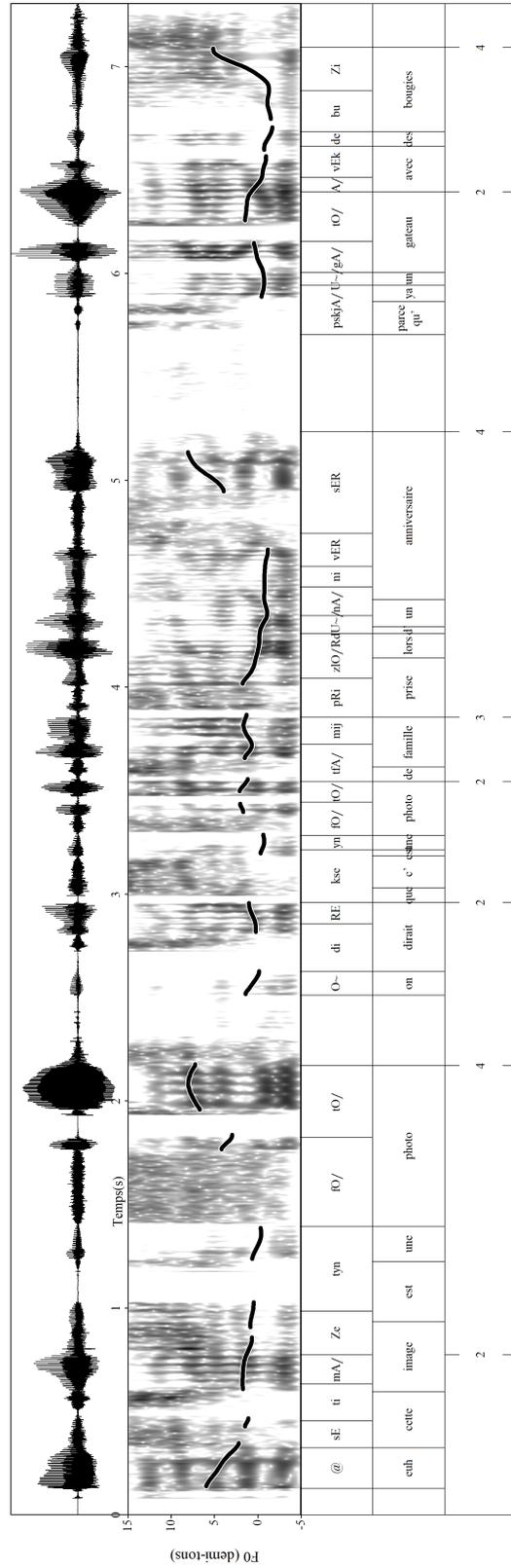


FIGURE 5.31 – Exemple de production d'un francophone (FR4) lors de la description de l'image 4. « Euh cette image est une photo + on dirait que c'est une photo prise lors d'un anniversaire + parce qu'y a un gâteau avec des bougies »



#### 5.4.4 Synthèse

Nous présentons ici les résultats obtenus dans notre étude de description d'images. Nous avons analysé les réalisations des apprenants japonophones et des francophones du groupe contrôle aux frontières de AP et de IP (focales). Comme pour la première expérience de production, nous avons analysé les proportions de contours montants produits, la durée de la voyelle accentuée ainsi que l'écart mélodique entre la voyelle accentuée et celle qui la précède. Les résultats concernant la production de contours montants sont les suivants :

- Les apprenants japonophones produisent, de façon générale, autant de montées mélodiques aux frontières de AP et de IP que les francophones du groupe contrôle. Nous n'observons pas de différence significative entre le groupe contrôle et les groupes d'apprenants.
- Nous avons cependant observé différentes stratégies chez les apprenants et certains d'entre eux produisent peu de montées mélodiques aux frontières de IP. Mais c'est aussi le cas d'une locutrice francophone du groupe contrôle.

Les résultats de l'analyse des données de F0 des contours montants sont les suivants :

- Les apprenants japonophones parviennent à produire, de façon générale, des montées mélodiques plus importantes aux frontières de IP que de AP. Nous n'observons pas de différence entre les montées produites par les francophones du groupe contrôle et les groupes d'apprenants A2-B1 et B1-B2.
- En revanche, la distinction entre les montées en frontière de AP et de IP n'est pas significative pour le groupe de débutants A1-A2, bien que ce résultat ne soit pas généralisable dû au faible nombre d'apprenants débutants (N=2).

Enfin, concernant la gestion de la durée de la voyelle accentuée aux frontières prosodiques, nous avons observé que :

- Les apprenants produisent, de façon générale, des voyelles plus longues aux positions accentuées que non-accentuées.

- Les apprenants produisent des voyelles de durée similaire aux frontières de AP et de IP, tandis que les francophones produisent des voyelles accentuées plus longues aux frontières de IP qu’aux frontières de AP. Trois apprenants tendent même à produire des voyelles plus courtes en frontière de IP que de AP et l’un d’entre eux ne fait pas de différences entre les voyelles accentuées et non-accentuées.
- La différence entre les voyelles en frontière de IP et de AP est significative chez les apprenants du groupe débutant, mais encore une fois, il est difficile d’interpréter ce résultat en raison du faible nombre de données.
- L’observation du ratio voyelle accentuée / voyelle non accentuée nous indique que l’allongement vocalique produit par les apprenants japonophones aux frontières de IP est insuffisant si on le compare aux allongements produits par les francophones. Celui produit en AP, en revanche, est relativement plus marqué que chez les francophones.

Ainsi, notre [première hypothèse](#) concernant un transfert négatif de la L1 sur les productions des japonophones n’est que partiellement confirmée. En effet, alors que nous nous attendions à des taux de montées mélodiques moins importants chez les apprenants aux frontières de IP que chez les francophones en raison de leur caractère facultatif en L1, ce ne fut pas observé. Si certains apprenants ont produit peu de montées mélodiques en frontières de IP, ce fut également le cas d’une locutrice du groupe contrôle. Les apprenants ont également été capables de produire des montées plus importantes en frontière de IP que de AP, et la seule prédiction issue de cette hypothèse que nous avons pu confirmer est la difficulté à produire un allongement vocalique plus important en frontière de IP que de AP. La différence entre la locutrice francophone du groupe contrôle et l’apprenant produisant très peu de montées mélodiques en frontière de IP est d’ailleurs également l’allongement bien plus important en IP que AP chez la francophone native.

Notre [seconde hypothèse](#) concernant la présence de stratégies L2 communes et ne

pouvant être expliquées par le transfert de la L1 n'a pas été confirmée. En effet, contrairement à nos prédictions, nous n'avons pas observé beaucoup de contours extra-montants aux frontières de IP chez les apprenants. Les montées de F0 des apprenants et des natifs du groupe contrôle n'étaient pas significativement différentes. Nous n'avons pas non plus observé de clair phénomène de surgénéralisation ou hypercorrection. Si certains apprenants ont par exemple produit des pourcentages plus élevés de montées mélodiques que les natifs cela reste une minorité.

Notre [troisième hypothèse](#) sur le rôle du niveau de langue sur les résultats n'est pas confirmée. En effet, tous les apprenants semblent capables d'associer les continuations majeures aux frontières de IP à une montée mélodique. Concernant la gestion de la F0 et de la durée, les seules différences que nous avons pu observer sont entre les deux apprenants du groupe débutant et les autres groupes d'apprenants. Outre la difficulté à généraliser ces résultats du fait du nombre d'apprenants dans ce groupe, ceux-ci sont également contradictoires puisque ces apprenants seraient mieux capables de gérer l'indice de durée mais n'arriveraient pas à produire des montées plus importantes en frontière de IP que de AP. Nous ne pouvons donc pas conclure à une amélioration générale de la gestion de ces indices avec le niveau de langue.

## 5.5 Mise en parallèle et discussion

Dans cette première partie du chapitre, nous avons présenté les résultats de deux expériences de production : une tâche de lecture de textes ainsi qu'une tâche de description d'images. Ces deux tâches impliquent un style de parole différent puisque la première permet d'éliciter de la parole lue et la seconde de la parole semi-spontanée. Nous commençons cette section par des observations croisées sur les résultats obtenus pour les deux tâches de production. Néanmoins, nous insistons sur le fait qu'il ne s'agit que d'observations qui ne s'appuient pas sur des tests statistiques. En effet, nous considérons que ces deux tâches de production, bien que comparables sur certains points, sont très différentes l'une de l'autre notamment concernant le type de frontières prosodiques : 5 niveaux ont été identifiés au préalable dans la tâche de lecture de textes et seulement 3 niveaux ont été analysés dans la tâche de description d'images dont le premier « non-accentué » issu d'une opération au-

tomatique excluant les segments notés manuellement accentués, et les deux suivants qui ont été manuellement annotés (AP et IP focal). Bien que nous ayons également noté des frontières de ip dans la tâche de description d'images, nous avons volontairement exclu ces (rares) occurrences de notre analyse. De plus, les groupes contrôles de francophones natifs ne sont pas les mêmes dans les deux tâches de production. Notons que c'est pour cette raison que nous avons essayé, pour les graphiques qui vont suivre, de séparer nettement les résultats des données des apprenants (directement comparables) des données des francophones natifs qui ont comme fonction principale de donner un point de repère dans la visualisation des résultats.

Avant l'analyse des résultats de ces deux études, nous avons formulé différentes prédictions, parmi lesquelles nous anticipions que les apprenants auraient plus de facilité à produire des contours montants en parole semi-spontanée qu'en parole lue, en raison de la présence de contours mélodiques montants et montants-descendants en japonais en parole spontanée (les BPM H% et HL%, cf. 2.3). Ces BPM partagent des fonctions semblables aux continuations majeures montantes du français en permettant le maintien du tour de parole et la continuation discursive. Nous nous attendions donc en particulier à observer ce transfert positif en frontière de IP focales, puisque ces positions prosodiques (à la différence des frontières de AP, ip ou de IP non focales à l'intérieur des énoncés) sont associées à la continuation discursive. Cependant, les résultats obtenus ne correspondent pas à nos prédictions. En effet, les apprenants japonophones, contre toute attente, n'ont pas eu de difficultés à réaliser autant de contours montants que les francophones natifs aux frontières prosodiques fortes (AP et IP dans les deux tâches ainsi que ip dans la tâche de lecture de textes). Ainsi, nous n'avons pas noté de différence globale entre les deux tâches concernant les pourcentages de production de contours montants, comme nous pouvons l'observer dans le graphique ci-dessous (Figure 5.33). Au niveau individuel, nous pouvons cependant observer que certains apprenants produisent des pourcentages de contours montants plus élevés dans la tâche de description d'images que dans la tâche de lecture de textes, notamment les apprenants JP16 et JP18 (voir graphique 5.34), ayant eu plus de difficultés à produire des contours montants dans la tâche de lecture que les autres apprenants. Il est donc probable que pour ces apprenants en particulier un transfert positif ait eu lieu, mais il faut noter que ce n'est pas le cas de tous les apprenants produisant peu de contours

montants en lecture de texte puisque l'apprenant JP9 montre un pourcentage de contours montants aussi bas dans les deux tâches.

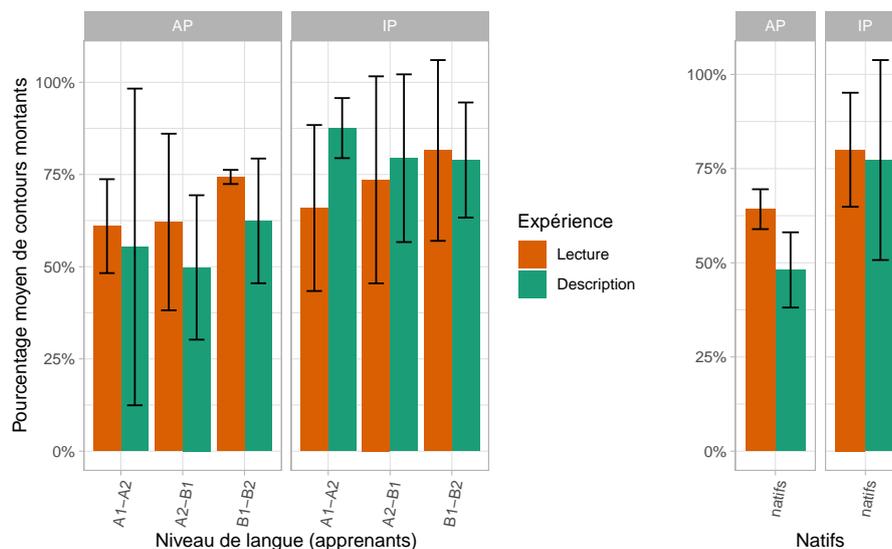


FIGURE 5.33 – Pourcentages moyens de contours montants aux frontières de AP et de IP dans les deux tâches de production

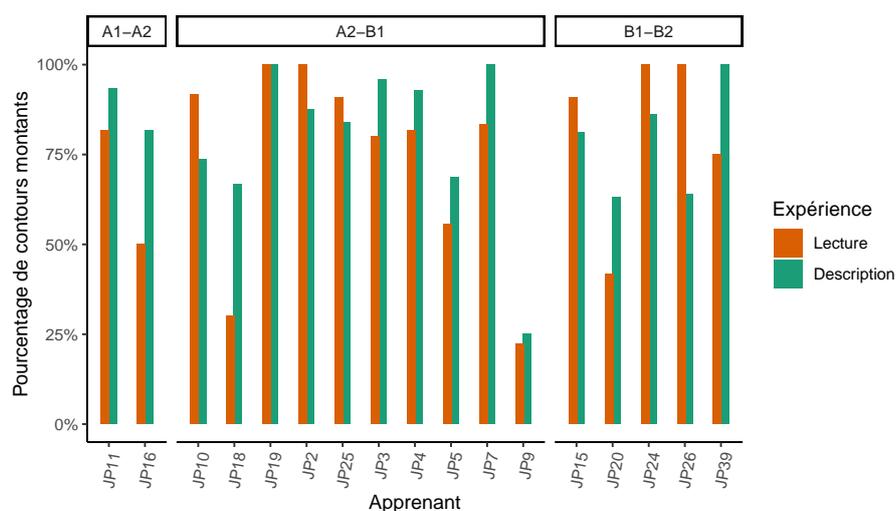


FIGURE 5.34 – Pourcentage de contours montants produits par les apprenants aux frontières de IP dans les deux tâches de production

En ce qui concerne les allongements vocaliques, nous observons la même difficulté à réaliser des allongements plus marqués en frontière de IP chez les apprenants dans les deux tâches. En effet, bien qu'il semblerait que les allongements vocaliques produits soient de façon générale plus importants dans la tâche semi-spontanée que dans la tâche de lecture de textes, le ratio des voyelles accentuées en frontière de AP et de IP par rapport

aux voyelles non accentuées est semblable chez les apprenants (voir Figure 5.35). Les francophones natifs des groupes contrôles, pour les deux expériences, ont produit des allongements comparables aux frontières de AP, tandis qu'on observe des allongements plus importants en frontière de IP chez les francophones de la tâche de description d'images. S'il est possible que cette différence puisse être relative au fait qu'il ne s'agisse pas du même groupe contrôle, nous pouvons cependant facilement l'expliquer par la plus forte portée continuative des continuations majeures en IP dans cette tâche discursive que dans la tâche de lecture de textes. En effet, les forts allongements vocaliques en fin de IP ont pour rôle d'indiquer que l'acte de parole continue – ici la description de l'image. Notons que nous n'avons conservé dans notre analyse que les continuations en position focale (fin d'énoncé) et exclu le dernier contour de la description de chaque image exprimant la finalité et étant par ailleurs souvent réalisé par un contour descendant, pour ne conserver que les contours ayant un réel rôle continuatif discursif. Encore une fois, nous n'avons pas observé le transfert positif au niveau de la durée que nous avions prédit qui se serait caractérisé par des allongements vocaliques plus importants en IP chez les apprenants dans la tâche de description d'images. Cette prédiction était basée sur les résultats de plusieurs travaux en japonais L1 faisant état d'allongements moraiques ou syllabiques finaux lors du maintien du tour de parole (Koiso et al., 1998) ou aux positions discursives faibles (Yoneyama et al., 2003), ayant donc un fort rôle continuatif similaire aux continuations majeures du français en frontière de IP focales. S'il est vrai que l'on observe des ratios vocaliques plus élevés dans la tâche semi-spontanée, c'est également le cas des contours produits aux frontières de AP auxquels on peut difficilement associer un rôle discursif similaire. Nous n'observons donc pas de différence dans la gestion de la durée entre les deux tâches qui pourrait être expliquée par un transfert positif de la L1.

La principale différence que nous avons pu noter lors de l'analyse des résultats se situe au niveau de la gestion de la F0 (voir Figure 5.36). En effet, dans la tâche de lecture de textes, les apprenants n'avaient pas réussi à distinguer les contours mélodiques en fin de IP des contours en fin de AP, tandis que les francophones du groupe contrôle avaient bien produit des montées plus importantes en frontière de IP (continuations majeures) que de AP (continuations mineures). Ce n'est pas le cas dans la tâche de description d'images où tous les groupes d'apprenants (à l'exception du groupe de débutants, composé de seulement

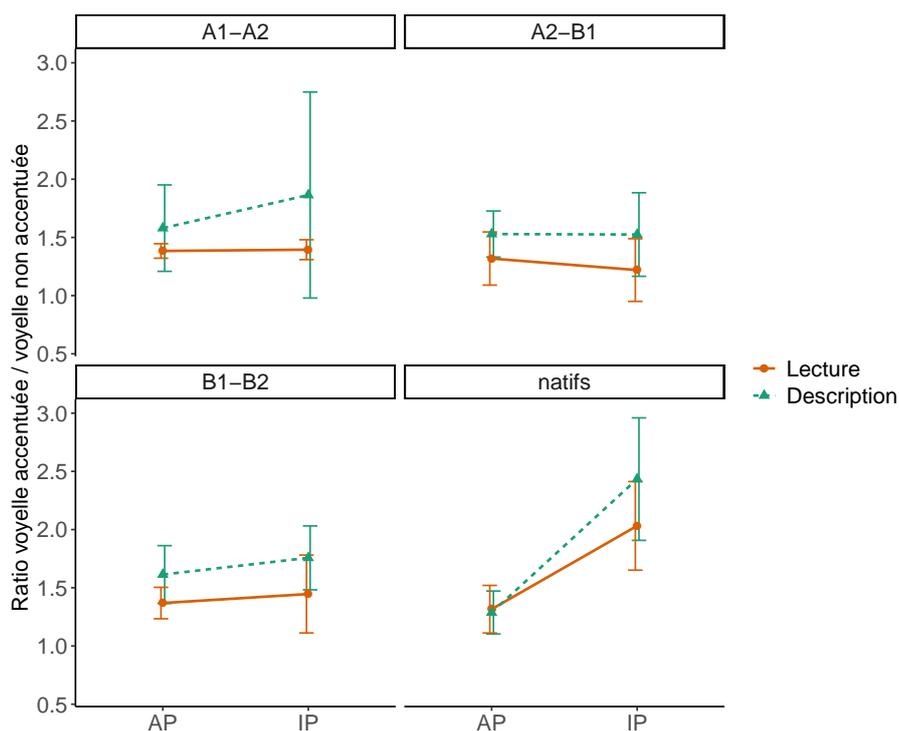


FIGURE 5.35 – Ratio voyelle accentuée en frontière de AP et de IP / voyelle non accentuée dans les deux tâches de production

deux apprenants) distinguent les deux types de frontières prosodiques en produisant des montées plus importantes en frontière de IP. En superposant les résultats des deux tâches, nous pouvons observer que si les apprenants produisent de façon générale des montées mélodiques similaires dans les deux tâches aux frontières de AP (3,1 st en moyenne dans les deux tâches de production), ils associent aux frontières de IP de plus fortes montées mélodiques dans la tâche de description d'images que dans la tâche de lectures de textes (respectivement 4,6 st contre 3,5 st en moyenne), produisant ainsi des continuations majeures semblables à celles des natifs en termes d'écart mélodique (5,1 st en moyenne). Ce résultat pourrait s'expliquer par un transfert positif de la L1 et par la réalisation de BPM montants aux mêmes positions discursives en japonais qui faciliteraient la réalisation de ces contours. Cependant, nous n'avons pas de point de comparaison dans la littérature du japonais nous permettant d'affirmer cette hypothèse. En effet, nous ne pouvons affirmer que les montées mélodiques des BPM sont bien similaires à celles produites par les japonophones en termes de F0. Autre point étonnant, l'absence de contours montants-

descendants qui sont pourtant fréquents en parole spontanée en L1. Est-ce que les apprenants ont réalisé que ce contour était très rare en français et ne l’ont volontairement pas intégré à leur interlangue ? Ou est-ce que la tâche demandée – qui plus est enregistrée – était stressante pour les apprenants qui ont donc produit l’équivalent de BPM H% plus fréquents en contextes formels et tendus tandis que les BPM HL% sont plutôt produits dans des situations informelles et relâchées (Maekawa, 2006) ?

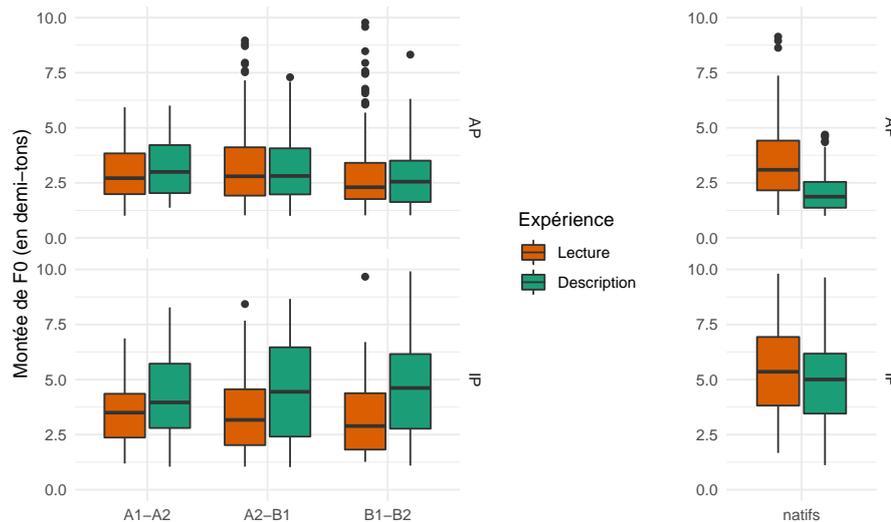


FIGURE 5.36 – Montées de F0 aux frontières de AP et de IP dans les deux tâches de production

Nos résultats nous ont aussi montré qu’il existait des stratégies très différentes selon les apprenants. En effet, un des apprenants en particulier (JP9) a produit une très faible proportion de contours montants et ainsi une prosodie semblable à celle du japonais. Il est intéressant de noter qu’il s’agit d’un des deux seuls apprenants de sexe masculin. Si cette remarque peut paraître anecdotique aux yeux de lecteurs francophones, les différences existantes en japonais entre le parler des femmes et des hommes font qu’elle ne l’est pas au regard de la langue japonaise. Rappelons-nous, par ailleurs, les contours montants descendants du japonais (BPM HL%) qui étaient à l’origine uniquement réalisés par les jeunes femmes avant de se généraliser. Il est probable que les apprenants japonophones produisent moins de contours montants que les apprenantes, et peut-être même que cette différence se retrouve en L1 à l’oral (moins de BPM montants ?). Malgré tout, le peu d’apprenants masculins dans notre étude et le comportement différent du deuxième apprenant

ne nous permettent pas d'aller plus loin dans cette réflexion. Cette réflexion met en avant la principale faiblesse de notre corpus qui est l'absence de production dans la L1 des apprenants. Nous avons en effet choisi de privilégier des enregistrements plus conséquents en L2 que nous avons cherché à analyser à travers le prisme théorique de la prosodie L1 plutôt que d'avoir des enregistrements plus réduits (ou par moins d'apprenants) dans les deux langues.

Il semblerait cependant que les apprenants différencient les frontières prosodiques du français puisque nos observations (non statistiques) montrent qu'ils associent plus fréquemment les frontières de IP à une pause que les frontières de AP, et cela dans les deux tâches. Si pour la tâche de description d'images cela peut s'expliquer par notre méthodologie ne conservant que les frontières de IP coïncidant avec une frontière discursive (focales), ce n'est pas le cas pour la tâche de lecture des textes. Il semblerait donc que les apprenants aient conscience d'une force de frontière plus importante en fin de IP et choisissent ainsi de plus souvent l'associer à des pauses. Notons cependant que la quasi-totalité des frontières de IP des textes du corpus COREIL que nous avons utilisés pour cette expérience ont lieu là où une virgule est présente dans le texte. En effet, les frontières de IP en français coïncident très souvent avec les virgules à l'écrit, non sans hasard. Il est ainsi probable que les apprenants aient conscience de la présence de frontières en français par ce moyen orthographique, qui a peut-être même pu leur être enseigné directement dans les cours de français lors d'exercices impliquant des lectures oralisées (« Il y a une virgule, il faut faire une pause »). En effet, la virgule en français étant même définie comme « marquant une pause de peu de durée »<sup>1</sup> il ne fait pas de doute que cette association entre les virgules et les pauses soit connue de tous les professeurs de français et puisse ainsi leur avoir été enseignée. Néanmoins, si les apprenants semblent bien capables de distinguer ces frontières en y associant plus fréquemment une pause silencieuse, nous avons observé lors nos analyses des deux tâches de productions qu'ils n'arrivaient pas à produire la continuation majeure qui y est attendue. La principale question qui se pose suite à ces résultats est si les apprenants sont capables de distinguer les frontières prosodiques et de les associer aux forces de continuation correspondantes au niveau perceptif, ou si ces difficultés se manifestent uniquement en production. Dans le chapitre suivant, nous présentons deux expériences

---

1. Définition donnée par le dictionnaire *Le Robert*

de perception et de reproduction d'énoncés qui tenteront de répondre à cette question et d'avoir un point de vue plus large sur ce phénomène observé en production.

# Chapitre 6

## Expériences de perception et de reproduction d'énoncés

### 6.1 Introduction

Dans le chapitre précédent nous nous sommes intéressés à la production des contours continuatifs par les apprenants japonophones. Cependant, l'analyse de la production a certaines limites notamment le fait qu'il soit difficile d'avoir une vision claire du système phonologique des apprenants. En effet, les indices phonétiques des réalisations des apprenants peuvent seulement nous donner des indications sur leur traitement de la continuation du français. Par exemple, nous avons observé qu'en lecture de textes, les apprenants ne distinguaient pas les continuations majeures (en frontière de IP) des continuations mineures (en frontière de AP) en produisant une montée mélodique et un allongement vocalique équivalents. Cependant, nous ne pouvons affirmer qu'il s'agit d'une difficulté au niveau phonologique et non simplement phonétique. C'est pourquoi nous considérons qu'il est primordial d'y associer des expériences de perception, mettant en avant la structure phonologique de l'apprenant.

Dans ce chapitre, nous allons présenter deux expériences perceptives : une expérience de reproduction d'énoncés et une expérience de complétion de phrases. La première expérience n'est pas exclusivement perceptive puisqu'elle associe la perception d'énoncés et leur (re)production par les apprenants. Cependant, son objectif est bien celui attendu, soit

l'analyse de la structure phonologique interne.

Les questions auxquelles nous tâchons de répondre dans ce chapitre sont les suivantes :

- Est-ce que les apprenants ont intégré la continuation du français dans leur représentation phonologique ?
- Est-ce qu'ils sont capables de percevoir les différentes frontières prosodiques du français ?

Nos hypothèses sont les suivantes :

**Hypothèse 1** – La perception d'une langue étrangère se fait à travers les connaissances acquises en L1 (transfert).

**Hypothèse 2** – les apprenants sont capables d'apprendre le système prosodique d'une langue cible.

Nos prédictions pour la **première hypothèse** sont que les apprenants vont avoir des difficultés à percevoir et intégrer la continuation du français dans leur système phonologique. Par exemple, ils auront des difficultés à associer certaines frontières syntaxiques à des montées mélodiques ou à percevoir les différentes frontières prosodiques. Cependant, nous considérons qu'il leur est possible d'acquérir ces connaissances phonologiques (**hypothèse 2**) et nous prédisons une amélioration avec le niveau de langue des apprenants.

## 6.2 Protocole

### 6.2.1 Participants

Pour toutes les expériences présentées dans ce chapitre les mêmes participants y ont participé : 38 apprenants japonophones et 12 francophones natifs. Le recrutement des apprenants s'est fait en grande partie à travers les cours de français LV2 de l'Université Sophia, mais aussi des cours de français en tant que spécialité ou par d'autres biais. 7 apprenants ont aussi participé aux expériences de production présentées dans le premier chapitre bien qu'ils restent minoritaires (JP2, JP3, JP4, JP5, JP11, JP18 et JP26). Tous les apprenants étaient des étudiants de l'université Sophia sauf pour deux d'entre eux, aussi étudiants, mais dans une autre université de Tokyo. Nous avons au départ enregistré 54 apprenants, mais afin d'éviter l'interférence d'autres facteurs, nous avons décidé d'écarter

## 6.2. PROTOCOLE

Participant	Age	Sexe	Classe	Durée d'apprentissage	Niveau	Séjour en pays francophone
JP2	21	H	LV2 avancé	3 ans	avancé	0
JP3	23	H	LV2 avancé	2 ans	avancé	0
JP4	21	F	LV2 avancé	4 ans	avancé	2 semaines (France)
JP5	22	F	LV2 avancé	3 ans	avancé	1 mois (France)
JP6	20	F	LV2 débutant	3 ans (lycée)	débutant	0
JP8	19	F	débutant	2 mois	débutant	0
JP11	19	F	LV2 débutant	6 ans (collège et lycée)	intermédiaire	3 semaines (France)
JP13	18	F	LV2 débutant	2 ans	débutant	0
JP14	19	H	débutant	2 mois	débutant	0
JP18	19	F	LV2 débutant	7 mois	débutant	0
JP21	19	H	débutant	2 mois	débutant	0
JP22	19	F	débutant	2 mois	débutant	0
JP23	18	H	débutant	2 mois	débutant	0
JP25	19	F	Spécialiste	3 ans	avancé	0
JP26	21	F	Spécialiste	3 ans	avancé	3 semaines (France)
JP28	20	F	LV2 intermédiaire	1,5 an	intermédiaire	0
JP30	20	F	LV2 intermédiaire	1 an	intermédiaire	0
JP32	19	F	LV2 intermédiaire	1,5 an	intermédiaire	0
JP33	20	F	LV2 intermédiaire	1 an	intermédiaire	0
JP34	20	F	LV2 intermédiaire	3 ans	intermédiaire	3 jours (France)
JP35	23	H	Spécialiste	4,5 ans	avancé	1 semaine (Canada)
JP37	20	F	LV2 intermédiaire	1 an	intermédiaire	0
JP38	21	F	Spécialiste	3 ans	avancé	3 semaines (France)
JP40	20	F	LV2 intermédiaire	1 an	intermédiaire	0
JP41	20	F	LV2 intermédiaire	1,5 an	intermédiaire	0
JP42	21	F	LV2 intermédiaire	2 ans	intermédiaire	1 semaine (France)
JP43	21	F	Spécialiste	2,5 ans	avancé	10 jours (France)
JP44	21	F	Spécialiste	2 ans	avancé	0
JP45	19	F	LV2 intermédiaire	1 an	intermédiaire	0
JP46	20	F	Spécialiste	1 an	intermédiaire	0
JP47	20	F	Spécialiste	1 an	intermédiaire	0
JP48	19	F	LV2 intermédiaire	3 ans	intermédiaire	0
JP49	21	F	Spécialiste	2,5 ans	avancé	2 semaines (France)
JP50	20	F	LV2 intermédiaire	1,5 an	intermédiaire	0
JP51	21	F	Spécialiste	3 ans	avancé	0
JP52	21	H	Spécialiste	3 ans	avancé	0
JP53	20	F	LV2 intermédiaire	2 ans	intermédiaire	3 semaines (Canada, Montreal)
JP54	28	F	LV2 intermédiaire	2 ans	intermédiaire	3 semaines (Suisse)

TABLEAU 6.1 – Données sur les apprenants japonophones ayant participé aux expériences de perception et de reproduction de phrases

les apprenants avec une seconde langue maternelle ayant vécu plusieurs années en dehors du Japon ainsi que les étudiants ayant résidé dans un pays francophone pendant une durée supérieure à un mois.

Quelques informations sur ces participants sont présentées dans le Tableau 6.1 sur leur âge, sexe, classe de français, durée d'apprentissage du français et potentiels séjours en pays francophone. Les francophones du groupe contrôle ont été recrutés dans un cours de linguistique de niveau L1 à l'Université Paris 7 (actuellement Université Paris Cité). Leur participation était basée sur du volontariat et n'a pas compté dans leur évaluation.

## 6.2.2 Conditions du recueil des données

Les apprenants japonophones ont passé les expériences dans une chambre sourde à l'Université Sophia (Tokyo). Les francophones du groupe contrôle ont quant à eux été enregistrés dans la chambre sourde de l'Université Paris 7 (Université Paris Cité). Nous leur avons fait compléter un formulaire de consentement ainsi qu'un questionnaire avant la passation des expériences. Bien que nous présentions uniquement les résultats de deux expériences dans cette thèse, les participants ont participé à quatre expériences différentes. Chaque expérience durait approximativement une dizaine de minutes. L'ensemble du créneau alloué pour ces expériences (incluant les pauses et le remplissage des formulaires) était d'une heure. Nous avons utilisé un ordinateur Asus Zenbook ainsi qu'un casque Sennheiser HD380pro. Pour les expériences de répétition, ils étaient également enregistrés à l'aide d'un micro Shure WH20XLR. Les enregistrements audios sont échantillonnés à 44100 Hz. Les participants japonophones enregistrés au Japon ont été dédommagés de 1000 yen pour leur participation (équivalent à environ 7 euros), et les participants du groupe contrôle enregistrés en France de 10 euros.

## 6.3 Expérience 3 : Reproduction de phrases au contour prosodique modifié

### 6.3.1 Introduction

L'objectif de cette expérience est de savoir si les apprenants japonophones ont intégré ou non la montée continuative du français dans leur interlangue. Pour ce faire, nous souhaitons tester leur capacité à corriger des contours prosodiquement non-acceptables en français, non-montants là où une montée mélodique est attendue. En parallèle, cette expérience permet également de vérifier que pour les francophones cette montée mélodique est bien essentielle et qu'ils vont donc naturellement corriger les contours.

Nos prédictions sont que :

- La montée continuative en français est difficile à acquérir pour les apprenants japonophones puisque dans leur langue maternelle un ton bas (L%) est habituellement produit à ces positions syntaxiques (en lien avec notre [hypothèse 1](#)).

- Les apprenants intègrent peu à peu cette notion de continuation dans leur représentation du français au cours de leur apprentissage (*hypothèse 2*)

Nous nous attendons donc à ce qu'une telle tâche soit globalement difficile pour les apprenants, notamment en comparaison avec les natifs, mais aussi à ce que les apprenants avec un meilleur niveau d'apprentissage corrigent plus fréquemment les contours prosodiques non-montants.

#### 6.3.2 Méthodologie

##### **Matériel linguistique**

Pour construire nos stimuli nous avons créé des phrases suivant une structure syntaxique simple : syntagme nominal (SN) + syntagme verbal (SV). Nous avons choisi des mots simples et ainsi accessibles pour des apprenants, même de niveau débutant.

Ces phrases varient suivant le nombre d'AP contenus dans le syntagme nominal (1 ou 2) ainsi que la structure syntaxique du syntagme (deux noms rattachés par une coordination 'ET' ou un syntagme nominal complexe composé de deux syntagmes prépositionnels). Nous avons également fait varier le nombre de syllabes dans le syntagme nominal. Pour faciliter la modification des stimuli, tous les syntagmes nominaux finissent par une voyelle et sont ensuite suivis par un syntagme verbal commençant par une consonne.

Nous avons enregistré 4 natifs francophones lisant ces phrases (deux hommes et deux femmes) et nous avons ensuite modifié le contour de F0 de ces enregistrements avec Praat (Boersma, 2001) pour créer des stimuli à la fois grammaticaux et agrammaticaux. Chaque contour de F0 de chaque AP a été modifié avec l'éditeur de son Praat pour être soit montant soit plat, donnant deux conditions possibles pour les stimuli avec 1 AP et quatre conditions pour ceux avec 2 AP. Les contours en fin de AP sont donc soit montants (H, high, 'haut') ou plats (L, low, 'bas'). Nous utilisons le terme de *grammaticalité* en raison de la forte congruence entre la syntaxe et la prosodie dans ce type d'énoncés. En effet, certains phrases prosodiques, tels que la présence de frontière de IP à des positions syntaxiques inhabituelles, sont jugés agrammaticaux (Feldhausen & Delais-Roussarie, 2012). Dans notre cas, les stimuli modifiés agrammaticaux ont subi un remplacement de leur contour montant en frontière de SN par un contour plat. D'autres réalisations grammaticales sans contour

montant à la fin du SN seraient possibles en cas d'une focalisation sur le SN (le contour pourrait ainsi être descendant) ou une production eurythmique avec l'accent produit sur le verbe suivant (notamment pour les stimuli avec un SN d'un seul AP). Toutefois, nos stimuli ne correspondent pas à des réalisations grammaticales, bien qu'il soit probable que les participants en produisent en restructurant prosodiquement l'énoncé (focus ou réalisation eurythmique) et sans pour autant produire une montée mélodique en fin de SN.

Conditions **grammaticales** où la frontière de SN coïncide avec une montée mélodique :

- Condition 1 : **H\_H** pour stimuli de 2 AP et **H** pour ceux de 1 AP
- Condition 3 : **L\_H** pour stimuli de 2 AP

Conditions **agrammaticales** où la frontière de SN est réalisée par un contour plat

- Condition 2 : **L\_L** pour stimuli de 2 AP et **L** pour ceux de 1 AP
- Condition 4 : **H\_L** pour stimuli de 2 AP

Tous les énoncés ont été naturellement produits par les francophones avec un contour montant à la fin de chaque AP (schéma prosodique H\_H ou H). Les modifications n'ont été faites que sur la première partie des énoncés (sur le syntagme nominal) et nous avons sélectionné les stimuli modifiés nous paraissant les plus naturels. Une attention particulière a été donnée aux stimuli comportant un seul AP, étant donné que leur courte longueur permettrait ou bien une réalisation congruente à la syntaxe où la montée mélodique coïncide avec la fin du syntagme nominal, ou une accentuation du verbe suivant le syntagme nominal – ici 'mange' ou 'boit' – en particulier pour l'énoncé « Marie mange une banane » où cette réalisation serait eurythmique (cf. 2.2.1). Pour ces deux stimuli, nous avons aligné le F0 du syntagme nominal avec celui du verbe. Tous les stimuli ont été modifiés, même ceux naturellement produits avec des contours montants : dans la condition H\_H nous avons aligné le pic de F0 du premier contour avec celui du deuxième. Un exemple de stimulus modifié dans les quatre conditions est présenté en Figure 6.1.

TABLEAU 6.2 – Stimuli expérience 3

1AP	N	Marie mange une banane
		Magali boit du café
2AP	N+N	Sarah et Marie sont professeurs
		Marie et Magali sont amies
	Long N	Le cours d'anglais s'est bien déroulé
		Les enfants de Rémi sont gentils

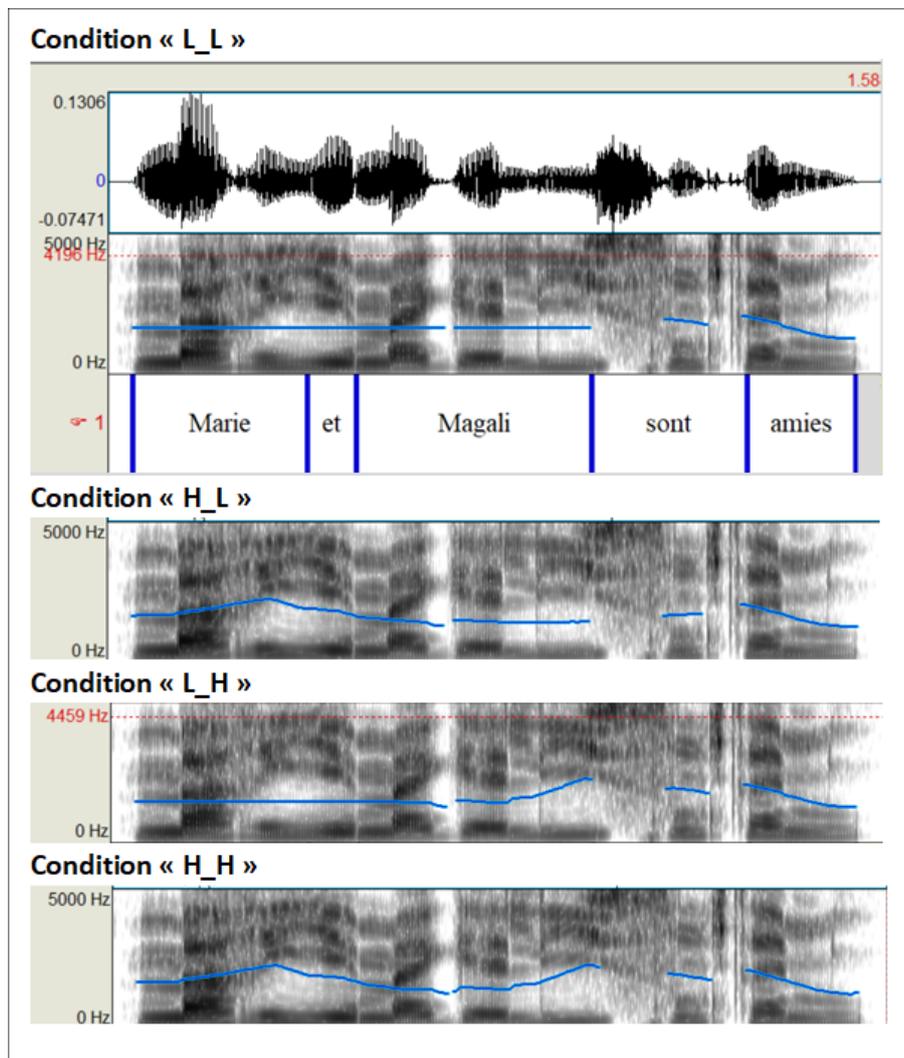


FIGURE 6.1 – Exemple de stimuli modifiés dans les quatre conditions de test. Le schéma réalisé naturellement par la locutrice francophone correspond à la condition H\_H

## Protocole

Pour la réalisation de cette expérience nous avons utilisé le logiciel Psychopy (Peirce, 2007). Afin d'éviter toute interférence entre chaque condition prosodique, nous avons séparé les stimuli de chaque condition dans des blocs, eux-mêmes séparés par une vidéo silencieuse d'une minute. Le but recherché par cette courte vidéo était de distraire les participants et ainsi d'essayer de leur faire oublier l'intonation des stimuli entendus dans le bloc précédent. Nous avons mélangé les stimuli dans chaque bloc avec des distracteurs (les stimuli représentant 1/3 des énoncés et les distracteurs les 2/3 restants). Nous avons également ajouté des blocs de 5 distracteurs après chaque vidéo pour éliminer le risque d'avoir le même énoncé dans deux conditions prosodiques différentes trop rapprochés l'un de l'autre. En effet, nous n'avons pas contrôlé l'ordre d'apparition des stimuli qui était décidé aléatoirement par le programme. L'expérience était également précédée d'un court entraînement, à l'issue duquel les participants pouvaient poser des questions à l'expérimentateur avant de continuer.

Les participants étaient enregistrés dans les conditions présentées en 6.2.2. Les instructions données étaient de reproduire naturellement les énoncés qu'ils allaient entendre. Ainsi, les participants étaient invités à écouter le stimulus puis, après une pause de 3 secondes, à le reproduire de façon naturelle. Ce délai de 3 secondes a pour but d'éviter le plus possible toute imitation directe de l'énoncé et pouvoir ainsi permettre de mieux mettre en évidence leur représentation phonologique.

## Participants

Les participants de cette expérience sont 38 apprenants japonophones et un groupe contrôle de 12 natifs francophones (voir 6.2.1 pour plus de détails). Nous avons écarté un des natifs francophones (FR10) qui n'avait pas respecté les consignes, portant le nombre de natifs à 11.

### 6.3.3 Résultats

Nous avons analysé les contours prosodiques des énoncés reproduits par les participants à la fin de chaque AP avec le logiciel Praat (Boersma, 2001) et noté si ceux-ci étaient

montants (1) ou non (0). Cette mesure a été faite manuellement avec des critères à la fois auditifs et acoustiques : pour noter un contour comme « montant » il faut que celui-ci soit audible et qu'il y ait une différence de plus d'un demi-ton entre la voyelle de la syllabe accentuée et celle de la syllabe précédente. Nous avons noté *NA* les productions incertaines des apprenants et celles que nous avons jugées trop éloignées de la cible, en particulier pour la première partie de l'énoncé, soit le syntagme nominal. Lorsqu'il n'était pas possible de déterminer si un contour était montant ou non, notamment en présence de voix craquée, nous l'avons également noté *NA*.

A partir de ces données nous avons établi les schémas prosodiques produits par les participants (H\_H, L\_H, H\_L ou L\_L pour les stimuli avec un syntagme nominal de deux AP et H ou L pour les stimuli avec un syntagme nominal d'un AP). Nous avons classé les schémas H\_H et L\_H et H, où la montée prosodique correspond à la fin du syntagme nominal, comme prosodiquement « grammaticaux » car ce sont les schémas prosodiques attendus pour ce type d'énoncé. Les schémas prosodiques L\_L, H\_L et L, quant à eux, ont été classés comme « agrammaticaux ».

La variable étant binaire (grammatical ou agrammatical, 1 ou 0) nous avons utilisé un modèle linéaire généralisé à effets mixtes (GLMM) avec le niveau de langue et la condition prosodique comme facteurs fixes. Nous avons également ajouté item et participant comme facteurs aléatoires dans notre modèle.

Selon notre hypothèse nous nous attendions à ce que les apprenants aient plus de difficultés que les natifs à corriger les contours prosodiquement agrammaticaux mais qu'il n'y aurait pas de différence lorsqu'ils sont exposés aux contours grammaticaux. C'est pour cela que nous avons ajouté à notre modèle une interaction entre le niveau de langue et la condition prosodique. Nous avons comparé le modèle avec interaction à un modèle sans interaction avec un test du rapport de vraisemblance, en utilisant la fonction *anova*, et nous avons trouvé une différence significative ( $\chi^2(9)=56,96, p<.001^{***}$ ), de sorte que nous avons gardé le modèle avec interaction pour notre analyse. L'analyse des contrastes de cette interaction a été faite en utilisant le package *emmeans* de R et nous avons opté pour l'ajustement Tukey.

Les résultats globaux de l'expérience dans les quatre conditions prosodiques rassemblant les stimuli 1AP et 2AP sont présentés dans la Figure 6.2. Tout d'abord, nous pouvons

contrast	estimate	SE	df	z.ratio	p.value
natif-avancé	1.42	0.69	Inf	2.05	0.17
natif-intermédiaire	2.14	0.69	Inf	3.13	0.009 **
natif-débutant	2.78	0.83	Inf	3.35	0.0046 **

TABLEAU 6.3 – Contrastes de l'interaction niveau de langue \* condition prosodique, condition 3 (H\_L)

observer que les apprenants de tous les groupes de niveau sont capables de reproduire les contours montants lorsqu'ils y sont exposés. Ainsi, il n'y a pas de différence entre les groupes d'apprenants et les natifs pour les conditions prosodiques grammaticales (L\_H et H\_H).

En revanche, pour les conditions prosodiques agrammaticales (L\_L et H\_L), nous observons un écart important entre les natifs et les apprenants. En effet, les apprenants ont bien plus de difficultés à corriger les contours non-montants et à produire un schéma grammatical. On observe cependant une certaine progression linéaire avec le niveau d'apprentissage bien qu'il n'y ait pas de différence significative entre tous les groupes d'apprenants dans les deux conditions. Cette absence de différence significative s'explique en partie par une variation importante dans tous les groupes d'apprenants que la Figure 6.3 permet de bien visualiser.

Les natifs quant à eux ont significativement plus corrigé les contours prosodiques agrammaticaux que les apprenants, excepté dans la condition 3 (H\_L) où l'on n'observe pas de différence significative entre les natifs et le groupe de niveau avancé. (voir Tableau 6.3)

Il est important de noter que la tâche en elle-même a été relativement difficile pour les apprenants, notamment pour les stimuli 2 AP. En effet, les apprenants avaient beaucoup de difficultés à reproduire les énoncés et nous n'avons pu conserver que 52% des productions pour le groupe de débutants, 57% pour le groupe intermédiaire, mais tout de même 82% pour le groupe de niveau avancé. Pour la condition prosodique la moins naturelle et donc la plus difficile, L\_L ce pourcentage de données analysables descend jusqu'à 35% pour les groupes d'apprenants débutants et intermédiaires.

Les stimuli 1 AP, en revanche, ont été bien plus facilement reproduits par les apprenants du fait de leur courte longueur et de leur simplicité. En effet, nous avons pu conserver plus

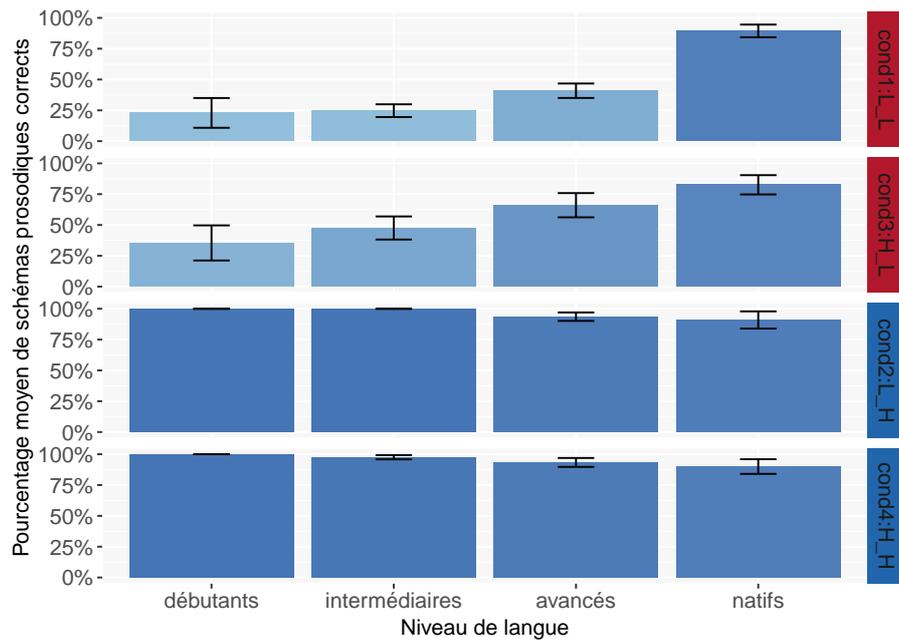


FIGURE 6.2 – Pourcentage moyen de schémas prosodiques grammaticaux par niveau de langue dans les quatre conditions prosodiques (conditions agrammaticales en haut et grammaticales en bas)

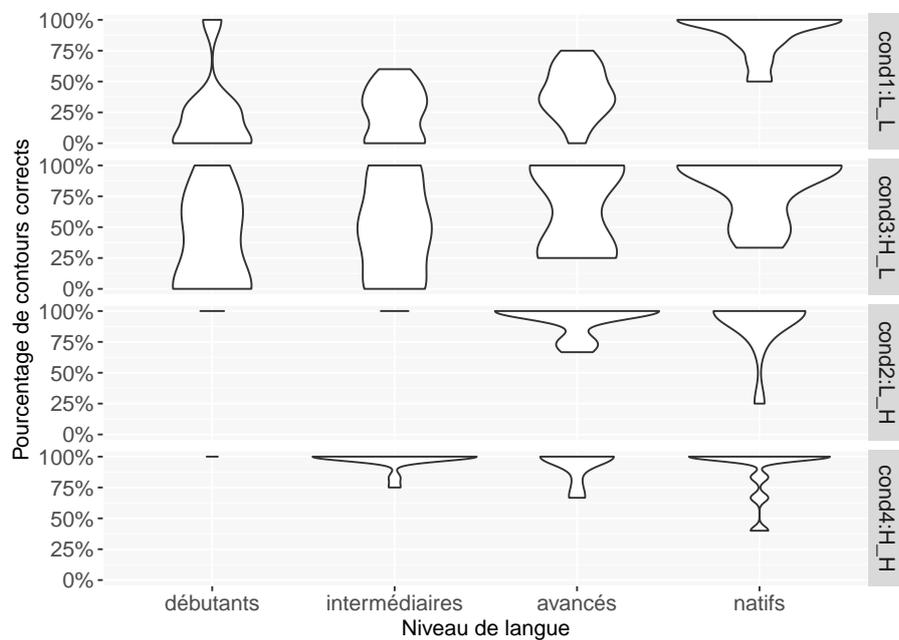


FIGURE 6.3 – Répartition des moyennes de productions de schémas prosodiques grammaticaux par niveau de langue dans les quatre conditions prosodiques (conditions agrammaticales en haut et grammaticales en bas)

de 90% des productions pour tous les groupes de participants. Ce sont ces données que nous allons maintenant présenter.

Nous pouvons voir sur la Figure 6.4 que le pourcentage moyen de contours corrigés tend à progresser avec le niveau d'apprentissage (19% pour les débutants, 29 % pour les apprenants intermédiaires et 46% pour les apprenants de niveau avancé). Mais nous observons toujours une variation très importante à l'intérieur de chaque groupe d'apprenants (voir Figure 6.5).

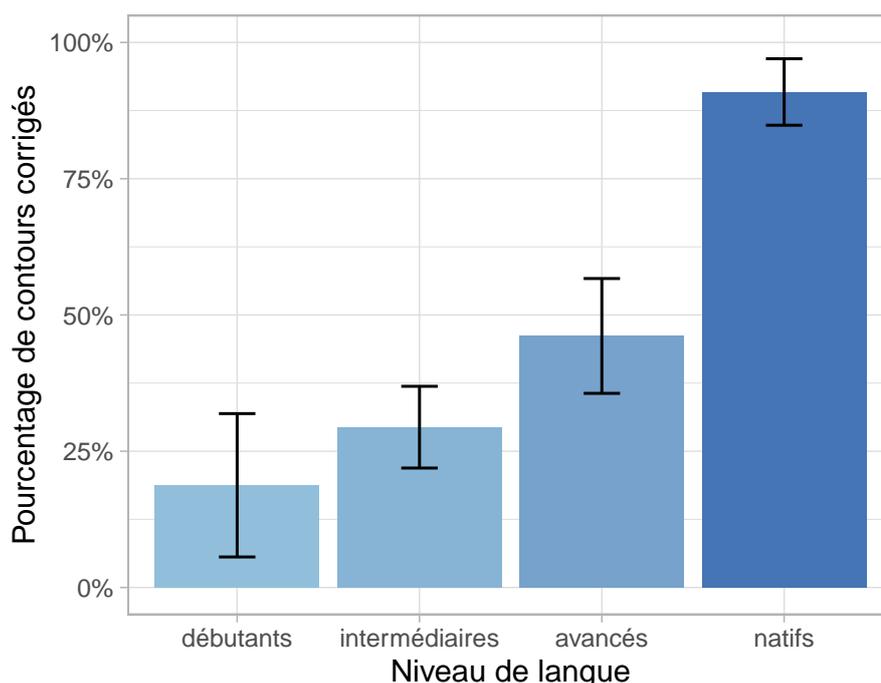


FIGURE 6.4 – Pourcentage moyen de contours corrigés suivant le niveau de langue pour les stimuli 1 AP, condition L

Pour les stimuli 2 AP, les participants ont réalisé plusieurs schémas prosodiques différents (Figure 6.6). Nous pouvons observer que leurs réalisations semblent influencées par la condition prosodique entendue et cela même, dans une moindre mesure, chez les natifs. Par exemple dans la condition 2 (L\_H), les participants ont tendance à produire un même schéma L\_H. De même pour la condition H\_L où l'on observe chez les natifs plus de schémas H\_L que dans les autres conditions. En revanche, l'influence de la condition L\_L, qui est la moins naturelle en français reste très marginale sur les productions du groupe contrôle.

Nous pouvons d'ailleurs noter qu'en dehors du pourcentage plus élevé de schémas

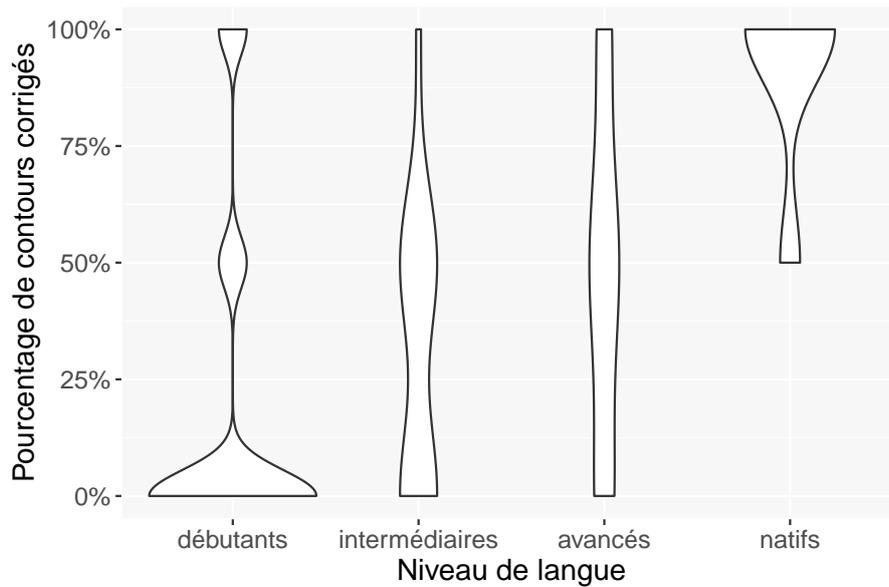


FIGURE 6.5 – Répartition des moyennes de contours corrigés suivant le niveau de langue pour les stimuli 1 AP, condition L

prosodiques grammaticaux produits dans les conditions agrammaticales, les natifs n'ont quasiment pas produit de contours L\_L là où les apprenants en ont produit, même dans la condition H\_L.

#### 6.3.4 Synthèse

Dans cette expérience de reproduction d'énoncés, nous avons cherché à savoir si la montée continuative du français était présente dans l'interlangue des apprenants japonophones. Pour cela, nous avons exposé des apprenants japonophones et des francophones du groupe contrôle à des stimuli modifiés grammaticaux et agrammaticaux ; ces derniers l'étant par la modification du contour montant à droite du SN en contour plat. Nos prédictions étaient que les apprenants auraient des difficultés à corriger les contours non-montants agrammaticaux en raison d'un transfert négatif de leur L1 ([première hypothèse](#)), mais que l'on observerait de meilleurs résultats chez les apprenants de niveau plus avancé car cette notion de continuation peut être acquise par les apprenants lors de leur apprentissage du français ([deuxième hypothèse](#)).

Voici les principaux résultats de la tâche de reproduction d'énoncés :

- Les apprenants japonophones sont capables de reproduire les contours montants

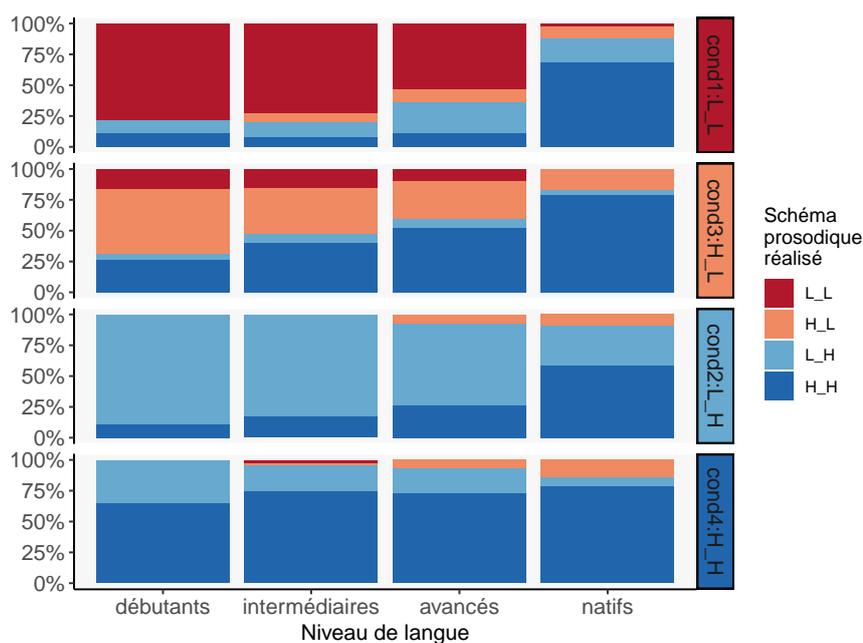


FIGURE 6.6 – Proportions des schémas prosodiques produits suivant le niveau de langue et la condition prosodique

lorsqu'ils y sont exposés

- Les francophones du groupe contrôle ont bien corrigé les contours jugés « agrammaticaux » en produisant une montée mélodique
- Les apprenants japonophones, quant à eux, ont eu des difficultés à corriger ces contours non-montants agrammaticaux.
- On observe une légère amélioration avec le niveau de langue. Dans une des deux conditions (H\_L), nous n'observons pas de différence significative entre le groupe contrôle et le groupe d'apprenants de niveau avancé. De plus, on observe une tendance à plus souvent corriger les contours non-montants dans les stimuli à 1 AP (qui ont également été les plus faciles à reproduire) chez les apprenants de niveau plus avancé.

Ainsi, notre **première hypothèse** est confirmée puisque l'on observe des difficultés chez les apprenants à corriger les contours non-montants agrammaticaux. En revanche, les apprenants sont capables de reproduire les montées mélodiques quand ils y sont exposés et

l'on n'observe donc pas de transfert négatif sur les énoncés grammaticaux.

Notre **seconde hypothèse**, quant à elle, n'est que faiblement confirmée puisque l'on n'observe pas d'amélioration clairement visible avec le niveau de langue. Cependant, on observe une tendance à mieux corriger les contours agrammaticaux, notamment dans la condition H\_L et pour les stimuli ayant un syntagme nominal composé d'un seul AP.

Cette expérience nous a permis d'analyser l'intégration phonologique de la continuation du français chez les apprenants japonophones. Cependant, cette tâche n'est pas exclusivement perceptive car elle demande également une réponse produite. Nous présentons dans la section suivante une expérience cette fois-ci entièrement perceptive ; une tâche de complétion de phrases. Si dans cette première expérience, nous avons pu analyser l'intégration de la continuation montante globale dans l'interlangue des apprenants, cette seconde expérience nous permettra d'explorer leur capacité à percevoir les différents types de continuation et les associer aux niveaux prosodiques du français.

## 6.4 Expérience 4 : Perception des différents niveaux prosodiques : tâche de complétion de phrases

### 6.4.1 Introduction

De précédentes études ont montré que les francophones natifs non seulement produisaient différemment les frontières de syntagmes accentuels des frontières de syntagmes intermédiaires mais qu'ils étaient également capables de les différencier perceptivement (D'Imperio & Michelas, 2009 ; Michelas, 2011). L'expérience utilisée était une tâche de complétion de phrases où les sujets devaient deviner quelle serait la suite la plus probable d'un énoncé tronqué à une frontière de syntagme accentuel ou intermédiaire. Le taux de bonne réponse était statistiquement plus élevé pour les stimuli de syntagme intermédiaire ; Michelas explique cette différence par la tendance universelle à choisir la structure syntaxique la moins complexe. En effet, la bonne réponse des stimuli tronqués en frontière de syntagme intermédiaire correspond également à la structure syntaxique la plus simple.

Nous avons voulu nous inspirer de l'expérience de Michelas (2011) avec des apprenants japonophones, afin de tester leur capacité à percevoir ces frontières prosodiques du français.

Nos prédictions sont les suivantes :

- Les apprenants japonophones auront des difficultés à distinguer les différents niveaux prosodiques du fait du système prosodique différent dans leur L1 (selon notre [première hypothèse](#))
- La tâche sera plus difficile pour les débutants que pour les apprenants d'un meilleur niveau (selon notre [deuxième hypothèse](#))
- Le taux de bonnes réponses pour tous les sujets (natifs et apprenants) sera plus élevé pour les stimuli de syntagme intermédiaire dû au choix « par défaut » de la structure syntaxique la plus simple (selon le principe d'attachement minimal)

## 6.4.2 Méthodologie

### Matériel linguistique

Nous avons repris des phrases de l'expérience de Michelas (2011) et construit de nouvelles que nous avons adaptées aux apprenants en utilisant un vocabulaire accessible. Nous avons obtenu un total de 18 phrases (voir Annexe D.2). La moitié de ces phrases sont construites avec un syntagme nominal comportant deux syntagmes prépositionnels. L'autre moitié des phrases sont construites avec un syntagme nominal composé de trois syntagmes prépositionnels. Dans les deux types d'énoncés, les prépositions utilisées dans les syntagmes prépositionnels sont toujours « de ». Nous avons enregistré quatre locuteurs francophones (deux hommes et deux femmes) lisant ces phrases à voix haute et sélectionné manuellement celles qui correspondaient le mieux à la structure prosodique décrite par Michelas. Ensuite, nous avons tronqué ces énoncés à la fin du deuxième AP correspondant, suivant leur nombre de syntagmes prépositionnels dans leur syntagme nominal, à soit une frontière de AP soit une frontière de ip. Nous avons également enregistré 18 distracteurs et 4 phrases pour l'entraînement. Les distracteurs ont également été tronqués à une position de fin de mot décidée aléatoirement. Contrairement à l'expérience de Michelas, les mêmes énoncés ne sont pas proposés dans les deux conditions (AP et ip), et chaque stimulus est donc différent. Nous avons fait ce choix car la mise en place de deux listes de stimuli différentes à alterner afin que la moitié des apprenants écoutent les stimuli dans une condition

et l'autre dans une autre nous semblait difficilement réalisable étant donné qu'il aurait fallu conserver un équilibre dans chaque groupe de niveau. Or, les apprenants faisaient partie de cursus différents (et à des années différentes, du début de licence jusqu'au master) et l'expérience se déroulait sur les mêmes créneaux pour tous ce qui ne permettait pas une telle organisation, d'autant plus qu'une partie des participants a dû être écartée car le questionnaire qu'ils ont complété nous a révélé qu'ils ne faisaient pas partie de nos critères de sélection (cf. 6.2.1). Cependant, nous pensons que l'utilisation de stimuli différents dans les deux conditions n'est pas si problématique car l'objectif principal est de comparer les résultats des francophones par rapport à ceux des apprenants.

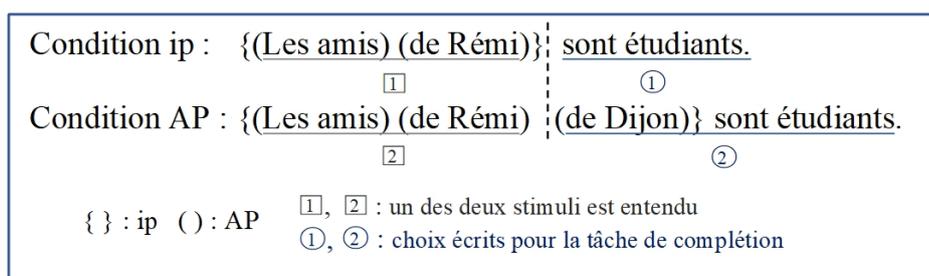


FIGURE 6.7 – Type de stimuli – Tâche de complétion de phrases. Dans les deux conditions les énoncés sont tronqués à la fin du second AP qui correspond soit à une frontière de ip (condition ip) soit de AP (condition AP)

Notons que les énoncés que nous avons sélectionnés diffèrent légèrement dans leur réalisation prosodique. Ainsi, dans certains stimuli de la condition AP la deuxième frontière de AP n'est marquée que par la durée (forme également attestée dans la thèse de Michélas, qui précise que cette frontière de AP reste bien perçue par les auditeurs) et le contour produit dans ces stimuli reste plat ou légèrement descendant au lieu d'être réalisé par une montée mélodique. Nous avons également parfois observé la réalisation d'accents initiaux (Hi) sur la première syllabe du mot lexical du premier AP, et ce dans les deux conditions (ip ou AP) – c'est-à-dire lorsque le SN comporte deux ou trois AP. Nous présentons en

Figure 6.8 un stimulus présentant ces deux particularités. Notre choix de conserver ces réalisations légèrement différentes et non prototypiques se justifie par notre volonté de tester des énoncés tels qu'ils pourraient être entendus par les apprenants en condition naturelle. Cependant, nous sommes conscients de la possibilité que ces différences influent sur les résultats, mais l'intégration du facteur aléatoire « item » à nos analyses statistiques permettra de prendre cela en compte.

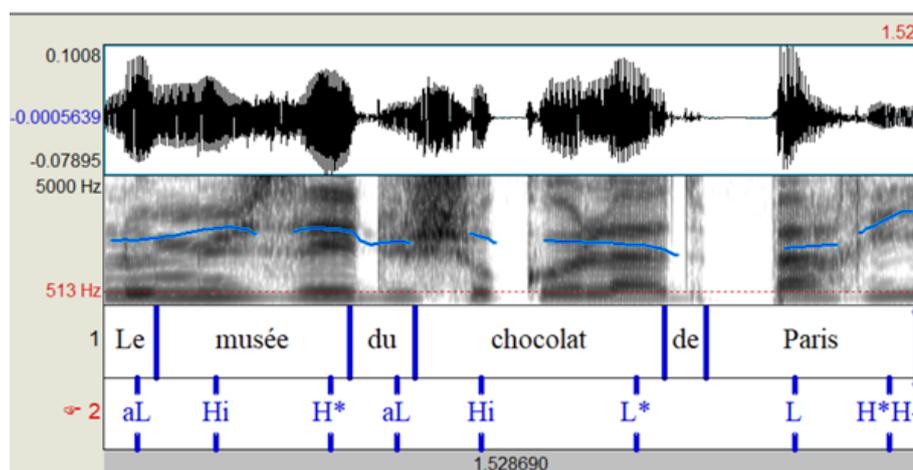


FIGURE 6.8 – Exemple de réalisation « non-prototypique » utilisée comme stimulus pour la condition AP (tronqué à la fin du deuxième AP dans l'expérience). « Le musée du chocolat de Paris [est très intéressant à visiter] ». On peut observer un accent initial sur la deuxième syllabe du premier AP ainsi qu'un ton bas  $L^*$  à la fin du deuxième AP, cependant bien marqué par un allongement final

## Protocole

L'expérience a été créée avec Psychopy (Peirce, 2007). Les sujets ont été enregistrés dans les conditions présentées en 6.2.2.

Nous avons demandé aux participants d'écouter les énoncés tronqués et de choisir parmi les deux propositions écrites la suite qui leur paraissait la plus probable. Nous avons également insisté sur le fait que toutes les phrases sont grammaticales et donc possibles et qu'il s'agit de décider par intuition « suivant la façon dont elles ont été prononcées ». Dans le cas des participants japonophones, les instructions étaient dans leur L1 pour faciliter la compréhension de l'exercice.

Les participants pouvaient réécouter les stimuli autant de fois que nécessaire avant de choisir leur réponse. Cette option n'était pas dans le protocole de Michelas et nous l'avons

ajoutée afin de faciliter la tâche aux apprenants pour qui plusieurs écoutes peuvent s'avérer nécessaires à la compréhension de l'énoncé.

L'expérience commence par un court entraînement et les participants peuvent poser des questions en cas de doute sur les instructions avant de continuer. La durée de cette expérience est d'une quinzaine de minutes.

### Participants

Les participants de cette expérience sont les mêmes que pour l'expérience précédente soit 38 apprenants japonophones et un groupe contrôle de 12 natifs francophones (voir 6.2.1 pour plus de détails).

### 6.4.3 Résultats

Les réponses aux différents items ont été automatiquement enregistrées par le programme Psychopy dans un fichier Excel. La réponse est donc binaire : soit correcte (1) soit fausse (0). Les résultats généraux de la tâche de complétion de phrases sont représentés par le graphique ci-dessous (Figure 6.9). Nous observons que les apprenants de tous les groupes de niveau de langue ont eu beaucoup de difficultés et leurs résultats ne dépassent pas le seuil de chance. A l'inverse, les natifs francophones ont bien réussi la tâche et affichent un taux de réussite global de 76%

Afin de confirmer nos observations, nous avons décidé d'utiliser un modèle statistique GLMM (*Generalised Linear Mixed Model*) permettant le traitement du type de données obtenu (réponse binaire). Nous avons ajouté comme facteurs fixes le niveau d'apprentissage (débutant, intermédiaire, avancé ou natif) ainsi que le type de stimuli (frontière AP ou ip). Nous avons également ajouté item et participant comme facteurs aléatoires dans notre modèle.

Nous nous attendions à observer, en plus d'un pourcentage global de réponses correctes plus élevé pour les natifs et les apprenants de meilleur niveau, une différence selon le type de stimuli (AP ou ip). En effet, il a été montré par Michéas que les natifs obtiennent de meilleurs résultats dans cette tâche pour les stimuli de type ip, correspondant à la structure syntaxique la plus simple. Selon elle, cela est lié à la théorie d'attachement minimal, où les participants à une expérience ont tendance à choisir les réponses les moins complexes

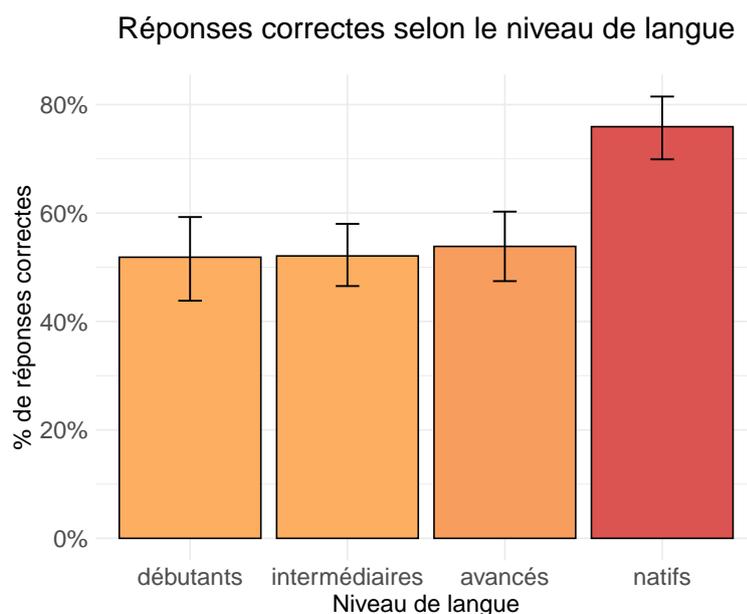


FIGURE 6.9 – Réponses globales correctes selon le niveau de langue

syntactiquement (moins de noeuds syntaxiques). Ainsi, nous avons prédit que les débutants ne seraient pas capables d'utiliser les indices prosodiques afin de deviner la suite des énoncés et donc répondraient plus souvent « par défaut » avec la réponse ip, soit la plus simple au niveau syntaxique. L'écart entre ip et AP serait donc plus important que chez les natifs et les apprenants de niveau plus avancé, où, bien qu'existant, serait réduit par la capacité des participants à utiliser les indices et ne pas répondre « par défaut ».

C'est pourquoi nous avons commencé par construire notre modèle avec une interaction entre le niveau d'apprentissage et le type de frontière prosodique (AP ou ip) comme effets fixes. Le modèle incluait également un intercept aléatoire par participant et une pente aléatoire par participant pour l'effet du type de frontière. Nous avons comparé le modèle avec interaction à un modèle sans interaction avec un test du rapport de vraisemblance, en utilisant la fonction *anova*, et nous n'avons pas trouvé de différence significative ( $\chi^2(3)=2.97$ ,  $p=0.395$ ). Cela signifie qu'il n'y a pas d'interaction entre le niveau de langue et le type de frontière prosodique. Nous avons donc conservé le modèle sans interaction entre type de frontière et niveau d'apprentissage.

Nous pouvons observer dans la Figure 6.10 le taux de réponses correctes selon le type de frontière. La tendance visible est, à l'inverse de nos prédictions, une sensibilité plus accrue

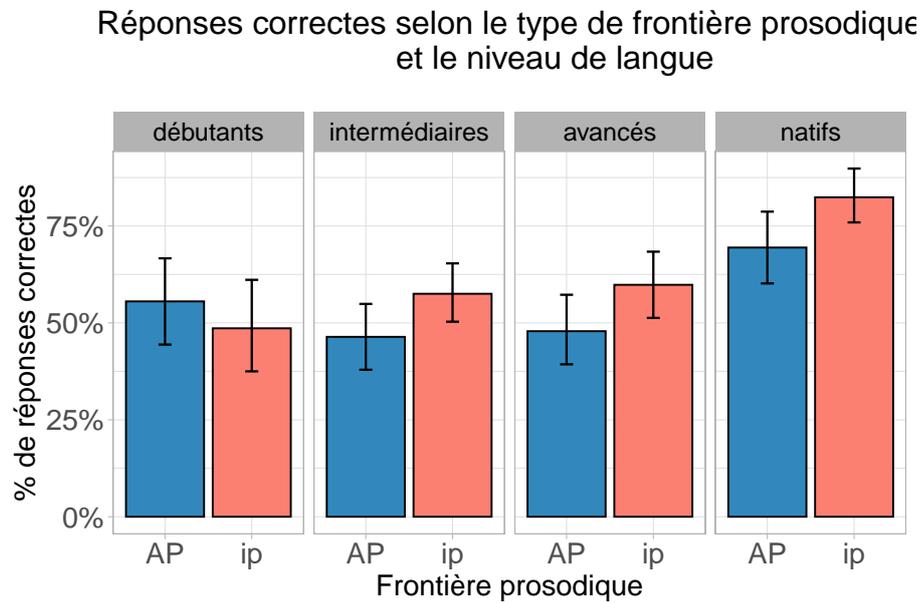


FIGURE 6.10 – Réponses correctes selon le niveau de langue et la frontière prosodique

Fixed effects	estimate	SE	z.value	p.value
(Intercept)	1.07	0.24	4.39	< 0.0001***
boundaryip	0.41	0.27	1.52	0.128
proficiencyadvanced	-1.11	0.23	-4.76	< 0.0001***
proficiencybeginner	-1.23	0.26	-4.78	< 0.0001***
proficiencyintermediate	-1.19	0.22	-5.31	< 0.0001***

TABLEAU 6.4 – Sommaire du modèle GLMM – résultats des effets fixes (le niveau de référence pour le niveau de langue est « natifs »)

aux frontières de ip que de AP suivant l'augmentation du niveau de langue. Cependant, l'absence d'interaction entre le niveau de langue et le type de frontière prosodique ainsi que le résultat non-significatif de l'effet fixe 'type de frontière prosodique' dans notre modèle statistique (voir Tableau 6.4) montrent que cette tendance n'est pas significative.

Lors de la passation des expériences, nous avons recueilli différentes données sur les apprenants à l'aide d'un court questionnaire (Annexe B). Nous leur avons par exemple demandé leur niveau de motivation sur une échelle allant de 0 à 5, ainsi que le nombre d'heures par semaine où ils sont exposés au français, en dehors des heures de cours. N'ayant pu nous appuyer sur de précédentes études pour le traitement de ce type de données, et ne voulant pas rendre notre analyse trop complexe nous avons choisi de ne pas les intégrer dans notre modèle statistique. Il nous semble cependant intéressant d'observer les résultats des apprenants selon ces facteurs, et nous présentons ainsi en Figure 6.11 le pourcentage de réponses correctes selon le niveau de langue et le nombre d'heures par semaine en dehors des heures de cours où les apprenants sont volontairement exposés au français (apprentissage, révisions des cours, visionnage de films, etc.). Nous choisissons de présenter ce facteur car, de notre point de vue, il est à la fois potentiellement lié à la quantité d'input auquel des apprenants sont exposés, mais il est également représentatif de leur motivation de façon probablement moins biaisée que l'échelle de motivation que nous leur avons fait compléter. En effet, beaucoup d'apprenants (10 d'entre eux) ont été capables de répondre qu'ils ne consacraient pas de temps à l'étude du français ou n'étaient pas exposés à la langue en dehors des cours (=0h), reflétant ainsi leur implication minimale dans cet apprentissage, bien que cela puisse aussi varier selon d'autres facteurs tels que la situation de l'apprenant et le nombre d'heures de cours dans la semaine (qu'il peut juger suffisant).

Nous observons que les apprenants les plus exposés au français (plus de 3 heures par semaine) tendent à obtenir de meilleurs résultats que les apprenants du groupe ayant répondu qu'ils ne consacraient pas de temps pour le français en dehors des cours (0 heure), en particulier dans le groupe de débutants et d'apprenants de niveau avancé. Ces simples observations ne peuvent donner lieu à de fortes affirmations, de par la complexité du facteur utilisé ainsi que ses nombreux biais, et l'absence de vérification statistique. Cependant, ces résultats rappellent ceux obtenus par Moyer (2004) (voir section 1.6.2) où celui-ci ob-

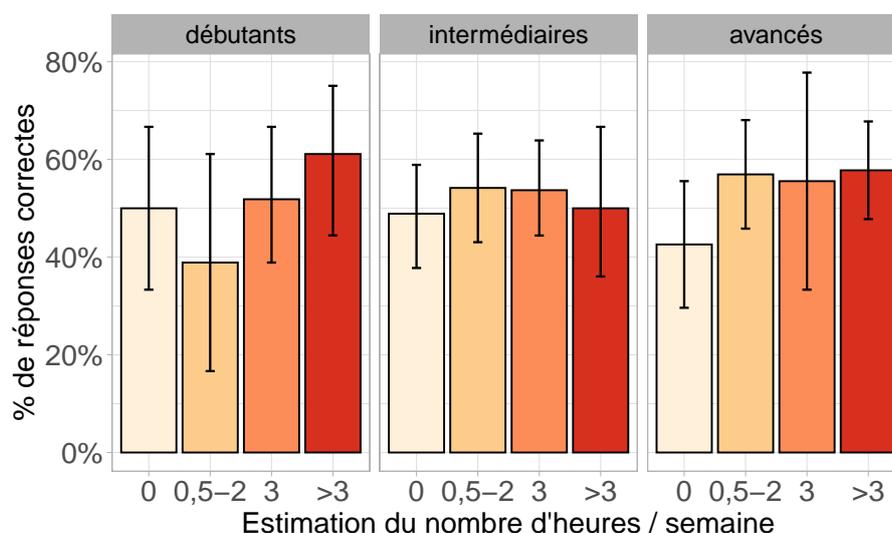


FIGURE 6.11 – Réponses correctes selon le niveau de langue et le nombre approximatif d’heures d’exposition au français / semaine en dehors des heures de cours

serve l’efficacité d’un *engagement optimal* se manifestant par la recherche active d’input et de *feedback* en L2 ainsi qu’une grande quantité d’input passif (lecture, visionnage de films, etc.) sur le degré d’accent étranger perçu. Il est ainsi possible que la motivation et l’exposition au français puissent jouer un rôle positif dans l’acquisition (inconsciente) des différentes frontières prosodiques du français.

#### 6.4.4 Synthèse

Dans cette expérience de complétion de phrases nous nous sommes intéressés à la perception des différentes frontières prosodiques du français par les japonophones. Les participants à cette expérience devaient sélectionner la suite la plus probable de stimuli perçus auditivement et ceci nous permettait de tester leur capacité à associer les différents types de continuation aux positions prosodiques appropriées, et en particulier à distinguer les frontières de AP des frontières de ip dans un syntagme nominal comportant plusieurs AP. Nous prédisions des difficultés à réaliser cette tâche en raison de l’absence d’un tel phénomène dans leur L1 induisant un transfert négatif (*hypothèse 1*), mais de meilleurs résultats chez les apprenants de niveau plus avancé qui auraient en partie intégré le système prosodique du français dans leur interlangue (*hypothèse 2*). De plus, nous avons une

troisième hypothèse spécifique à cette expérience, le principe d'attachement minimal, qui prédisait de meilleurs résultats pour les stimuli ip (la structure syntaxique la plus simple) que ce soit chez les apprenants ou chez les francophones du groupe contrôle.

Voici les principaux résultats de la tâche de complétion de phrases :

- Les apprenants n'ont globalement pas réussi cette tâche et affichent un pourcentage de réussite ne dépassant pas le seuil de chance.
- Les francophones du groupe contrôle ont quant à eux bien réussi l'expérience
- Contrairement à nos prédictions, les apprenants japonophones débutants n'ont pas choisi plus fréquemment la réponse ayant la structure syntaxique la moins complexe (ip).
- On observe une tendance à obtenir de meilleurs résultats avec les stimuli 'ip' chez les francophones du groupe contrôle ainsi que les apprenants intermédiaires et avancés. Cette tendance n'est cependant pas significative.

Ainsi, notre [première hypothèse](#) est confirmée puisque l'on observe des difficultés chez les apprenants à distinguer les frontières prosodiques et à y associer la continuation adaptée. Notre [seconde hypothèse](#), quant à elle, n'a cependant pas été confirmée puisque l'on n'observe pas d'amélioration visible avec le niveau de langue.

## 6.5 Discussion

Dans ce chapitre, nous avons évalué l'intégration de la continuation en français L2 par les apprenants japonophones. Dans un premier temps, nous avons analysé leur capacité à corriger des énoncés modifiés prosodiquement agrammaticaux en français à l'aide d'une tâche de reproduction de phrases. Les résultats que nous avons obtenus nous montrent que les apprenants japonophones ont des difficultés à corriger les contours non-montants mais qu'ils en semblent mieux capables lorsque leur niveau de langue est plus élevé. Il semblerait donc qu'ils arrivent progressivement à acquérir la notion générale de continuation montante et à l'intégrer dans leur interlangue. Cette progression est mieux visible sur les

stimuli les plus simples (avec un SN ne comportant qu'un seul AP), ce qui nous indique également que cette expérience était peut-être trop complexe pour les apprenants de niveau débutant, malgré l'apparente facilité des phrases utilisées. Une autre critique qui pourrait nous être adressée est le fait que les participants auraient tendance à répéter les énoncés et ne pas faire appel à leur représentation phonologique mais seulement à leurs capacités phonétiques. En effet, comme Dupoux et al. (2008) l'ont mis en évidence par les résultats obtenus dans leurs expériences, selon la méthodologie utilisée les compétences évaluées ne sont pas les mêmes. En particulier, une importante charge cognitive est nécessaire pour évaluer les compétences phonologiques des participants. S'il est vrai que nous pensons que les trois secondes de délai entre l'écoute du stimulus et la répétition demandent une charge cognitive suffisante pour accéder à la représentation phonologique de ces énoncés, nous ne pouvons écarter la possibilité que durant ces trois secondes, les participants ne cherchent pas à respecter les consignes de reproduire ces énoncés naturellement et se contentent de les répéter mentalement tels quels en boucle. Cette stratégie pourrait être plus fréquente chez les apprenants débutants pour qui la tâche est plus difficile, et il est ainsi possible que leurs résultats bas puissent s'expliquer en partie par ce biais méthodologique. Dans ce cas, il faut imaginer que le poids du transfert négatif de la L1 serait moins important que les résultats ne le laissent paraître. De plus, les apprenants montrent de très bons résultats pour les conditions grammaticales, ce qui nous permet d'affirmer que le transfert négatif de la L1 reste malgré tout limité.

Les résultats des francophones du groupe contrôle sont très intéressants puisqu'ils semblent naturellement corriger les contours non-montants en contours montants. En effet, les francophones natifs ont privilégié la stratégie de correction du contour non-montant en fin de SN plutôt que de restructurer la prosodie de l'énoncé pour le rendre grammatical, que ce soit en produisant un focus sur le SN ou en faisant une réalisation eurythmique. Même avec les stimuli où le pic mélodique H\* à la fin du premier AP avait été supprimé (condition L\_H, désaccentuation possible en français lorsque le débit de parole est rapide), ils ont choisi majoritairement de produire une montée mélodique à la fin des deux AP. Ce schéma est identique aux stimuli avant leur modification, et c'est le schéma prosodique prototypique attendu. Cependant, le fait d'avoir gardé les indices temporels des phrases originales lors de la création de nos stimuli modifiés peut également expliquer ce

choix ; en effet, comme l'allongement en fin du premier AP est conservé, les francophones peuvent percevoir l'accent et ainsi naturellement l'associer à une montée mélodique, même si celle-ci n'était pas présente dans le stimulus d'origine. Bien que cette expérience de reproduction d'énoncés a été menée en priorité sur des apprenants L2, les résultats montrent qu'elle peut également nous renseigner sur les représentations phonologiques en langue maternelle. Nous pouvons ainsi imaginer que ce type d'expérience puisse être adapté pour une étude en prosodie L1. Il serait par exemple possible de modifier le paramètre de durée en plus des contours mélodiques pour pouvoir observer si les contours sont toujours corrigés en l'absence d'allongement accentuel ou si d'autres stratégies de restructuration prosodique sont mises en place.

Ensuite, la deuxième expérience de ce chapitre (inspirée de Michelas 2011) nous a permis d'évaluer la capacité des apprenants à percevoir les différents niveaux prosodiques du français par une tâche de complétion de phrases. Les résultats des apprenants n'excédant pas le seuil de chance nous suggèrent que si notre première expérience nous a montré que les apprenants avaient en partie intégré la notion de continuation en français, ils ont plus de difficultés à intégrer les différents niveaux prosodiques du français. En effet, ils n'étaient pas capables de percevoir les différences entre les frontières prosodiques de AP et de ip et de les associer à la structure syntaxique appropriée. Les participants du groupe contrôle ont bien mieux réussi l'exercice, confirmant les résultats obtenus par Michelas et la perception de différences prosodiques à ces frontières.

Nos résultats se distinguent cependant de ceux de Michelas sur un point particulier : les francophones de son expérience ont montré significativement de meilleurs résultats avec la condition ip que AP, ce qu'elle explique par le principe d'attachement minimal (les réponses « ip » étant celles comportant la structure syntaxique la plus simple), tandis que nous n'avons observé qu'une tendance non-significative chez nos francophones du groupe contrôle. De plus, si cette tendance est également visible chez les apprenants japonais intermédiaires et avancés, ce n'est pas le cas chez les débutants alors que nous nous attendions au contraire à ce que cette différence soit plus marquée car ils auraient plus facilement tendance à répondre « par défaut ». Ainsi, nous proposons une autre hypothèse pouvant expliquer ces résultats : les restructurations prosodiques possibles selon les stimuli. En effet, le principal inconvénient de ce protocole expérimental est qu'il per-

met de faire percevoir tous les indices prosodiques pertinents... à l'exception des pauses silencieuses. En effet, dans cette expérience, un stimulus est entendu et sa suite (la plus probable, selon l'auditeur) doit être sélectionnée à l'écrit. Ainsi, l'auditeur peut tout à fait imaginer l'insertion d'une pause silencieuse entre le stimulus et la suite de l'énoncé. Or, les pauses contribuent à la perception d'une frontière prosodique forte (Astésano et al., 2012). Il est ainsi possible que les stimuli de la condition AP (coupés en fin du deuxième AP d'un syntagme nominal comportant trois AP), puissent plus facilement être interprétés comme étant une frontière de ip par l'association (virtuelle) à une pause silencieuse, et engendrer un nombre plus important d'erreurs. A l'inverse, il nous semble plus difficile d'imaginer une réalisation prosodique acceptable qui transformerait la frontière de ip des stimuli de la condition ip en frontière de AP. En effet, le fort contour montant généralement réalisé à cette frontière prosodique peut difficilement être perçu par les natifs comme une frontière de AP, et « couperait » l'unité de sens que forme le syntagme nominal. Cela expliquerait la tendance à obtenir de meilleurs résultats pour la condition ip chez les francophones natifs et les apprenants de niveau plus avancé que pour la condition AP. Nous illustrons en Figure 6.12 notre hypothèse explicative : en ① une restructuration prosodique qui rendrait possible la perception d'une frontière de ip à la place de la frontière de AP réalisée dans le stimulus, et en ② une réalisation agrammaticale lorsque la frontière de ip est conservée et « coupe » l'unité de sens. Notons que nous avons choisi de représenter une potentielle réalisation agrammaticale en ② avec un contour montant en fin de SN (sur « de **Dijon** ») car il s'agit de la réalisation la plus fréquente à cette frontière, mais que toute autre réalisation (contour légèrement montant, plat, descendant...) serait aussi difficilement acceptable. Autre remarque, le contour représenté à la fin du deuxième AP est représenté comme plus montant que celui à la fin du premier AP, or ce n'est pas nécessairement le cas bien qu'il s'agisse de la forme la plus souvent observée dans nos stimuli. En effet, la frontière de ip est caractérisée par Michélas par un retour au niveau de référence mélodique, et elle note également un allongement accentuel plus important qu'en AP.

Selon cette hypothèse, le fait que les apprenants japonophones de niveau plus avancé ont tendance à obtenir de meilleurs résultats en condition ip pourrait donc s'expliquer par une meilleure intégration des niveaux prosodiques du français dans leur interlangue, et cela refléterait ainsi une (légère) amélioration par l'apprentissage. Cependant, ceci n'est



FIGURE 6.12 – Illustration de notre hypothèse explicative concernant la tendance à de meilleurs résultats en condition ip dans l'expérience de complétion de phrases. En ① restructuration prosodique AP vers ip avec insertion d'une pause, en ② réalisation agrammaticale illustrant la difficulté à percevoir une autre frontière que de ip

pas confirmé par les analyses statistiques et cela reste donc des suppositions. Les faibles résultats des apprenants montrent que cette expérience a été très difficile pour les apprenants, qui n'étaient principalement exposés à la langue française que dans leurs cours à l'université. La perception des niveaux prosodiques et leur association aux frontières syntaxiques est une tâche subtile qui demande certainement une plus longue exposition à la langue française pour en être capable. Nos observations – non statistiques – concernant les résultats selon le nombre d'heures d'exposition au français semblent par ailleurs aller dans ce sens puisque les apprenants les plus motivés et exposés à plus d'input en dehors de la salle de classe tendent à obtenir de meilleurs résultats que les apprenants les moins motivés. Il serait intéressant de réitérer cette expérience avec des apprenants japonophones vivant en France (ou dans un autre pays francophone) ayant accès à un input plus important et ayant ainsi potentiellement développé de meilleures capacités perceptives. Les résultats de notre étude pilote (qui n'a pas été présentée dans cette thèse), menée à Paris sur quatre apprenants japonophones, semble aller dans cette direction puisque ceux-ci ont globalement bien réussi l'expérience, mais il serait intéressant de vérifier cela à plus grande échelle.

**Troisième partie**

**Bilan et perspectives**



# Chapitre 7

## Discussion générale

Dans la partie précédente, nous avons présenté les résultats de deux expériences de production des contours continuatifs du français (lecture de textes et description d'images) chez des apprenants japonophones ainsi que deux expériences nécessitant l'utilisation de compétences perceptives (tâche de reproduction d'énoncés et tâche de complétion de phrases).

Des recherches antérieures, présentées en 1.5, ont mis en avant les liens existants entre production et perception de la L2. Des études en phonétique L2 montrent l'amélioration des résultats des apprenants à l'aide de paradigmes pré-test/post-test intégrant un entraînement dans la compétence opposée. Ainsi, Akahane-Yamada et al. (1996) et Okuno et Hardison (2016) ont observé qu'un entraînement perceptif permettait aux apprenants d'améliorer leurs résultats en production lors du post-test, et, dans l'autre sens, Kartushina et al. (2015) ont montré l'effet d'un entraînement en production sur la perception. Inversement, d'autres études notent un décalage entre la production et la perception (Detey & Racine, 2015; Sheldon & Strange, 1982; Tateishi & Winters, 2013). Les liens production/perception et la présence d'un potentiel décalage entre les deux compétences nous ont incités à examiner notre problématique du point de vue de la production et la perception, et d'intégrer des expériences faisant appel aux compétences perceptives plutôt que de se concentrer uniquement sur les productions des apprenants.

Les résultats que nous avons obtenus dans les expériences de perception et de reproduction d'énoncés ne sont pas directement comparables à ceux que nous avons obtenus

dans les tâches de production, dans la mesure où les protocoles expérimentaux sont très différents, plus ou moins difficiles pour les apprenants et plus ou moins contrôlés. Cependant, nous avons abouti à des conclusions similaires dans les deux types d'expériences.

Tout d'abord, concernant notre première hypothèse – le transfert de la L1 sur la L2 – celle-ci n'a pu être confirmée que partiellement dans nos études de production, notamment en ce qui concerne le type de contours produits. En effet, alors que nous nous attendions à observer une plus grande proportion de contours non-montants chez les apprenants japonophones que chez les francophones natifs, ce ne fut pas le cas, à l'exception de quelques apprenants. De même, les résultats de notre étude de reproduction d'énoncés montrent que les apprenants parviennent, dans une certaine mesure, à corriger les énoncés agrammaticaux en produisant une montée mélodique. Il est vrai que les résultats de cette expérience sont largement inférieurs à ceux des natifs, alors que dans les tâches de production les apprenants ont produit significativement autant de montées mélodiques que les francophones, mais cela peut en partie s'expliquer par la difficulté de la tâche (ne pas simplement répéter l'énoncé perçu) et le niveau différent des apprenants ayant participé à l'expérience (plus de débutants que chez les participants des tâches de production), d'autant plus que les apprenants de niveau plus avancé montrent de meilleurs résultats. Ainsi, les apprenants japonophones auraient en grande partie intégré les contours montants à leur représentation phonologique du français.

Notre première hypothèse (transfert de la L1) était liée à deux prédictions séparées : la difficulté des apprenants à réaliser les contours montants du français et la difficulté à gérer les indices de F0 et de durée en produisant des montées mélodiques et des allongements plus importants aux continuations majeures que mineures. Nous avons discuté précédemment de notre première prédiction qui n'a pas pu être confirmée, en particulier en production. Cependant, notre deuxième prédiction a bien été confirmée, que ce soit en production ou en perception. Les résultats obtenus dans nos expériences de production nous montrent en effet que si les apprenants japonophones sont bien capables de produire des contours montants aux frontières prosodiques fortes du français (AP, ip et IP), ils ne sont pas capables d'associer les frontières de IP à un allongement final plus important (continuation majeure) et une montée mélodique plus ample qu'en frontière de AP (continuation mineure), à l'exception de la tâche de description d'images où les apprenants ont

---

réussi à produire des montées mélodiques plus importantes en frontière de IP que de AP – et proches de celles réalisées par les francophones natifs en termes de F0. De façon générale, nos résultats montrent que les apprenants ont des difficultés à gérer les indices prosodiques du français et à distinguer différents types de continuations dans leur production. Et nous retrouvons cette même difficulté en perception, où les résultats obtenus dans la tâche de complétion de phrases nous révèlent qu'ils ne sont pas capables de distinguer les réalisations prosodiques aux frontières de ip de celles produites aux frontières de AP, et de les associer aux structures syntaxiques correspondantes. Ainsi, bien que les apprenants semblent intégrer progressivement la notion de continuations montantes dans leur interlangue, ils n'auraient cependant pas intégré les différents types de continuation à leur représentation phonologique du français. Les observations que nous avons faites sur les pauses dans nos analyses des tâches de production laissent à penser qu'il ne s'agit pas d'absence de hiérarchie prosodique dans leur interlangue, mais d'une méconnaissance des rôles des contours continuatifs dans le système prosodique du français. Le lien entre syntaxe et prosodie, pourtant bien attesté en français (cf. 2.2.1), en particulier en lecture oralisée, ne semble pas avoir été entièrement établi par les apprenants.

Au cours de notre discussion en 6.5, nous avons en effet proposé une autre hypothèse que celle de l'attachement minimal qui permettrait d'expliquer la tendance des francophones natifs à obtenir de meilleures réponses pour les stimuli ip que AP dans la tâche de complétion de phrases. Plus précisément, nous notions que s'il était possible pour des francophones d'interpréter une frontière de AP comme frontière de ip en « ajoutant » perceptivement une pause, une frontière prosodique forte à la fin du deuxième AP facilite l'interprétation d'une frontière ip car, autrement, elle briserait l'unité de sens formée par le syntagme nominal. Si nous suivons cette réflexion, cela nous indique que si les apprenants ne réussissent pas l'exercice c'est qu'ils ne sont donc pas capables de percevoir les frontières fortes produites au milieu d'un groupe de sens comme non-natives et gênant la compréhension et l'intégrité du groupe de sens (ici un syntagme nominal). Encore une fois, ce résultat perceptif est à mettre en parallèle avec les productions des apprenants. Bien qu'il ne s'agisse pas d'une stratégie systématique, certains apprenants, en particulier à la fin d'énoncés plus complexes, ont produit des frontières fortes à la fin de courts *chunks* (pause et fortes montées mélodiques), cassant ainsi les unités de sens. Il est inté-

ressant de noter qu'un même phénomène est mis en évidence par Xi et al. (2020) dans son étude sur les apprenants mandarinophones. S'il est possible d'expliquer ces productions par une stratégie universelle découlant d'une certaine insécurité linguistique (semblable à celle décrite par Santiago 2014, bien que dans notre cas nous n'observons pas de contours extra-montants), les résultats bas obtenus dans la tâche de complétion de phrases nous laissent penser que cela n'est pas la seule explication, et que les apprenants ne perçoivent pas ces frontières trop importantes à l'intérieur des groupes de sens comme inadéquates.

Il est intéressant de faire le parallèle avec les études neurolinguistiques, qui ont mis en évidence les ondes cérébrales jouant un rôle dans la perception des frontières prosodiques, et qui notent de plus grandes difficultés chez les locuteurs à supprimer une frontière prosodique superflue que de traiter une frontière manquante (Pauker et al., 2011). Dans sa thèse, Pélissier (2018) s'est intéressée à la capacité des apprenants francophones à associer frontières syntaxiques et prosodiques en recueillant des données EEG (électroencéphalogramme) lors de la perception d'énoncés à prosodie congruente ou incongruente, tels que « As the man is hunting the ducks # the pigeons fly away » (prosodie congruente) et « As the man is hunting # the ducks # the pigeons fly away » (prosodie incongruente en raison d'une frontière prosodique superflue). Elle a montré que tous les apprenants percevaient bien les frontières prosodiques car une onde spécifiquement active à ces frontières, *Closure Positive Shift* (CPS), pouvait être observée. De plus, tout comme les natifs, les apprenants avaient moins de difficultés à (perceptivement) rétablir une frontière manquante qu'à supprimer une frontière superflue. Cependant, certains apprenants ne montraient pas de P600 suite aux frontières superflues, c'est-à-dire un pic de l'activité électrique cérébrale généralement activé par la perception d'erreurs grammaticales ou lors de réanalyse syntaxique, mais qui peut également être déclenché par une prosodie incongruente. Les apprenants plus compétents et les natifs, eux, montraient bien un P600 lors de la perception de stimuli incongruents. Nous pouvons faire le parallèle avec les apprenants japonais et proposer l'hypothèse que, comme ces apprenants francophones de l'anglais, ils soient moins sensibles au lien entre la prosodie et la syntaxe que les natifs et qu'ils n'ont pas la même réponse cérébrale qu'eux face à des énoncés prosodiquement incongruents. Pélissier a également testé l'effet d'un entraînement sur les liens prosodie-syntaxe, qui a montré une amélioration de la capacité à discriminer certains stimuli et le déclenchement

---

d'un P600 lors des stimuli incongruents chez les apprenants les moins performants dans le post-test. Il semblerait donc bien que les apprenants puissent être capables d'être sensibilisés à la prosodie et ses liens avec la syntaxe, montrant l'importance que peuvent avoir les entraînements spécifiques sur la prosodie des apprenants. C'est ce point que nous allons approfondir dans la section suivante, où nous allons nous intéresser à l'enseignement de la Prosodie L2 et des contours continuatifs en particulier, pour enfin proposer des **applications didactiques** concrètes exploitant les résultats et conclusions obtenus à l'issue de notre travail de recherche.



# Chapitre 8

## Proposition d'application didactique

Dans ce chapitre, nous nous intéressons aux implications didactiques des résultats que nous avons obtenus dans notre recherche. Nous allons dans un premier temps faire une courte revue des différentes méthodes pour l'enseignement de la prosodie L2, éprouvées par la recherche expérimentale, afin d'avoir une vision sur les possibilités didactiques qui s'offrent à nous. Ensuite, nous présentons la place actuelle de l'enseignement des contours continuatifs en FLE et les (rares) exercices proposés dans des méthodes de langue. Enfin, nous tentons de combler ces lacunes en proposant nos propres idées d'activités didactiques pour l'enseignement de la continuation du français.

### 8.1 Méthodes d'entraînement en Prosodie L2

#### 8.1.1 Les tâches de *shadowing* et de répétition

Les tâches de *shadowing* ou de répétition sont fréquemment utilisées dans l'enseignement des langues étrangères. En quoi consistent-elles ? Lorsqu'on les oppose, la tâche de *shadowing* décrit la répétition quasi-instantanée et simultanée d'un stimulus auditif (décrite comme répétition *on-line* par Shiki et al. 2010) tandis la simple répétition (« *off-line* ») correspond à la répétition de groupes de mots ou de phrases en décalé, après une pause silencieuse. Selon Hamada 2019, la tâche de *shadowing* demande aux apprenants plus d'attention aux aspects phonologiques du stimuli entendu que la tâche de répétition. Ainsi, plusieurs travaux de recherche ont été menés dans le but de tester l'efficacité du *shado-*

*wing* dans l'apprentissage de la prosodie L2. La technique de *shadowing* ayant été principalement développée au Japon pour les apprenants japonophones de l'anglais depuis plus d'une vingtaine d'années (Hamada, 2019), la grande majorité de ces travaux portent sur ce public ou sur les apprenants du japonais.

Kusumoto (2015) a voulu tester l'efficacité du *shadowing* sur la prosodie des apprenants japonophones de l'anglais. Les résultats montrent que certains apprenants sont jugés par des anglophones natifs comme ayant progressé entre le pré-test et le post-test (13 apprenants sur 30). Cependant, dans une deuxième expérience où l'auteure a analysé les productions de la courbe mélodique des apprenants par rapport aux modèles, elle n'observe pour la majorité des apprenants qu'une amélioration modérée voire nulle pour certains apprenants. Les effets positifs du *shadowing* sur l'apprentissage de la prosodie semblent donc limités.

Rongna et Hayashi (2012) se sont intéressés à l'efficacité du *shadowing* sur la prosodie des apprenants du japonais. Ils ont enregistré 15 apprenants du japonais (11 L1 chinois et 4 L1 mongol) partagés en deux groupes de niveau avant et après les sessions d'entraînement. Suite à trois sessions réparties sur trois jours de *shadowing* et de lecture d'un dialogue, les apprenants ont amélioré leur production du *pitch* japonais et augmenté leur débit de parole. Cet effet prolongé de l'entraînement n'était pas présent dans une précédente étude des deux auteurs, où ils avaient observé une forte régression après l'entraînement et donc une durée limitée dans le temps de l'impact positif de la technique de *shadowing* sur la production de l'accent japonais. Mori (2011) a quant à elle observé une amélioration de la prosodie (rythme, intonation et allongement final) de 20 apprenants japonophones de l'anglais suite à des sessions d'entraînement alternant *shadowing* et tâches de lecture à voix haute.

Bien que le *shadowing* soit la méthode d'imitation la plus testée pour l'apprentissage de la prosodie L2, certaines études se sont intéressées à la technique de répétition. C'est le cas notamment des travaux de Escudero et al. (2018) sur les apprenants de l'espagnol. Ils ont entraîné 5 apprenants japonophones et 9 anglophones américains par des lectures oralisées de phrases suivies de répétitions. Ils ont observé un effet positif de la tâche de répétition par rapport aux lectures oralisées, mais uniquement pour la fluence. Yoshimura et MacWhinney (2007) se sont aussi intéressés à la fluence des apprenants en entraînant

30 apprenants du japonais sur des phrases. Leur protocole intégrait l'écoute de la phrase lue par un modèle et lectures à voix haute de la même phrase apparaissant sous forme écrite à l'écran six fois, avant d'être à nouveau répétée sans le support écrit. Dans leur étude, ils notent une amélioration de la fluence au cours au fil des six lectures à voix haute, indiquant que l'écoute suivie du stimulus précédant la première répétition ne suffit pas à atteindre une bonne fluence et que les lectures à voix haute permettent une meilleure amélioration. Cependant, les deux tâches (répétition et lecture à voix haute) n'étant pas séparées cela ne permet pas leur comparaison, et d'évaluer l'ampleur de leur impact sur la fluence. Les travaux de Pellegrino et Vigliano (2015) ont une approche originale en s'intéressant à l'efficacité de la répétition d'énoncés non pas issus d'un modèle natif comme c'est habituellement le cas, mais originellement produits par les apprenants avant d'être synthétisés pour correspondre à une réalisation native. Les auteurs justifient cette procédure par l'impact positif de la proximité de la voix du modèle sur la réussite de l'entraînement en prononciation L2 (citant Probst et al., 2002). Ainsi, les apprenants peuvent répéter les stimuli énoncés avec la voix leur étant la plus proche (la leur !) et cette « auto-imitation » permettrait donc une amélioration de leur prosodie en L2. Leur étude porte sur 7 apprenants japonophones de l'italien L2, sur trois conditions variant prosodiquement en italien : l'ordre, la demande et la concession. L'entraînement a permis aux apprenants d'améliorer leurs productions dans les trois conditions de sorte à ce que les italophones natifs soient mieux capables de reconnaître les fonctions pragmatiques des énoncés produits, l'effet de l'entraînement étant particulièrement important pour la condition de concession (+ 38,66% d'identifications correctes). L'exercice n'a cependant pas permis aux apprenants d'améliorer leur degré d'accent étranger perçu. Enfin, il est dommage que les auteurs n'aient pas cherché à comparer la technique d'auto-imitation à la classique répétition d'énoncés lus par un modèle natif, pour vérifier que la première méthode donne bien de meilleurs résultats que la seconde.

En conclusion, les méthodes de *shadowing* ou de répétition semblent généralement aider les apprenants à améliorer leur prosodie, mais cet effet est variable et ne persiste pas nécessairement dans la durée. Il paraît donc nécessaire de les associer à d'autres méthodes d'entraînement.

### 8.1.2 La visualisation

Percevoir la prosodie d'une langue étrangère n'est pas une tâche aisée pour les apprenants. En cours de langue, les explications de l'enseignant peuvent leur paraître confuses, voire abstraites. En effet, il n'est pas rare qu'un apprenant ne parvienne pas à produire les contours mélodiques appropriés (par exemple, il produit des contours descendants à la place des contours montants), et que non seulement il ne parvienne pas à se corriger malgré les explications et l'exemple donné par l'enseignant, mais aussi qu'il ne perçoive pas les contours qu'il produit tels qu'ils le sont et qui lui paraissent mêmes adéquats. Dans ce cas, les techniques de visualisation de la courbe mélodique peuvent être un réel atout. Ainsi, des études se sont intéressées aux techniques de *feedback* visuel pour l'apprentissage de la prosodie L2.

De Bot (1983) a cherché à évaluer l'efficacité du *feedback* audio-visuel par rapport à une simple tâche de répétition sur la prosodie de l'anglais L2 chez des apprenants néerlandophones, en se basant sur une étude antérieure n'ayant pu observer de bénéfice de l'utilisation d'un *feedback* audio-visuel sur la prosodie L2<sup>1</sup>. 59 apprenants ont été répartis en quatre groupes : deux groupes s'entraînant avec un *feedback* audio-visuel ou par des répétitions, où ils ont la possibilité de répéter chaque stimulus jusqu'à ce qu'ils soient satisfaits de leur production, un groupe n'ayant eu que les instructions initiales sur la prosodie de l'anglais (aussi présentes dans les deux groupes de test) et un groupe n'ayant uniquement participé qu'aux pré-test et post-test. Les tests de jugement montrent que ce *feedback* permet une meilleure amélioration de la prosodie des apprenants que la tâche de répétition. Les auteurs notent un autre point important : les apprenants répètent plus souvent les stimuli lorsqu'ils ont un *feedback* visuel que lorsqu'ils n'en ont pas, certainement car cela les motive à tenter de corriger leurs erreurs. Le groupe de contrôle n'ayant eu accès qu'aux instructions concernant la prosodie de l'anglais montre également une amélioration, bien que moins importante que celle du groupe test de *feedback* audio-visuel.

James (1976), dans une étude menée sur des étudiants de l'université de Toronto, a cherché à évaluer l'efficacité de la visualisation. Trois groupes ont été constitués : un groupe

1. Wichern, P. U. M., & Boves, L. (1980). Visual feedback of F0 curves as an aid in learning intonation-contours. In *Proceedings Institute of Phonetics Nijmegen* (Vol. 4, pp. 53-63), cité par De Bot 1983

suivant la méthode traditionnelle de répétition de phrases, un deuxième pouvant visualiser la courbe mélodique des stimuli et un troisième ayant également accès à un *feedback* visuel de leurs propres réalisations. James observe que seuls les apprenants du dernier groupe expérimental montrent une amélioration de leurs productions. Le deuxième groupe n'ayant pas obtenu de meilleurs résultats que le premier, la simple visualisation de la courbe mélodique native ne semble donc pas suffisante.

Anderson-Hsieh (1992) a testé l'utilisation d'un logiciel de visualisation (Visi-Pitch) en cours d'anglais pour l'enseignement de la prosodie des apprenants dont la L1 était majoritairement le chinois ou le coréen. Elle observe que grâce à ce *feedback* visuel, la plupart des apprenants parviennent au bout de quelques essais à corriger, ou du moins améliorer, leurs productions en anglais. Il faut cependant noter que ce *feedback* est associé à des explications de la part de l'instructeur. L'auteure remarque que les apprenants semblent plus à l'aise et moins gênés vis-à-vis de leur prononciation non-native car fascinés par la présentation visuelle de leurs réalisations.

Plus récemment, Sandryhaila (2019) a montré l'efficacité d'un entraînement audio-visuel sur des apprenants sinophones du français par l'utilisation du logiciel WinPitch (Germain & Martin, 2000), les apprenants du groupe expérimental ayant significativement mieux amélioré leur production des contours mélodiques que ceux du groupe contrôle n'ayant effectué qu'une tâche de répétition sans *feedback* visuel. Dans une seconde étude sur des apprenants anglophones américains – toujours en français L2 – elle observe que l'association d'explications et de commentaires de la professeure permettent une meilleure amélioration de la prosodie des apprenants que l'utilisation du logiciel en totale autonomie, montrant à nouveau l'intérêt des instructions prosodiques pour observer des progrès substantiels.

Guyot-Talbot et al. (2016) ont analysé l'effet d'un entraînement avec *feedback* visuel sur la production des contours mélodiques des questions ouvertes et fermées en anglais chez des apprenants francophones, ayant tendance à incorrectement produire des montées mélodiques sur les questions ouvertes. Les auteurs ont mesuré la corrélation entre les contours prosodiques des apprenants et ceux du modèle natif et ont observé que les apprenants avaient progressé dans leur production des questions ouvertes et fermées, tandis qu'aucune amélioration n'était visible dans le groupe contrôle n'ayant pas bénéficié de

l'entraînement.

Le *feedback* visuel semble donc bien être efficace pour l'apprentissage de la prosodie L2. Il permet en effet un retour immédiat sur les productions des apprenants et les aide à cibler et corriger leurs erreurs. De plus, les apprenants semblent apprécier ces méthodes qui les motivent à se corriger. Le point négatif de ces techniques de *feedback* visuel est qu'elles nécessitent, pour être facilement mises en place et relativement efficaces, de se dérouler dans un laboratoire de langues. Or, si la plupart des universités possèdent un ou des laboratoires de langue et que certains cours de langue peuvent s'y dérouler, ce n'est pas nécessairement le cas des plus petites structures comme les écoles de langue, ou les classes d'enseignement secondaire. Et s'il est toujours possible pour l'enseignant de proposer des activités de visualisation à tour de rôle sur un seul ordinateur, ce n'est pas la solution la plus adaptée et cela limite grandement les possibilités.

Il existe des techniques tentant d'intégrer la visualisation prosodique dans les supports écrits, pour aider l'apprenant à percevoir et comprendre la prosodie de la langue qu'il étudie. C'est le cas notamment des « écritures prosodiques » (« *prosodic writing* ») qui essaient de reproduire certains traits prosodiques en jouant sur la taille et la hauteur des lettres ou les polices de caractère.

Dans une expérience sur des apprenants japonophones de l'allemand, Rude (2012) a testé l'utilisation d'une écriture prosodique pour enseigner l'accentuation de l'allemand. Ainsi, 15 apprenants ont lu un texte en allemand avant de recevoir un entraînement alliant explications prosodiques et tâches d'imitation d'exemples audio retranscrits sur support écrit avec deux méthodes : soulignement des syllabes accentuées et écriture prosodique. Dans le post-test, les apprenants devaient à nouveau lire le texte, cette fois-ci entièrement retranscrit en écriture prosodique. L'auteur note une forte amélioration de l'accentuation chez les apprenants dans la lecture du texte en écriture prosodique par rapport au pré-test, en prenant pour référence le mot « *zentral* » accentué en deuxième syllabe et correctement réalisé par plus de deux tiers des apprenants dans le post-test contre seulement un seul apprenant dans le pré-test. Rude note également la réduction des voyelles non-accentuées et un ratio similaire à celui des natifs, donnant lieu à un rythme plus proche de celui de la langue cible. Cependant, il nous paraît important de nuancer les résultats obtenus par cette étude car ils ne portent que sur un ensemble de données relativement limité et que l'amé-

lioration de la production par l'écriture prosodique n'a pas été par la suite testée sur des supports écrits classiques. Les bons résultats obtenus, bien qu'encourageants et permettant certainement aux apprenants de mieux comprendre l'accentuation de l'allemand et de produire des réalisations proches de celles des natifs, ne reflètent peut-être que la capacité des apprenants à bien prononcer l'accentuation de l'allemand en suivant l'écriture prosodique sans pour autant être capables de reproduire ces résultats dans d'autres contextes.

Kamiyama et Yamamoto (2007) ont eux aussi mis en place un entraînement prosodique basé sur l'utilisation d'écritures prosodiques. Ils se sont intéressés à l'apprentissage des expressions de demandes en japonais par les apprenants francophones. En effet, les apprenants francophones auraient tendance à produire des montées finales trop importantes sur la particule interrogative, qui rendent l'énoncé « impoli ». Après un entraînement en classe associant explications et répétitions avec l'aide de représentations visuelles par des écritures prosodiques, les auteurs rapportent que les apprenants produisaient moins de demandes dont la prosodie était jugée « impolie ». Ils notent également le retour positif des apprenants qui semblent avoir apprécié cette méthodologie, qui les a aidés à mieux prendre conscience de la prosodie du japonais.

Niebuhr et al. (2017) ont cherché à comparer six techniques de visualisation prosodique : un tracé continu de la mélodie de l'énoncé au-dessus du texte (*continuous contour*), un tracé simplifié de la mélodie de l'énoncé représentant uniquement les principaux changements mélodiques (*hat pattern*), un tracé ajoutant au contour mélodique des points sur les syllabes plus ou moins gros selon l'intensité (*tadpole*), une technique de visualisation où le texte suit la hauteur mélodique – une écriture prosodique (*meandering text*) – ainsi que deux systèmes de visualisation habituellement utilisés par les prosodistes, GToBI (Baumann et al., 2000) et GAT2. Ils ont recruté 18 apprenants néerlandophones de l'allemand et leur ont proposé de lire par paire deux énoncés par technique de visualisation, et de s'enregistrer lisant les énoncés à voix haute après avoir discuté sur comment les énoncés devraient être réalisés selon eux. Ensuite, ils devaient classer les techniques de visualisation selon leur facilité. Parmi les six méthodes proposées, ce fut la technique de tracé *hat pattern* qui a donné les meilleurs résultats et qui a été jugée la plus facile à utiliser. Le tracé continu (*continuous contour*) et la technique *tadpole* arrivent en seconde place mais seule la première technique a bien été classée par les participants (en troisième place). L'écriture

prosodique (*meandering text*) a été classée en seconde place pour sa facilité mais a donné de moins bons résultats que les autres techniques évoquées précédemment. Les deux techniques d'annotation prosodique GToBI et GAT2 sont sans surprise celles qui ont été les moins efficaces et les plus difficiles à utiliser pour les participants, confirmant qu'elles ne sont pas adaptées à une utilisation en enseignement L2. Ainsi, selon ces résultats, les techniques de visualisation directe de la courbe mélodique semblent les plus efficaces pour la production de prosodie L2, et une courbe épurée ne marquant que les principaux changements mélodiques semble être la représentation la plus simple et la plus efficace pour les apprenants.

Bien qu'aucune étude, à notre connaissance, n'ait testé l'efficacité des écritures prosodiques et autres techniques de visualisation sur les productions des apprenants L2 à plus grande échelle, il semblerait qu'elles puissent au minimum aider et motiver les apprenants à percevoir la prosodie L2, et potentiellement leur permettre de progresser. Il faut toutefois rappeler l'étude de James (1976) ayant conclu à l'absence de progression par la simple visualisation, tandis que le *feedback* visuel apparaissait comme bien plus efficace. Nous pensons donc que ces techniques de visualisation ne peuvent avoir un impact aussi important que le *feedback* visuel mais que, associées à des explications de la part de l'enseignant, elles peuvent être un outil utile dans l'enseignement de la prosodie.

### 8.1.3 Les gestes

Ces dernières années, un certain nombre d'études se sont intéressées à la gestuelle en cours de langue et à l'utilisation des gestes pour l'enseignement de la prononciation. Associer des gestes aux explications est une stratégie spontanée et naturelle et qu'utilisent fréquemment les enseignants. Lorsqu'ils tentent d'illustrer des phénomènes prosodiques, ces gestes permettent de proposer une visualisation des différents paramètres. Ainsi, certains gestes représentent la hauteur mélodique et les mouvements de F0 par des mouvements de la main (« *pitch gestures* »), par exemple de droite à gauche et de bas vers le haut pour représenter un contour montant. D'autres représentent le paramètre de durée (« *durational gestures* »), par exemple en mettant les mains parallèles et rapprochées l'une de l'autre puis en écartant la main gauche pour représenter un allongement syllabique.

Les premières études sur les gestes représentant la durée – les *durational gestures* –

montrent des résultats peu concluants. Hirata et Kelly (2010) ont testé l'efficacité des gestes dans l'apprentissage du contraste vocalique du japonais chez 60 locuteurs naïfs, monolingues anglophones dans quatre conditions d'entraînement : audio, audio-visuel, audio associé à des gestes marquant la durée de la voyelle et audio-visuel associé à des gestes. Seul l'entraînement audio-visuel a permis une meilleure amélioration que l'entraînement audio (contrôle) lors des tâches d'identification, montrant l'absence d'efficacité particulière des gestes sur l'apprentissage de ce contraste. Hirata et al. (2014) réitérèrent le même type d'expérience en testant également un entraînement où les participants produisent les gestes, et en comparant ces résultats à celui où ils ne font que les visualiser. Les résultats sont à nouveau décevants car ils n'observent pas de différence significative entre les différents types d'entraînement, les gestes n'étant pas associés à de meilleurs résultats, qu'ils soient produits par les participants ou simplement visualisés. Cependant, Li et al. (2020) considèrent que les résultats peu encourageants décrits par ces études sont liés à la méthodologie utilisée par les auteurs et notamment le type de gestes : gestes de battement (« *beat gestures* ») pour les voyelles courtes et mouvement horizontal des mains pour les voyelles longues. Ainsi, ils choisissent dans leur étude d'utiliser des mouvements horizontaux d'ampleur différente selon la longueur de la voyelle. Li et al. (2020) ont mené leur expérience sur 50 locuteurs catalanophones naïfs et se sont intéressés à leur perception du contraste vocalique japonais par une tâche d'identification, mais aussi à leur production via une tâche d'imitation dans le pré-test et le post-test et les participants se sont entraînés en reproduisant les paires minimales et les gestes leur étant associés. Cette fois-ci, l'entraînement avec les gestes a été positif et a permis une meilleure amélioration des productions lors des tâches d'imitation que le groupe contrôle. En revanche, les résultats des tests d'identification montrent que cet effet positif ne se retrouve pas en perception car aucune différence statistique n'a été observée entre le groupe expérimental et le groupe contrôle concernant l'amélioration des résultats.

Gluhareva et Prieto (2017) se sont intéressées aux gestes de battement (« *beat gestures* ») dans l'acquisition de la prosodie de l'anglais. Leur étude menée sur 20 apprenants catalanophones de l'anglais montre que la visualisation de mouvements de battement sur les syllabes accentuées améliore les productions des apprenants dans une tâche de production guidée semi-spontanée (phrases élicitées par un contexte donné), notamment pour

les contextes difficiles, par rapport à un groupe contrôle n'ayant pas bénéficié de la visualisation de gestes durant leur entraînement.

Quelques études ont cherché à évaluer l'efficacité des *pitch gestures* dans l'acquisition de la prosodie L2. Yuan et al. (2019) se sont intéressés à l'efficacité des *pitch gestures* sur la production d'énoncés de l'espagnol (déclaratives, questions totales, requêtes) de 89 apprenants mandarinophones. Le groupe expérimental a reçu un entraînement associant des énoncés à des *pitch gestures* correspondant à des mouvements de la main imitant les trois contours mélodiques finaux (L\*L%, L\*H% et L\*HL%). Leurs résultats montrent que les deux groupes d'apprenants ont mieux utilisé les contours mélodiques appropriés lors du post-test pour les questions totales et les requêtes, mais que l'amélioration des résultats est plus importante chez les apprenants ayant bénéficié de l'entraînement avec les gestes et ayant de bonnes aptitudes musicales. L'utilisation des gestes semble cependant ne pas avoir eu d'impact positif sur les apprenants ayant de faibles aptitudes musicales.

Baills et al. (2019) se sont penchés sur le potentiel intérêt d'imiter les gestes contre simplement les visualiser lors de l'apprentissage. Dans une première étude menée sur 49 locuteurs naïfs catalanophones et portant sur l'apprentissage de mots en mandarin et des tons lexicaux correspondants, ils confirment l'intérêt de la visualisation de *pitch gestures*, qui permet ici aux participants d'obtenir de meilleurs résultats pour l'identification des tons mais aussi pour apprendre le vocabulaire du mandarin. Dans une deuxième expérience sur 56 autres locuteurs catalanophones, semblable à la première et ne différant que par les instructions données, ils montrent que l'imitation des gestes lors de la phase d'entraînement sur les mots du mandarin permet aux participants encore d'obtenir de meilleurs résultats que lorsqu'ils se contentent de répéter les mots et de visualiser les gestes associés sans les reproduire.

Ainsi, les gestes semblent aider les apprenants à mieux percevoir et produire la prosodie L2, mais comme pour les autres méthodes (visualisation, *shadowing*, répétition...) les résultats des études expérimentales montrent un certain nombre de limites. Ces méthodes semblent être efficaces sur certains faits prosodiques en particulier, certains apprenants (par exemple selon leurs aptitudes musicales), certains types de gestes, etc. Il paraît donc pertinent d'avoir recours à la multimodalité en combinant plusieurs méthodes d'entraînement dans les cours de L2 afin de bénéficier de leurs potentiels apports, tout en suscitant

au maximum l'intérêt des apprenants pour la prosodie par la grande variété des exercices proposés.

## 8.2 La place actuelle de l'enseignement des continuatifs en FLE

Actuellement, très peu de place est accordée à la prononciation en FLE, comme le déplorent Sauvage et Billières (2019). Ainsi, la plupart des manuels généralistes se contenteront de proposer quelques exercices de prononciation ou encadrés explicatifs sur des points de phonétique en particulier. Pour avoir des ressources plus complètes il sera donc nécessaire de se tourner vers des manuels spécialisés dans l'enseignement de la prononciation. Cependant, même parmi ces manuels spécialisés on observe malheureusement peu d'exercices axés sur la prosodie, notamment concernant les continuatifs. Dans cette section, nous proposons d'analyser la place de l'enseignement des continuatifs dans des méthodes de FLE généralistes visant un public d'adultes ou adolescents, relativement récentes (à partir des années 2000) ainsi que des manuels spécialisés pour l'enseignement de la prononciation. Nous avons analysé les onze méthodes de FLE suivantes : *Alter Ego+*, *A Plus!*, *Défi*, *Edito*, *Génération*, *Merci!*, *Nouveau Rond-Point*, *Tendances*, *Totem*, *Vite et Bien*, *Entre Nous*. Quant aux manuels de phonétique, nous avons analysé les quatre collections suivantes : *Phonétique progressive du français*, *Phonétique essentielle*, *Phonétique en dialogues*, *La prononciation en classe* (voir tableau 8.1 pour plus de détails).

Titre	Année	Maison d'édition	Niveau(x)			
Alter Ego +	2012	Hachette	1 : A1	2 : A2	3 : B1	4 : B2
A Plus!	2015/2016	Editions Maison des Langues	1 : A1	2 : A2.1	3 : A2.2	4 : A2.2
Défi	2018-2020	Editions Maison des Langues	1 : A1	2 : A2	3 : B1	4 : B2
Edito	2016/2018	Didier	1 : A1	2 : A2	3 : B1	4 : B2
Génération	2016	Didier	1 : A1	2 : A2	3 : B1	4 : B2
Merci!	2016	Didier	1 : A1	2 : A2	3 : B1	
Nouveau Rond-Point	2011/2013	Editions Maison des Langues	1 : A1/A2	2 : B1	3 : B2	
Tendances	2016	CLE International	1 : A1	2 : A2		
Totem	2014/2015	Hachette	1 : A1	2 : A2	3 : B1	
Vite et Bien	2009/2010	CLE International	1 : A1/A2	2 : B1		
Entre Nous	2015/2016	Editions Maison des Langues	1 : A1	2 : A2	3 : B1	
Phonétique progressive du français	2001-2006	CLE International	débutant	intermédiaire	avancé	
Phonétique en dialogues	2006	CLE International	débutant			
La prononciation en classe	2014	PUG	A1 à B2			
Phonétique essentielle du français	2016/2017	Didier	A1/A2	B1/B2		

TABLEAU 8.1 – Méthodes de FLE analysées

Dans trois des méthodes de FLE (*Nouveau Rond-Point*, *A Plus!*, *Vite et Bien*) nous n'avons

pu trouver aucun exercice sur la prosodie du français. Pour les autres méthodes, les exercices et explications prosodiques sont généralement placés dans un encadré (souvent intitulé « Phonétique ») et ne sont pas directement intégrés à la leçon. Sans surprise, le nombre d'exercices phonétiques présents dans ces manuels est plutôt réduit, en particulier en comparaison avec les exercices grammaticaux ou lexicaux. La majorité d'entre eux se focalisent sur des points segmentaux, mais il est intéressant de voir que beaucoup de méthodes préfèrent commencer par des exercices axés sur le suprasegmental avant d'aborder le segmental. C'est le cas notamment de *Totem* dont les premières leçons sont associées à des exercices sur le rythme, l'accentuation et l'intonation du français, tandis que les phonèmes sont travaillés dans les leçons suivantes. La méthode *Défi* privilégie également la prosodie dans son premier opus où la majorité des exercices proposés sont axés sur l'étude des phénomènes suprasegmentaux, bien que l'opus suivant (A2) propose principalement des exercices segmentaux et que ces derniers soient eux-mêmes absents des niveaux supérieurs (B1 et B2).

Dans toutes les méthodes que nous avons analysées, la prosodie du français est abordée d'une manière similaire : les exercices proposés traitent les paramètres de la durée et d'intonation séparément. Dans un premier temps, la définition des « groupes rythmiques » ou « mots phonétiques » correspondant généralement aux syntagmes accentuels est le plus souvent présentée avec des indications sur l'allongement produit à leur fin : « l'accentuation porte sur la dernière syllabe du mot phonétique. La dernière syllabe est toujours plus longue » (*Totem 1*, p.21), la dernière syllabe du groupe rythmique est « plus tonique et plus longue » (*Défi A2*), « écoutez et comptez le nombre de syllabes puis répétez en allongeant la dernière syllabe de chaque énoncé » (*Alter Ego+ 1*, p.55) « le rythme du français est marqué par l'allongement de la dernière syllabe des groupes rythmiques. » (*Entre Nous 2*, p.77) Dans certains cas, l'« accentuation » n'est simplement pas définie en termes de paramètres prosodiques :

- Dans un exercice sur les groupes rythmiques : « Observez. Ecoutez l'accent. Répétez » (*Tendances A1*, p.13)
- « Ecoutez et observez la phrase. Où est placé l'accent ? » « il y a un **accent tonique** sur chaque groupe rythmique. L'accent tonique est toujours a| au début b| à la fin du

groupe rythmique, sur la dernière syllabe du dernier mot » (*Edito A1*, p.30)

- « Placez l'accent sur la dernière syllabe » (*Entre nous 1*, p.29)
- « Dans un mot, l'accent tonique est généralement placé sur la première voyelle ou la dernière voyelle ? » (*Génération A1*, p.21)

Certaines méthodes évoquent également la présence de pauses après les groupes rythmiques et les variations mélodiques :

- « [l'allongement de la dernière syllabe des groupes rythmiques] est suivi d'une pause plus ou moins longue. » (*Entre nous 2*, p.77)
- « la frontière entre les mots phonétiques est très importante. C'est une micro-pause (une respiration) précédée d'une légère variation d'intonation vers le haut » (*Edito B1*, p.15)
- « les énoncés sont composés de syllabes, ayant toutes la même durée, sauf la dernière, qui est accentuée et qui est deux fois plus longues que les autres. Ces syllabes sont enveloppées dans un schéma mélodique montant ou descendant selon que le groupe est un groupe rythmique ou un groupe de souffle (*sic*) » (*Alter Ego+ 1 [guide pédagogique]*, p.55)

Nous constatons bien ici que les mouvements mélodiques passent en second plan et ne sont pas clairement définis. La très fréquente présence de contours montants aux frontières du français n'est ainsi pas décrite, et seul le paramètre de durée est mis en avant. Dans certains cas cependant, la valeur continuative des contours montants est décrite. C'est le cas notamment dans *Défi A1* où le contour montant est décrit comme partant d'un niveau moyen à un niveau supérieur et permettant d'indiquer la continuation du tour de parole : « on monte la voix de chaque dernière syllabe du groupe de mots pour indiquer qu'on veut encore parler », « une intonation montante signale un élément important, en cours ou à venir. Son but est de maintenir l'attention de l'interlocuteur. » Le traitement séparé des paramètres prosodiques dans l'enseignement de l'accentuation du français nous paraît problématique. En effet, s'il n'est pas insensé de plus fortement insister sur l'allongement final que sur les mouvements mélodiques – notamment car une syllabe peut être

perçue comme accentuée sans variation mélodique uniquement grâce à la durée, en particulier lorsque le débit est rapide (Michelas, 2011) – il nous semble toutefois étrange de ne pas évoquer les contours mélodiques que les apprenants peuvent bien percevoir lors des exemples donnés. Se focaliser sur la durée en omettant la mélodie pourrait induire les apprenants en erreur en les faisant considérer que celle-ci n'est pas importante et qu'elle peut être ignorée. S'il est vrai que l'intonation est plus difficile à enseigner (les contours peuvent être montants, descendant ou même de diverses autres formes selon leur emplacement, le type de parole, le locuteur, etc.) elle reste très importante pour permettre entre autres une bonne segmentation du phrasé et ainsi une bonne compréhension de la part des interlocuteurs. Il serait donc plus judicieux d'attirer l'attention des apprenants sur les contours produits aux frontières prosodiques, leur faisant réaliser que ceux-ci sont généralement montants en français, et de travailler sur tous les paramètres prosodiques en même temps.

Les types d'exercices proposés dans les méthodes se ressemblent également. Il s'agit souvent d'écouter des phrases, répéter, souligner les syllabes accentuées. Celles-ci sont d'ailleurs généralement marquées visuellement en étant soulignées, en gras ou représentées par des rectangles ou ronds plus grands par rapport aux syllabes qui précèdent. Par exemple, dans *Alter Ego+ 1* (p.55) le groupe rythmique « à bientôt » est représenté par le schéma suivant : ● ● ●● . Ces visualisations sont aussi fréquemment associées à des logatomes (« dadadaaa » « titiTA »). Les gestes ou les métaphores sont aussi très souvent mis à profit, par exemple l'image d'un élastique qui se tend pour représenter la syllabe accentuée est donnée dans *Défi A2*. Dans *Tendance 1*, les apprenants sont invités à marquer les syllabes en tapant sur la table ou dans leurs mains et à taper plus fort sur la dernière syllabe tout en avançant d'un pas par syllabe et en tapant avec le pied pour marquer l'accent final. La méthode *Entre nous 1* propose dans son guide d'associer des mouvements des bras et de la main vers le haut quand l'intonation est montante et vers le bas quand elle est descendante (des *pitch gestures*, comme nous l'avons vu dans la section précédente 8.1.3). Enfin, dans *Edito B1* est proposée une activité originale où les apprenants doivent scander des noms de villes ou de pays en levant le poing sur la syllabe accentuée, à la manière d'une manifestation.

Les explications et exercices proposés dans les manuels spécialistes sont très proches

de ceux que nous avons observés dans les méthodes généralistes, bien que plus développés. On y retrouve le même traitement séparé des paramètres prosodiques et la focalisation sur l'allongement final. Dans le premier tome de *Phonétique progressive du français*, l'allongement final est d'abord travaillé, s'ensuit des leçons sur l'intonation avant de proposer une « synthèse rythmique et mélodique » où des activités travaillant les deux paramètres sont proposées. La plupart des autres méthodes, en plus de les traiter de façon séparée, ne mettent pas directement en lien l'accentuation et l'intonation. Les exercices proposés sont ici aussi des exercices d'écoute, de répétition et de repérage des syllabes accentuées ainsi que des exercices intégrant des méthodes de visualisations (rond ou carrés pour représenter les syllabes). On retrouve également des exercices où une phrase est allongée progressivement (en jouant en général avec la longueur d'un syntagme nominal) pour que l'apprenant prenne conscience de la mobilité de l'accent du français qui se place systématiquement en position finale.

Par exemple, dans *Phonétique en dialogues* (p.85) :

« Marie est revenue. Marie ? Marie-Claire ? Marie-Claire Lubert ? Marie-Claire Lubert est revenue ? ».

Contrairement aux méthodes généralistes, des activités sont proposées pour l'intonation non-finale, s'appuyant sur la visualisation en utilisant des flèches montantes et descendantes comme dans *Phonétique progressive du français* niveau avancé où sont intégrées des activités où les apprenants doivent identifier et noter les mouvements mélodiques. Dans *Phonétique progressive du français* niveau avancé et *Phonétique essentielle du français*, des exercices mettent en avant la fonction continuative des contours montants en demandant aux apprenants de noter si la phrase est finie ou non (selon si le contour mélodique est descendant ou montant). *La prononciation en classe* se distingue des trois autres méthodes en travaillant simultanément la durée et l'intonation. Ils proposent diverses activités basées sur des chansons et utilisent abondamment la visualisation, les gestes et les métaphores dans leurs activités. Par exemple, ils proposent dans une activité aux apprenants d'imiter le geste d'un archet pour représenter les contours montants et descendants, dans une autre de marquer physiquement l'allongement final en gardant les bras le long du corps lors des syllabes non-accentuées et de les étendre lorsque la syllabe est accentuée,

etc.

Dans toutes les méthodes et manuels que nous analysés nous n'avons pas observé de distinction entre les continuations mineures et majeures, et c'est ce point-ci que nous considérons comme étant leur plus grande lacune. En effet, les résultats de notre travail de thèse nous ont montré la difficulté des apprenants japonophones à distinguer les différents niveaux prosodiques et à réaliser des continuations différentes à leurs frontières, et nous avons mis en avant les difficultés de compréhension que cela peut poser. Alors que cette distinction entre les contours mineurs et majeurs est connue et largement diffusée depuis le modèle de Delattre – modèle semblant pourtant avoir inspiré les manuels spécialisés en phonétique du FLE, et étant par exemple directement intégré dans le manuel *La prononciation en classe* – il est surprenant de constater qu'aucune explication ni activité n'est proposée aux apprenants pour apprendre à différencier les contours mineurs des contours majeurs. Lorsque les énoncés présentés sont associés à une visualisation utilisant des flèches, elles sont identiques qu'elles soient en frontière de syntagme accentuel (donc de continuation mineure) ou en frontière de syntagme intonatif (continuation majeure). Il n'est nulle part fait mention de contours montants plus amples que ceux que l'on retrouve en fin de simples groupes rythmiques, et qui sont plus souvent associés à une pause. Ainsi, nous retrouvons des exercices où les représentations proposées sont très éloignées de la réalisation des énoncés.

Par exemple, dans *Phonétique essentielle du français niveau A1-A2* (p.36) un exercice donne comme consigne de relier des schémas alternant carrés et rectangles pour représenter des syllabes plus ou moins longues à des énoncés écrits et lus dans un enregistrement audio. La première phrase, « Tu vois, ce carré vert ! » (qui se distingue de la deuxième « Tu vois ce carré vert ! ») doit être reliée au schéma suivant : □□□□□□□. Or, dans cet énoncé la première frontière non finale correspond à une frontière de syntagme intonatif (« Tu **vois**, ») et la seconde (« ce **car** ») à une frontière de syntagme accentuel. Non seulement cette visualisation semble uniquement prendre la durée en compte, tandis que la forte montée mélodique de la continuation majeure à la frontière du syntagme intonatif (effectivement produite dans l'enregistrement associé) n'est pas rendue visible, mais elle n'est pas non plus fidèle au rythme de la phrase car la continuation majeure est suivie d'une forte pause silencieuse qui n'est pas représentée ici. Dans *La prononciation en classe*,

ce sont des lignes courbes qui représentent les contours mélodiques aux frontières des groupes rythmiques, et, à nouveau, les mêmes lignes sont utilisées pour les continuations mineures et majeures.

Ainsi, dans un exercice l'énoncé « Le **train** / pour Abidjan / part à sept heures » (p.59) est représenté par ce schéma : . Or, si des contours équivalents peuvent être produits à la fin des deux premiers groupes rythmiques, la continuation à la fin du premier groupe accentuel « **train** » est toutefois souvent naturellement produite comme mineure et celle à la fin du second groupe accentuel (frontière de syntagme intermédiaire) comme majeure. Un autre point pose problème : la segmentation en groupes rythmiques mise en évidence dans la phrase par les barres obliques et le fait que les courbes présentées soient séparées les unes des autres en laissant un espace entre elles. Une telle représentation pourrait induire en erreur les apprenants et leur laisser penser que chaque groupe rythmique est suivi d'une pause, et produire ainsi une frontière forte avec une pause importune à la fin du premier groupe accentuel. D'autant plus que nous avons pu observer un tel phrasé non natif perturbant la compréhension des unités de sens dans les productions de nos apprenants japonophones !

Autre exemple, toujours dans le même manuel, le poème de Maurice Carême « Le chat et le soleil » est utilisé dans une activité où chaque phrase est « retranscrite » en logatomes, surmontée de courbes représentant l'intonation (p.47).

Ainsi, la première phrase du poème « Le chat ouvrit les yeux, le soleil y rentra » est associé aux courbes suivantes : .

Le contour réalisé sur le premier groupe rythmique « le chat » serait le même que celui réalisé à la fin du deuxième groupe rythmique « ouvrit les yeux », alors qu'une continuation mineure est habituellement produite à la fin du premier et une continuation majeure à la fin du second (frontière de syntagme intonatif). Une critique similaire a été faite par Martin (2019), qui appuie le fait que le premier contour ne peut être réalisé de cette manière car en suivant le principe de contraste de pentes il devrait être produit comme descendant ou neutralisé selon son modèle (qui, rappelons-le, est basé sur la forme des mouvements mélodiques à l'intérieur de la voyelle ; la syllabe accentuée pouvant tout à fait être à une cible haute H\*, ce que nous considérons comme une montée mélodique).

### 8.3 Idées et propositions didactiques : comment enseigner les continuatifs en FLE ?

Dans cette section, nous proposons quelques idées d'activités sous la forme de fiches pédagogiques pour enseigner les contours continuatifs du français. Le public visé est plutôt des adultes (potentiellement aussi des adolescents), et nous précisons pour chaque activité les compétences et le niveau minimum requis. Dans ces exercices, nous suivons les principes de l'approche communicative en essayant d'intégrer le travail sur les continuatifs à des activités à but communicatif (se présenter, faire une description, etc.). Rappelons par ailleurs que cette manière d'enseigner la prononciation, à la fois par des exercices contrôlés mais aussi en l'intégrant à des activités communicatives a été jugée comme plus efficace que lorsqu'elle est uniquement travaillée par des exercices contrôlés (*focus-on-form* vs. *focus-en-forms*, Saito 2012). L'idée est toutefois d'intégrer les activités que nous proposons dans des séquences didactiques plus larges, que l'enseignant se chargera de préparer et d'adapter selon ses objectifs pédagogiques. C'est pour cela que nous avons choisi d'avoir recours à l'écrit, généralement utilisé dès les premiers cours de niveau débutant en FLE, bien que son introduction précoce dans l'apprentissage pourrait avoir un impact négatif sur la fluence des apprenants (Alazard, 2013). Les objectifs communicatifs et linguistiques sont renseignés en préambule de chaque activité.

### Activité 1 : Présentation

**Niveau : A1 et +**

**Objectifs communicatifs et linguistiques :**

- Découvrir les contours continuatifs majeurs et mineurs et être capable de les distinguer (perception et production)
- Savoir se présenter et parler de ses loisirs (⚠️ révisions – il ne s’agit pas d’une introduction à ces compétences qui sont un pré-requis)

**Texte**

Je m’appelle Emilie. Je suis belge et je vis en France.  
J’habite à Montgeron, une petite ville à côté de Paris,  
dans un grand appartement. Je travaille pour une émission de télévision. J’aime écouter de la musique et faire de la pâtisserie. De temps en temps, je vais au cinéma avec des amis.

**Exercice 1 : Mise en route**

1. *L’enseignant fait une première lecture à voix haute du texte en suivant le phrasé prosodique proposé dans le corrigé ci-dessous.*
2. *Il pose quelques questions de compréhension et, selon le niveau des apprenants, il n’hésite pas à faire quelques révisions si nécessaire en comparant avec les étudiants (« Et vous, vous vous appelez comment ? », « Quelle est votre nationalité ? »), etc.*

3. *Il fait à nouveau une lecture à voix haute, en utilisant des gestes différents pour les continuations mineures et majeures, partant d'un même point central à sa mi-hauteur. Pour les continuations majeures, il fait un grand et long geste du bras et de la main, et pour les continuations mineures, un geste plus court et montant moins en hauteur. Bien que ce soit généralement plus difficile à réaliser, il est conseillé de réaliser le geste avec le bras gauche (et vers la gauche), pour que les élèves puissent le voir dans le « bon » sens.*
4. *L'enseignant demande à un élève de lire la première phrase et corrige son intonation avec les gestes. De même pour la deuxième phrase. Il explique qu'il y a deux continuations, mineure (petite), et majeure (grande). Il fait lire la deuxième phrase à tous les élèves de la classe, et les corrige, et peut utiliser des logatomes (dadada) si certains élèves sont bloqués par le segmental.*

### **Exercice 2 : Perception**

1. *L'enseignant distribue le texte à compléter et après avoir expliqué les instructions il interroge les élèves pour compléter ensemble la première ligne (les deux premières phrases).*
2. *L'enseignant lit à nouveau le texte à voix haute, en respectant bien l'intonation proposée dans le corrigé, et les élèves complètent les cases. Cette fois-ci, il ne fait pas de gestes.*
3. *Correction de l'exercice. L'enseignant peut faire noter aux apprenants que les continuations majeures sont là où il y a des virgules, ou deux phrases « complètes » reliées par la conjonction de coordination « et ».*

Texte à compléter

Je m'appelle Emilie. Je suis **belge** et je **vis** en France.

J'**habite** à Montgeron, une petite **ville** à côté de Paris,

dans un grand appartement. Je **travaille** pour une émis-

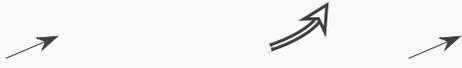
**sion** de télévision. J'aime écouter de la **musique** et **faire**

de la pâtisserie. De temps en **temps**, je vais au **cinéma**

avec des amis.

**Consignes** : Dessinez la flèche de la continuation que vous entendez. Continuation mineure :  Continuation majeure : 

## Corrigé texte


  
 Je m'appelle Emilie. Je suis belge et je vis en France.


  
 J'habite à Montgeron, une petite ville à côté de Paris,


  
 dans un grand appartement. Je travaille pour une émis-


  
 sion de télévision. J'aime écouter de la musique et faire


  
 de la pâtisserie. De temps en temps, je vais au cinéma

avec des amis.

**Exercice 3 : Se présenter****Consignes :**

Présentez-vous exactement de la même manière qu'Emilie. Par exemple : « Je m'appelle Natacha. Je suis russe et je vis en France. J'habite à Dijon, dans un petit appartement... »

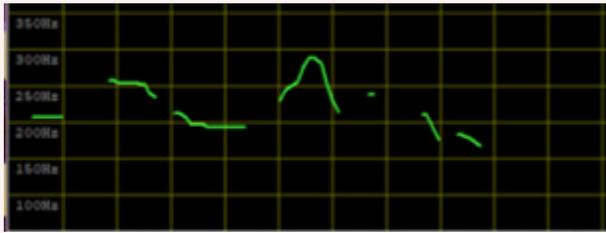
Dans l'activité suivante, nous proposons un exercice de visualisation avec le logiciel à accès libre en ligne RTPitch, qui, bien que rudimentaire, a le net avantage de permettre la visualisation de la courbe mélodique en temps réel tout en restant très facile d'utilisation (d'autant plus qu'il n'est pas nécessaire de le télécharger pour l'utiliser). Nous notons cependant que d'autres logiciels plus sophistiqués peuvent être utilisés en cours de langue, tel que WinPitch LTL (Germain & Martin, 2000), mais plutôt axés pour des séances spécifiques en laboratoire de langue et demandant de la part de l'enseignant une prise en main en amont.

#### **Activité 2 : Visualisation de la continuation**

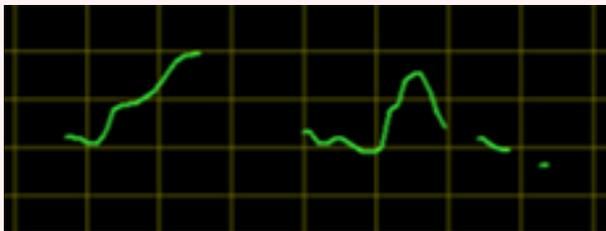
**Niveau : A2 et +**

**Objectifs communicatifs et linguistiques :** Être capable de distinguer les continuations mineures et majeures en production à l'aide de la visualisation

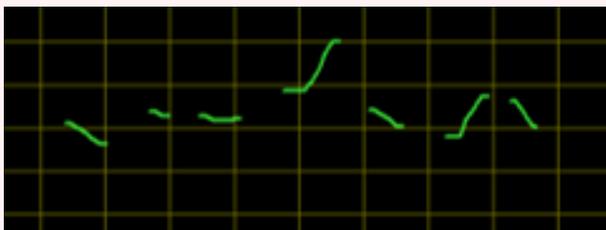
**Matériel :** Cette activité nécessite un microphone et un ordinateur, elle n'est donc réalisable qu'en laboratoire de langue. Cependant, si les apprenants ont le matériel nécessaire chez eux, cette activité peut être proposée en devoir maison (en demandant à l'apprenant d'envoyer les captures d'écran de leurs productions, par exemple).



La **soeur** de Marie est étudiante.



Hier, je suis **allée** à Paris.



La prochaine **fois**, je ne ferai pas ça.

**Consignes :**

Lisez et enregistrez les phrases avec le logiciel en ligne RTPitch en respectant l'intonation et les continuations mineures et majeures. Essayez de faire comme le modèle.

### Activité 3 : La liste des courses

**Niveau : A1 et +**

**Objectifs communicatifs et linguistiques :**

- Découvrir les contours d'énumération / de liste
- Être capable de produire des allongements suffisamment longs
- Lexique de l'alimentation, faire les courses

#### Dialogue

Marc : Je vais faire les courses, qu'est-ce qu'il te faut ?

Aurélie : Hmm... Il me faut du pain, du chocolat, des yaourts, des pommes, des bananes, une boîte de petits pois, des carottes, des bouteilles d'eau, du beurre... Voilà, ce sera tout, je pense.

**Exercice :**

1. *L'enseignant fait une première lecture à voix haute du texte en prononçant bien les allongements à la fin de chaque objet énuméré (contour de liste).*
2. *Il explique le contour de liste, il peut à nouveau s'aider des gestes et utiliser un mouvement différent de celui des continuation mineures et majeures, par exemple un mouvement plus long et plus horizontal que la continuation majeure et/ou en rejoignant les deux mains au centre, parallèles, et en écartant la main gauche suivant l'intonation.*

3. *Les apprenants doivent faire leur liste des courses. Et vous, qu'est-ce qu'il vous faut? (avec ou sans préparation préalable suivant le niveau)*

### Activité 4 : Faire une description

**Niveau : A2 et +**

**Objectifs communicatifs et linguistiques :**

- Découvrir les contours d'énumération / de liste
- Savoir produire des allongements finaux suffisamment longs
- Décrire une image, lexique basique de la maison et des meubles (compétences pouvant être travaillées simultanément)
- Révisions sur la position (à droite, à gauche, sur...) ⚠ pré-requis



FIGURE 8.1 – La Chambre à Arles, de Vincent van Gogh

**Exercice 1 :** « Décrivez la pièce où nous sommes. Il y a... » *L'enseignant incite les élèves à prononcer les allongements à la fin de chaque objet énuméré (contour de liste).*

**Exercice 2 :** « Décrivez le tableau de Vincent Van Gogh »

Enfin, nous encourageons également l'intégration d'exercices prosodiques aux **activités orales**. Par exemple, il est possible de :

- Travailler les continuations lors de l'apprentissage des habitudes et du quotidien.  
« Tous les **matins**, je me **lève** à 8 heures. Je me **douche**, je **prends** mon petit-déjeuner, puis je **sors** de chez **moi**... »
- Travailler les continuations majeures lors de l'apprentissage du futur / passé, à l'aide des ajouts à gauche. Exemple : « Qu'est-ce que vous avez fait hier ? » « **Hier**, j'ai fait ... » / « Qu'est-ce que vous avez prévu de faire ce weekend ? » « Ce **weekend**, je vais ... »
- Travailler la distinction continuations mineures / majeures avec les loisirs (comme nous l'avons abordé dans l'activité 1), corriger l'apprenant s'il produit des pauses importunes ou des continuations majeures à l'intérieur des groupes de sens « J'aime [jouer aux jeux vidéos] et [regarder des films.] »

# Conclusion

## Principaux résultats

Dans cette thèse, nous nous sommes intéressés à la production et la perception des contours continuatifs du français L2 chez des apprenants japonophones, étudiant le français dans leur pays natal. Nous avons mené quatre expériences : deux expériences de production (lecture de textes et description d'images) ainsi que deux expériences utilisant les compétences perceptives des apprenants (tâche de reproduction d'énoncés et de complétion de phrases). Cette recherche nous a conduits à un certain nombre d'observations et de conclusions :

1. Les apprenants japonophones semblent facilement capables d'intégrer la notion de continuations montantes dans leur interlangue : ils produisent statistiquement autant de montées mélodiques que les francophones du groupe contrôle et sont capables, dans une certaine mesure, de modifier des énoncés prosodiquement agrammaticaux en produisant une montée mélodique. Ces résultats ne vont pas dans le sens de l'hypothèse d'un transfert négatif de la L1 qui prédisait des difficultés pour produire des montées mélodiques aux frontières prosodiques, étant donné que celles du japonais sont généralement réalisées par un ton bas (L%).
2. Les apprenants japonophones ont toutefois des difficultés à distinguer les différentes frontières prosodiques et à les associer à différentes forces de continuation. En effet, contrairement aux francophones du groupe contrôle, ils ne produisent pas de montées mélodiques plus amples et d'allongements plus importants en frontière de IP (où une continuation majeure est attendue) que de AP (continuation mineure). De même, en perception, ils n'arrivent pas à distinguer les frontières de ip et de AP et à

les associer aux frontières syntaxiques appropriées.

3. En parole semi-spontanée (description d'images), les apprenants japonophones produisent bien des montées mélodiques plus amples en frontière de IP que de AP, comme les francophones natifs. Ils semblent ainsi associer les frontières discursives focales, liant les énoncés dans un même acte de parole, à une montée mélodique plus ample. Ce phénomène pourrait s'expliquer par un transfert positif de leur L1 et par la présence des BPM montants en japonais ayant une même fonction discursive.

Nos résultats ont également confirmé l'importance de la montée continuative en français. En effet, les francophones produisent bien une majorité de contours montants aux frontières prosodiques fortes et, face à des stimuli dont la prosodie est agrammaticale, privilégient la modification du contour plat en frontière de syntagme nominal en un contour montant plutôt que de restructurer prosodiquement l'énoncé. De plus, les résultats que nous avons obtenus en réitérant l'expérience de complétion de phrases de Michélas (2011) nous confirment que les francophones perçoivent les frontières de ip et les distinguent de celles produites à droite du deuxième AP d'un syntagme nominal en comportant trois. De même, en production, ils réalisent des continuations différentes aux frontières de IP que de AP en les associant à un allongement final plus conséquent et une montée mélodique plus ample. La question du ip en production semble cependant plus complexe car nous avons obtenu des résultats différents dans notre tâche de lecture de textes. En effet, les francophones ont produit des allongements de la voyelle accentuée significativement plus importants en frontière de ip que de AP, mais identiques à ceux produits en IP, et à l'inverse, des montées mélodiques moins importantes en frontière de ip que de IP mais identiques à celles produites en frontière de AP. Si les résultats de F0 peuvent s'expliquer par le fait que seul un « retour au niveau de référence » a été décrit aux frontières de ip, les résultats de durée ne correspondent pas aux descriptions qui ont été faites de ce niveau prosodique. Cependant, notre corpus n'avait pas pour ambition de décrire ce phénomène en particulier et il nous est impossible d'en tirer des conclusions, d'autant plus que notre classement dans la catégorie « ip » de certains items peut être discuté car il ne s'agissait pas nécessairement d'énoncés typiques semblables à ceux qui ont été analysés dans les études précédentes sur le syntagme intermédiaire.

---

Les résultats que nous avons obtenus dans notre thèse affirment l'importance d'enseigner les contours continuatifs aux apprenants du FLE. Nous nous sommes donc intéressés au traitement des continuatifs dans les méthodes de FLE généralistes et les manuels spécialisés dans l'enseignement de la prononciation. Un certain nombre de lacunes nous sont apparues, la plus flagrante étant l'absence de description des différentes sortes de continuations (mineures et majeures). Les difficultés des apprenants résidant principalement dans la gestion des différentes forces de frontières prosodiques, il nous semble important d'enseigner cette distinction entre continuations mineures et majeures ; nous avons donc proposé quelques idées d'activités sur ce sujet à la fin de cette thèse, ayant pour ambition d'appliquer concrètement les connaissances que nous avons acquises ou confirmées par notre travail de recherche.

## Contributions à l'hypothèse de transfert et de l'interlangue

Le transfert de la L1 sur la L2 est l'hypothèse principale que nous avons suivie et qui a servi de fil conducteur pour toutes nos expériences. Ainsi, nous avons tâché de comparer tous nos résultats à nos prédictions issues de cette hypothèse. Au final, celles-ci ont certes été en partie confirmées par nos résultats, mais nous n'avons pas observé de transfert direct de la L1 sur la L2. En effet, alors que nous prédisions des difficultés chez les apprenants japonophones à produire des contours montants aux frontières prosodiques du français, nous avons été surpris de constater qu'ils n'ont eu, dans la grande majorité des cas, aucune difficulté à produire ces montées mélodiques et qu'aucune différence significative entre les apprenants et les francophones du groupe contrôle n'a été observée sur ce point. Nos résultats ne confirment pas la tendance des apprenants japonophones à produire plus de contours descendants ou plats aux frontières prosodiques décrites dans les études antérieures (Anan, 1981 ; Kamiyama, 2005 ; Kojima, 1978 ; Lepetit, 1989).

Les bons taux de contours montants réalisés aux frontières prosodiques fortes sont cependant à nuancer. En effet, une analyse plus poussée des différents items nous a montré que les apprenants avaient tendance à généraliser les contours montants, là où les francophones produisaient des taux variables de montées mélodiques. Et, concernant quelques items en particulier (notamment des frontières finales de syntagme nominal ou verbal), les apprenants ont montré des taux de réalisation très bas, tandis que les francophones

produisaient systématiquement une montée.

Si un transfert direct de la L1 sur les taux de réalisation des continuations montantes n'a pas été observé, nous avons cependant bien noté un transfert négatif du japonais sur le français, se manifestant par la difficulté des apprenants à gérer les différents types de continuation et leurs réalisations phonétiques. En lecture de textes, les apprenants ont des difficultés à réaliser des montées mélodiques suffisamment amples aux continuations majeures. Cela confirme les observations faites précédemment par Anan (1981) sur la réalisation de montées mélodiques insuffisantes par les apprenants japonophones. Ils ont également des difficultés à produire un allongement final suffisamment important aux frontières de IP, en lecture de textes et en parole spontanée. Ces résultats sont aussi semblables à ceux obtenus par Marchand et Le Gac (2020) chez les turcophones qui réalisent correctement les schémas accentuels du français sans toutefois être capables de bien gérer les paramètres de durée et de F0 selon le type de frontière prosodique.

Notre recherche met donc en avant les limites de l'interférence de la L1 sur les réalisations prosodiques des apprenants. Si le phénomène de transfert permet bien d'expliquer certaines difficultés, il reste insuffisant, et d'autres facteurs semblent jouer un rôle. Ces constatations nous mènent donc à appuyer la notion d'interlangue (Selinker, 1972) – langue transitoire de l'apprenant marquée par le transfert de la L1, mais également d'autres facteurs, tels que des généralisations de schémas accentuels (Pešková, 2019 ; Ramírez Verdugo, 2006 ; Rasier & Hiligsmann, 2007), la production et généralisation de contours extra-montants (Santiago, 2014) ou la sursegmentation liée à l'insécurité linguistique (Barquero-Armesto, 2012 ; Kaglik & Boula de Mareüil, 2010 ; Lee, 2014). Par ailleurs, si nous n'avons pas observé de contours extra-montants dans nos données, nos apprenants japonophones ont bien tendance à sursegmenter leur phrasé en produisant un nombre important de pauses aux frontières non finales et ainsi à produire de courts *chunks*.

## L'importance des différences individuelles et du type de parole

Notre travail a mis en évidence l'importance d'aller au-delà de la généralisation – aussi nécessaire soit-elle pour nous permettre de tirer nos conclusions de recherche – et de prendre en compte les différences individuelles. Nous avons vu dans nos expériences de production que certains apprenants montraient des stratégies individuelles tout à fait dif-

---

férentes de la majorité d'entre eux. Ainsi, trois apprenants en particulier ont produit bien moins de contours montants que les autres, et les réalisations de l'un d'entre eux étaient semblables à celles que nous imaginions en cas de transfert du japonais sur le français (notamment pour la tâche de lecture de textes) : des contours bas aux frontières prosodiques, plats ou descendants, et l'absence d'allongement de la voyelle accentuée. Nous avons également noté des différences assez marquantes dans la prosodie des apprenants, telles qu'un empan mélodique très réduit chez certains. Mais ces différences individuelles ne concernent pas que les japonophones car nous avons également observé différentes stratégies chez les francophones, en particulier dans la tâche de description d'images où une des locutrices du groupe contrôle a réalisé des contours plats ou descendants aux frontières de IP, en les associant toutefois à de très forts allongements de la syllabe accentuée permettant d'exprimer la continuation. Si les variations des réalisations prosodiques est déjà bien attestée (voir par exemple les différents contours possibles lors des énumérations, Di Cristo 2016), il est intéressant de noter que dans notre étude elles semblent principalement liées à l'individu.

Les résultats différents obtenus dans la tâche de lecture de textes et de description d'images concernant la gestion de la F0 aux frontières de IP mettent également en avant l'importance de la variation situationnelle dans la prosodie des apprenants. Il est vrai qu'analyser la prosodie en parole spontanée peut être un exercice long et difficile, mais il est intéressant de voir que même dans le cas de descriptions d'images, et donc de parole semi-spontanée car cela reste une tâche contrôlée, nous pouvons obtenir des résultats différents de ceux de la parole lue.

Par cela, nous ne pouvons que partager le même point de vue que Christodoulides (2020) lorsque celui-ci affirme que « un effort doit être fait non seulement pour étudier les phénomènes phonétiques et prosodiques à travers les différents styles de parole, mais aussi pour intégrer la variation individuelle dans les analyses » (p.116). Si cela est vrai pour les analyses de la prosodie des locuteurs natifs, nous affirmons que cela l'est tout autant en prosodie L2.

## La question du niveau d'apprentissage

Avant la réalisation et l'analyse de nos expériences, nous avons émis l'hypothèse que les apprenants seraient capables d'apprendre la continuation du français : cela impliquait que les apprenants ayant un niveau de langue plus avancé obtiendraient de meilleurs résultats (plus proches de ceux des francophones natifs) que les apprenants de moindre niveau, en production et en perception. Cependant, les résultats que nous avons obtenus ne confirment pas clairement cette hypothèse :

- en production, le seul groupe d'apprenants montrant des résultats différents des deux autres est le groupe de débutants (A1-A2). Or celui-ci n'est composé que de deux apprenants, rendant toute généralisation un peu risquée. Malgré tout, nous pouvons noter que les apprenants tendent à produire moins de contours montants aux frontières de AP et de IP dans la tâche de lecture de textes, et en frontière de AP dans la tâche de description d'images, ce qui irait dans le sens d'une acquisition progressive de la continuation du français, mais qui peut tout à fait être dû au hasard ou à des différences individuelles.
- dans la tâche de complétion de phrases, tous les groupes de niveau ne réussissent pas l'expérience et nous n'observons aucune différence statistique entre eux.
- dans la tâche de reproduction d'énoncés, on remarque une (faible) amélioration pour certains stimuli, les groupes A1-A2 et B1-B2 les ayant significativement moins corrigés que les francophones tandis que les résultats du groupe de niveau le plus avancé (B1-B2) ne diffèrent pas significativement de ceux des francophones du groupe contrôle.

Ainsi, selon les expériences, nous n'observons pas ou peu de différences selon le niveau de langue des apprenants. Si un lien positif entre niveau de langue et meilleurs résultats a pu être établi dans d'autres études (Duběda, 2009 ; Santiago, 2014, entre autres) certains travaux révèlent que le niveau des apprenants n'est pas toujours corrélé à leur réussite, notamment en perception (Dupoux et al., 2008). Il est ainsi possible que le niveau de langue ne soit pas un facteur suffisamment important pour permettre de noter une claire amélioration entre la prosodie des apprenants débutants et celle des apprenants de

---

niveau intermédiaire. En particulier lorsque ceux-ci – comme c’est le cas pour nos apprenants – ne sont pas exposés à leur langue cible au quotidien, et ne travaillent que rarement leurs compétences phonético-phonologiques dans cette même langue, que ce soit en cours ou par eux-mêmes. Il est très probable que nos résultats concernant le niveau de langue puissent s’expliquer par ces différents facteurs, et que si l’on répliquait l’expérience sur des apprenants vivant dans un pays francophone ou de niveau bien plus avancé nous obtiendrions de meilleurs résultats, bien qu’encore une fois il soit possible qu’ils varient selon les individus.

## Limites et perspectives

Il existe un certain nombre de limites à notre recherche, principalement méthodologiques. Premièrement, il a été difficile pour nous de recruter des apprenants de niveau débutant, en particulier pour les tâches de production. Par cela, le groupe A1-A2 ne comportait que deux apprenants et il n’a ainsi pas été possible d’avancer de réelles conclusions sur la présence ou non d’une progression avec le niveau de langue. Ensuite, nous n’avons pas enregistré de japonophones dans leur langue maternelle, et ce point de repère nous a manqué pour la validation d’un potentiel transfert positif de la L1 en parole spontanée dans la production de montées continuatives majeures en ce qui concerne la fréquence fondamentale. Aussi, si nous avons pu montrer que les apprenants avaient des difficultés à produire les différents types de continuations (mineures et majeures), il serait intéressant de valider les difficultés communicatives qu’impliquent ces réalisations prosodiques non-natives. Nous pouvons par exemple imaginer une tâche de jugement par des francophones de stimuli resynthétisés en transférant la durée et la F0 d’énoncés non-natifs sur des stimuli natifs (à la manière de Kamiyama et Sakamoto, 2010 ou de Duběda, 2009, par exemple). Nous pensons également qu’il serait intéressant de comparer les résultats des japonophones à ceux d’apprenants de différentes autres L1. En effet, il est probable que les difficultés que nous avons décrites chez les apprenants japonophones et les spécificités de leurs réalisations prosodiques se retrouvent chez des apprenants d’autres L1, confirmant davantage la nécessité d’enseigner les contours continuatifs en classe de FLE.



# ANNEXES



## **Annexe A**

### **Formulaires de consentement**



UNIVERSITE PARIS DISDEROT

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

LABORATOIRE DE LINGUISTIQUE FORMELLE

(UMR 7110)

Fiche de participation à l'expérience : **Perception et production du français****Description de l'expérience**

Cette expérience s'inscrit dans le cadre général des expériences non invasives de linguistique. Cette expérience est composée de 4 tâches distinctes, les trois premières étant des tâches de répétition de suites de sons et de phrases et la quatrième étant une tâche de complétion de phrase ; il s'agit d'écouter le début d'une phrase et de sélectionner la suite la plus probable. Les 3 premières tâches sont enregistrées. Les données sont anonymisées et ne seront utilisées que dans le cadre de la recherche.

Cette étude nous permettra de mieux comprendre les phénomènes de perception et de production des apprenants du français en comparaison avec les résultats des natifs francophones.

Chaque tâche durant approximativement 10 min, la durée totale de l'expérience est d'environ 1 heure.

**CONSENTEMENT DE PARTICIPATION**

Je soussigné(e) ..... comprends la nature des recherches auxquelles je vais participer et comprends que j'ai le droit d'arrêter ma participation à tout moment sans avoir à fournir d'explication. Je comprends que le paiement sera relatif à la participation.

Date et signature

**Sujet :**

Nom, Prénom :

Adresse personnelle :

E-mail :

<i>Date</i>	<i>Nombre d'heures</i>	<i>€</i>
<b>Total</b>		

Je déclare avoir participé à cette expérience (dont j'ai lu la description) aux dates et durées mentionnées ci-dessus et avoir été rémunéré de la somme ci-dessus.

250 Fait à , le / /

Signature



## 同意書

上智大学  
SOPHIA UNIVERSITYuniversité  
PARIS  
DIDEROT  
PARIS 7U<sup>S</sup>PC  
Université Sorbonne  
Paris Cité

私は、下記の内容について説明を受け、了承の上で実験に参加します。

研究テーマ

日本人のフランス語学習者の発音

目的

本研究は日本人のフランス語学習者の発音問題について、調査することを目的としています。また、日本語の発音システムはフランス語の発音にどのように影響を与えているのかについても考察します。本実験の実施にあたり、パリディドロ大学と上智大学倫理委員会の承認を得ております(UMR 7110)。

内容

上智大学内の防音室にて、PC上で実験を行います。実験ではいくつかのタスクを行っていただきます。その後、アンケートにお答えいただきます。

実験協力者として選定される理由

日本語を母語とするフランス語学習者と認められた場合。

所要時間と謝礼

本実験の所要時間は1時間程度(休憩含む)を予定しています。途中、休憩や中断を希望する場合にはすぐに申し出てください。なお、本実験への参加に同意をした後も、口頭により同意を撤回することができます。同意を撤回しても参加者の方が不利益を被ることはありません。また、実験への参加に対して、謝礼を差し上げます。

危険性

本実験に参加することによる身体・精神への危険性は極めて低いですが、考え得る最大リスクとしては、倦怠感や疲労が挙げられます。もし、実験の途中でストレスや負担を感じ、実験の中止を希望する場合は、我慢せずすぐに申し出てください。速やかに実験を中止します。なお、実験の中止に伴う参加者の方への不利益は生じません。

参加について

本実験への参加は、参加者の自由です。仮に実験にご参加いただけなかったり、途中で実験を中止されても、いかなる不利益も被ることはありません。本実験は参加者に個人的な利益を与えるものではありませんが、発音習得に関して何かしらの貢献がもたらされることが期待されます。

機密保持とデータ保存

本実験で収集されたデータや個人情報は、研究目的のみで使用し、第三者に公開されることはありません。個人情報は、記号、あるいは、参加者管理番号によって管理・保存されるため、個人が特定されることはありません。なお、本実験で収集したデータは、パスワード付きの電子ファイル、もしくは、鍵のかかる保管庫内に保存されます。

本実験で得られた結果は、学会誌への投稿論文・記事として報告される可能性があります。また、国内での学会や研究会で報告する可能性もあります。しかし、研究結果の報告にあたり、研究対象者の個人情報を公開することはなく、研究対象者が特定されることはありません。

参加者控え A.2. FORMULAIRE EN JAPONAIS

---

当てはまるものにチェックを入れてください:

私は、私の匿名のデータ及び録音を教育目的で使用することに同意します。

私は、私の声の音声クリップがオンライン刊行物を通じて入手可能になることに同意します。

私は、フォローアップがあった場合には本研究の第二段階のために連絡を受けることに同意します。連絡を受けるのはこの目的のみとします。

その他

本実験への参加は、大学の授業における評価や成績とは一切関係がなく、実験参加は義務ではありません。また必要に応じて、支障のない範囲で研究計画書及び方法に関する資料の入手、閲覧ができます。なお、実験終了後も同意を撤回することができます。同意を撤回する場合も、不利益を被ることはありません。

問い合わせ先

研究実施者(Rachel ALBAR)、及び、研究責任者(Dr. Hiyon YOO) はいつでもご質問、ご相談をお受けいたします。メールあるいは直接口頭で、ご自由にご相談ください。

- PhD student, Rachel ALBAR: \*\*\*\*\*@gmail.com  
[Université Paris Diderot, Laboratoire de linguistique formelle, UMR 7110  
Visiting researcher at: Sophia University, Graduate School of Science and Technology & Graduate School of Linguistics 7-1 Kioi-cho, Chiyoda-ku, Tokyo, 102-8554 JAPAN (B229 Phonetics Laboratory & 4-295c Arai laboratory)]

- Dr. Hiyon YOO: yoo@linguist.univ-paris-diderot.fr,  
[Université Paris Diderot, Laboratoire de linguistique formelle, UMR 7110.]

参加者署名:

私、 \_\_\_\_\_ は、上記研究プロジェクトについて十分に説明を受け、本研究に参加することに同意致します。私は上記の説明文を全て読み、本研究に何が含まれているのかを理解しました。

記入日: \_\_\_\_\_年 \_\_\_\_\_月 \_\_\_\_\_日

氏名(自筆): \_\_\_\_\_

研究者署名:

私、 \_\_\_\_\_ は、本研究がどういったもので何が求められているか、また考えられるリスクについて(該当するもの)、志願者に対して注意深く説明致しました。

記入日: \_\_\_\_\_年 \_\_\_\_\_月 \_\_\_\_\_日

氏名(自筆): \_\_\_\_\_



## **Annexe B**

### **Questionnaire**

NOM : \_\_\_\_\_

Prénom : \_\_\_\_\_

Sexe (entourez votre réponse) : masculin / féminin

性別 男 女

Date de naissance : \_\_\_\_\_

生年月日

Lieu de naissance (Pays et région) : \_\_\_\_\_

出身(国、地域)

Langue(s) maternelle(s) (si plusieurs, le dire): \_\_\_\_\_ dialecte \_\_\_\_\_

母語

方言

Lieu de domiciliation actuel : \_\_\_\_\_

住所(町)

Niveau d'études/ d'instruction (entourez votre réponse) :

BAC BAC + 3 BAC + 5 Doctorat et plus

高校卒業 学士 修士 博士/以上

Autre Précisez :

他

Profession (si besoin, mettre étudiant, mère/ père au foyer, etc.) : \_\_\_\_\_

職業

## 学習歴

Années d'études du français : \_\_\_\_\_ フランス語専門 か 第二外国語又は第三外国語

フランス語を学習した期間

Fréquence (exemple nombre d'heures de cours par semaine) : \_\_\_\_\_

学習頻度

Motivation pour apprendre le français

フランス語に対してのモチベーション



Cours suivis jusqu'à présent :

今まで受けたフランス語の授業 \_\_\_\_\_

Voyage(s) en France &amp; durée : \_\_\_\_\_

フランスでの滞在期間

Utilisation actuelle du français : très fréquente...fréquente...moyenne...rare = environ...../semaine

最近のフランス語の使用 かなり 結構 普通 まれ 週に\_\_\_\_時間ほど

Niveau de français (entourez la réponse) A1 A2 B1 B2 C1 C1+

フランス語のレベル

Autre(s) langue(s) étudiées et niveau : \_\_\_\_\_

勉強した言語とレベル (フランス語以外)

Faites-vous ou avez-vous déjà fait de la musique ? Oui (précisez)..... Non

音楽を習ったことがありますか。

Fait le 256

日付

Signature

署名

## **Annexe C**

### **Expériences de production**

Il faut que je sois à Grenoble, samedi, vers quinze heures. Pourriez-vous m'indiquer à quelle heure est la correspondance à Valence ? Si je dois partir avant midi de Marseille, j'aimerais savoir s'il y aura un wagon restaurant. Il paraît que sur cette ligne, la SNCF prévoit un wagon de jeux pour les enfants. Est-ce que c'est vrai ? Vous prenez les réservations par téléphone ?

**Texte 2**

Est-ce que vous pourriez me donner la liste des restaurants de mon quartier ? J'habite à Saint-Barnabé, rue Montaigne. Je recherche quelque chose qui sorte un peu de l'ordinaire, d'un peu exotique : un restaurant japonais, par exemple. Je préférerais que ce ne soit pas une cuisine végétarienne.

Tu sais, maman, je me régale à Cassis. Il fait un temps magnifique et la mer est très bonne. Hier, je suis allé me promener dans les calanques: il faisait tellement de vent que j'ai failli être emportée ! Pour le bronzage, c'est l'idéal, mais pour la ligne, on peut pas en dire autant, avec tous ces marchands de glaces, et moi qui sais pas résister !

**Texte 4**

Ma sœur a une peur bleue de l'obscurité. Elle refuse absolument de sortir seule dès qu'il fait nuit. Elle veut toujours que quelqu'un l'accompagne. Mon père lui conseille d'emmener le chien avec elle : au moins, il la protégerait en aboyant au moindre danger.

## Annotation phonologique

clitic pw AP ip IP

## Texte 1

Il faut <sup>1</sup> que je <sup>2</sup> sois <sup>3</sup> à Grenoble <sup>4</sup>, samedi, <sup>5</sup> vers quinze <sup>6</sup> heures. Pourriez-<sup>7</sup>  
 vous <sup>8</sup> m'indiquer <sup>9</sup> à quelle <sup>10</sup> heure <sup>11</sup> est la <sup>12</sup> correspondance <sup>13</sup> à Valence ? Si  
 je <sup>14</sup> dois <sup>15</sup> partir <sup>16</sup> avant <sup>17</sup> midi <sup>18</sup> de Marseille, <sup>19</sup> j'aimerais <sup>20</sup> savoir <sup>21</sup> s'il y aura  
<sup>22</sup> un wagon <sup>23</sup> restaurant. Il paraît <sup>24</sup> que sur <sup>25</sup> cette <sup>26</sup> ligne, <sup>27</sup> la SNCF <sup>28</sup> prévoit <sup>29</sup>  
 un wagon <sup>30</sup> de jeux <sup>31</sup> pour les enfants. Est-ce que <sup>32</sup> c'est vrai ? Vous prenez <sup>33</sup> les  
 réservations <sup>34</sup> par téléphone ?

## Texte 2

Est-ce que <sup>35</sup> vous <sup>36</sup> pourriez <sup>37</sup> me donner <sup>38</sup> la liste <sup>39</sup> des restaurants <sup>40</sup> de mon <sup>41</sup>  
 quartier ? J'habite <sup>42</sup> à Saint-Barnabé, <sup>43</sup> rue Montaigne. Je recherche <sup>44</sup> quelque  
<sup>45</sup> chose <sup>46</sup> qui sorte <sup>47</sup> un peu <sup>48</sup> de l'ordinaire, <sup>49</sup> d'un peu <sup>50</sup> exotique : un  
 restaurant <sup>51</sup> japonais, <sup>52</sup> par exemple. Je préférerais <sup>53</sup> que ce <sup>54</sup> ne soit <sup>55</sup> pas <sup>56</sup>  
 une cuisine <sup>57</sup> végétarienne.

## Texte 3

Tu sais, <sup>58</sup> maman, <sup>59</sup> je me <sup>60</sup> régale <sup>61</sup> à Cassis. Il fait <sup>62</sup> un temps <sup>63</sup> magnifique <sup>64</sup> et  
 la <sup>65</sup> mer <sup>66</sup> est très <sup>67</sup> bonne. Hier, <sup>68</sup> je suis <sup>69</sup> allé <sup>70</sup> me promener <sup>71</sup> dans les <sup>72</sup>  
 calanques: il faisait <sup>73</sup> tellement <sup>74</sup> de vent <sup>75</sup> que j'ai failli <sup>76</sup> être emportée ! Pour  
 le <sup>77</sup> bronzage, <sup>78</sup> c'est l'idéal, <sup>79</sup> mais pour <sup>80</sup> la ligne, <sup>81</sup> on peut <sup>82</sup> pas <sup>83</sup> en dire <sup>84</sup>  
 autant, <sup>85</sup> avec <sup>86</sup> tous <sup>87</sup> ces marchands <sup>88</sup> de glaces, <sup>89</sup> et moi <sup>90</sup> qui sais <sup>91</sup> pas <sup>92</sup>  
 résister !

## Texte 4

Ma sœur <sup>93</sup> a une <sup>94</sup> peur <sup>95</sup> bleue <sup>96</sup> de l'obscurité. Elle refuse <sup>97</sup> absolument <sup>98</sup> de  
 sortir <sup>99</sup> seule <sup>100</sup> dès qu'il <sup>101</sup> fait <sup>102</sup> nuit. Elle veut <sup>103</sup> toujours <sup>104</sup> que quelqu'un <sup>105</sup>  
 l'accompagne.  
 Mon père <sup>106</sup> lui conseille <sup>107</sup> d'emmener <sup>108</sup> le chien <sup>109</sup> avec <sup>110</sup> elle : au moins, <sup>111</sup> il  
 la <sup>112</sup> protégerait <sup>113</sup> en aboyant <sup>114</sup> au moindre <sup>115</sup> danger.

Image 1



Image 2



Image 3



Image 4



## **Annexe D**

### **Expériences de perception**

## D.1 Stimuli et consignes Expérience 3 (tâche de reproduction d'énoncés modifiés)

TABLEAU D.1 – Stimuli Expérience 3

1AP	N	Marie mange une banane
		Magali boit du café
2AP	N+N	Sarah et Marie sont professeurs
		Marie et Magali sont amies
	Long N	Le cours d'anglais s'est bien déroulé Les enfants de Rémi sont gentils

### Consignes expérience 3

Merci pour votre participation

Vous allez entendre des phrases en français.

Reproduisez naturellement les phrases suivantes.

Les consignes sont en français pour tous les participants, y compris les apprenants japonophones, en raison de l'absence de traduction adéquate de l'expression "reproduire naturellement". Des précisions ont cependant été données en japonais au moment de la passation d'expérience, lorsque c'était nécessaire.

## D.2 Stimuli et consignes Expérience 4 (tâche de complétion de phrases)

Les phrases des stimuli ont été lues par quatre francophones natifs et enregistrées dans leur intégralité, sans les "... " que nous avons ajouté ici pour séparer ce qui correspond aux stimuli audio et écrit. La première partie des phrases (avant "...") est entendue par les participants, et ils doivent sélectionner la suite de la phrase à l'écrit (après "...").

### Condition AP

1. Le gérant du restaurant ... d'Aurélié parle avec un client
2. Les amis de Rémi ... de Dijon étudient tous à l'université
3. Le niveau de japonais ... de Marina est meilleur chaque année
4. Le musée du chocolat ... de Paris est très intéressant à visiter
5. L'agenda du patron ... d'Héléna n'est pas sur son bureau
6. Le diner de Sarah ... de ce soir commence à 19 heures
7. Le match de tennis ... de cette semaine n'aura pas lieu
8. La porte de la maison ... de Julien ne ferme pas
9. Le câble du téléphone ... de Julie ne fonctionne pas

### Condition ip

Ici, les mots entre parenthèse ne font pas partie de la phrase initiale mais indiquent la deuxième proposition écrite pour la complétion de la phrase

10. L'école de musique ... (de Marie) ferme le mois prochain
11. Le magasin de bricolage ... (du quartier) se trouve assez loin
12. La vitre de la voiture ... (de Marc) s'est cassée
13. L'aéroport de la ville ... (de Lyon) se trouve à une demie heure du centre
14. Les professeurs d'anglais ... (de l'école) sont très sévères
15. Cette bouteille de vin ... (de Bourgogne) coûte 30 euros
16. Les étudiants de l'université ... (de Strasbourg) organisent une grande fête
17. Le concert de piano ... (de jeudi) va durer 3 heures
18. Le livre de cuisine ... (de ma mère) coûte une dizaine d'euros

## Consignes expérience 4

### Version en japonais

参加してくれてありがとうございます。  
今から聞く文章は途中で切れています。その続く文章として2つの文章が表示されます。文法的にはどちらも正しい文章です。聞いた文章の発音から予想される続きとしてより自然だと思うものを選んでください。  
左記の文章を選ぶ場合はAを押してください。  
右記の文章を選ぶ場合はLを押してください。

### Version en français

Vous allez entendre un début de phrase puis voir apparaître l'écran deux suites possibles pour l'extrait que vous venez d'entendre.  
Pour chaque extrait, vous devez deviner quelle est la suite la plus probable tel qu'il a été prononcé.  
Si vous pensez qu'il s'agit des mots à droite, appuyez sur la touche L. Si vous pensez qu'il s'agit des mots à gauche, appuyez sur la touche A. Attention, en aucun cas vous ne devez chercher à terminer les phrases mais bien à deviner quels étaient les mots qui faisaient suite aux extraits que vous entendez.

# Bibliographie

- ADDA-DECKER, M., NEMOTO, R. & DURAND, J. (2009). Stratégies de démarcation du mot en français : une étude expérimentale sur grand corpus. *Journées d'Etudes Linguistiques de Nantes (JEL 2009)*.
- AKAHANE-YAMADA, R., TOHKURA, Y., BRADLOW, A. & PISONI, D. (1996). Does training in speech perception modify speech production? *Proceeding of Fourth International Conference on Spoken Language Processing, ICSLP '96*, 2, 606-609 vol.2. <https://doi.org/10.1109/ICSLP.1996.607434>
- ALAZARD, C. (2013). *Rôle de la prosodie dans la fluence en lecture oralisée chez des apprenants de Français Langue Étrangère* (Thèse de doctorat). Université Toulouse le Mirail-Toulouse II.
- ALBAR, R. (2020). Production de la continuation du français par des apprenants japo-phones : gestion de la F0 et de la durée. *6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 1 : Journées d'Études sur la Parole*, 10-18.
- ALBAR, R. & YOO, H. (2019). Phonological Awareness of French Rising Contours in Japanese Learners. *Interspeech 2019*, 2548-2552. <https://doi.org/10.21437/Interspeech.2019-2856>
- ALBAR, R. & YOO, H. (2020). The production of French continuation contours at different prosodic boundaries by Japanese learners. *Speech Prosody 2020*, 890-894. <https://doi.org/10.21437/SpeechProsody.2020-182>

- ALBIN, A. & WANG, R. (2020). When does intonational transfer occur? A comparative study of interrogative rises in four groups of L2 Japanese learners. *Speech Prosody 2020*, 857-861. <https://doi.org/10.21437/SpeechProsody.2020-175>
- ANAN, F. (1981). Furansugo kyōiku to intonēshon kanshō no mondai ni tsuite [notre traduction : le problème de l'interférence prosodique lors de l'enseignement de l'intonation]. *furansu bungaku ronshū*, 16, 16-23. [https://doi.org/10.20767/elfk.16.0\\_16](https://doi.org/10.20767/elfk.16.0_16)
- ANDERSON-HSIEH, J. (1992). Using electronic visual feedback to teach suprasegmentals. *System*, 20(1), 51-62. [https://doi.org/10.1016/0346-251X\(92\)90007-P](https://doi.org/10.1016/0346-251X(92)90007-P)
- ARAI, T. & GREENBERG, S. (1997). The temporal properties of spoken Japanese are similar to those of English. *Fifth european conference on speech communication and technology (Eurospeech 1997)*, 1011-1014.
- ASTÉSANO, C. & BERTRAND, R. (2016). Accentuation et niveaux de constituance en français : enjeux phonologiques et psycholinguistiques. *Langue française*, 191(3), 11-30. <https://doi.org/10.3917/lf.191.0011>
- ASTÉSANO, C., BERTRAND, R., ESPESSE, R. & NGUYEN, N. (2012). Perception des frontières et des proéminences en français. *Proceedings of the Joint Conference JEP-TALN-RECITAL 2012, volume 1 : JEP*, 353-360.
- BAILS, F., SUÁREZ-GONZÁLEZ, N., GONZÁLEZ-FUENTE, S. & PRIETO, P. (2019). Observing and producing pitch gestures facilitates the learning of Mandarin Chinese tones and words. *Studies in Second Language Acquisition*, 41(1), 33-58. <https://doi.org/10.1017/S0272263118000074>
- BAKER, W. & TROFIMOVICH, P. (2006). Perceptual paths to accurate production of L2 vowels : The role of individual differences. *IRAL*, 44(3), 231-250. <https://doi.org/doi:10.1515/IRAL.2006.010>
- BARQUERO-ARMESTO, M. Á. (2012). A comparative study on accentual structure between Spanish learners of French interlanguage and French native speakers. *Speech Prosody 2012*.
- BARTKOVA, K., BONNEAU, A., COLOTTE, V. & DARGNAT, M. (2012). Productions of continuation contours by French Speakers in L1 (French) and L2. *Speech Prosody 2012*.
- BAUMANN, S., GRICE, M., BENZMÜLLER, R. et al. (2000). GToBI-a phonological system for the transcription of German intonation. *Prosody*, 21-28.

- BECKMAN, M. E. (1982). Segment duration and the 'mora' in Japanese. *Phonetica*, 39(2-3), 113-135.
- BECKMAN, M. E. & PIERREHUMBERT, J. B. (1986). Intonational structure in Japanese and English. *Phonology*, 3, 255-309. <https://doi.org/10.1017/S095267570000066X>
- BERTRAND, R., BLACHE, P., ESPESSE, R., FERRÉ, G., MEUNIER, C., PRIEGO-VALVERDE, B. & RAUZY, S. (2008). Le CID - Corpus of Interactional Data - Annotation et Exploitation Multimodale de Parole Conversationnelle. *Revue TAL*, 49(3), pp.105-134.
- BERTRAND, R. & PORTES, C. (2012). Pour une approche phonologique de la prosodie dans l'interaction. *Langue française*, 175(3), 19-36. <https://doi.org/10.3917/lf.175.0019>
- BERTRAND, R., PORTES, C. & SABIO, F. (2007). Distribution syntaxique, discursive et interactionnelle des contours intonatifs du français dans un corpus de conversation. *TRANEL. Travaux Neuchâtelois de Linguistique*, (47), 59-77.
- BEST, C. T. (1995). A direct realist view of cross-language speech perception. *Speech perception and linguistic experience*, 171-206.
- BIGI, B. (2015). SPPAS-multi-lingual approaches to the automatic annotation of speech. *The Phonetician*, 111(112), 54-69.
- BOERSMA, P. (2001). Praat, a system for doing phonetics by computer. *Glott. Int.*, 5(9), 341-345.
- BONGAERTS, T., MENNEN, S. & SLIK, F. v. d. (2000). Authenticity of pronunciation in naturalistic second language acquisition : The case of very advanced late learners of Dutch as a second language. *Studia Linguistica*, 54(2), 298-308. <https://doi.org/10.1111/1467-9582.00069>
- BONGAERTS, T., van SUMMEREN, C., PLANKEN, B. & SCHILS, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19(4), 447-465. <https://doi.org/10.1017/S0272263197004026>
- BORDEN, G., GERBER, A. & MILSARK, G. (1983). Production and perception of the /r/-/l/ contrast in Korean adults learning English. *Language learning*, 33(4), 499-526.
- BRIÈRE, E. J. (1966). An Investigation of Phonological Interference. *Language*, 42(4), 768. <https://doi.org/10.2307/411832>
- CHRISTODOULIDES, G. (2020). Variation prosodique des styles de parole et interface syntaxe-prosodie : Étude sur corpus à grande échelle. *6e conférence conjointe Journées d'Études*

- sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 1 : Journées d'Études sur la Parole, 109-117.
- CORDER, S. (1967). The Significance of Learners' Errors. *Reprinted from IRAL, V/4.*
- DE BOT, K. (1983). Visual feedback of intonation I : Effectiveness and induced practice behavior. *Language and speech*, 26(4), 331-350. <https://doi.org/10.1177/002383098302600402>
- DELAIS-ROUSSARIE, E. (2000). Vers une nouvelle approche de la structure prosodique. *Langue française*, 126(1), 92-112. <https://doi.org/10.3406/lfr.2000.991>
- DELAIS-ROUSSARIE, E. (2005). Phonologie et Grammaire : études et modélisation des interfaces prosodiques. *Mémoire d'HDR. ERSS, Université de Toulouse 2.*
- DELAIS-ROUSSARIE, E. & POST, B. (2008). Unités prosodiques et grammaire de l'intonation : vers une nouvelle approche. *Actes des Journées d'étude sur la Parole JEP-TALN, 8.*
- DELAIS-ROUSSARIE, E., POST, B., AVANZI, M., BUTHKE, C., DI CRISTO, A., FELDHAUSEN, I., JUN, S.-A., MARTIN, P., MEISENBURG, T., RIALLAND, A., SICHEL-BAZIN, R. & YOO, H.-Y. (2015). Intonational phonology of French : Developing a ToBI system for French. In S. FROTA & P. PRIETO (Éd.), *Intonation in Romance* (p. 63-100). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199685332.003.0003>
- DELAIS-ROUSSARIE, E., POST, B. & YOO, H. (2020). Unités prosodiques et grammaire intonative du français : vers une nouvelle approche. *6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 1 : Journées d'Études sur la Parole*, 145-153.
- DELAIS-ROUSSARIE, E. & YOO, H. (2011). Learner corpora and prosody : from the COREIL corpus to principles on data collection and corpus design. *Poznań Studies in Contemporary PSCIL*, 47, 28-39.
- DELAIS-ROUSSARIE, E., YOO, H. & POST, B. (2011). Quand frontières prosodiques et frontières syntaxiques se rencontrent. *Langue française*, 170(2), 29-44. <https://doi.org/10.3917/lf.170.0029>

- DELATTRE, P. (1938). L'accent final en français : accent d'intensité, accent de hauteur, accent de durée. *French Review*, 141-145.
- DELATTRE, P. (1939). Accent de mot et Accent de Groupe. *The French Review*, 13(2), 141-146.
- DELATTRE, P. (1966a). Le mot est-il une entité phonétique en français? *Studies in French and Comparative Phonetics* (p. 145-153). De Gruyter.
- DELATTRE, P. (1966b). Les Dix Intonations de base du français. *The French Review*, 40(1), 1-14.
- DELL, F. (1984). L'accentuation dans les phrases en français. In F. DELL, D. HIRST & J. VERGNAUD (Éd.), *Forme sonore du langage : structure des représentations en phonologie* (p. 65-122). Hermann.
- DETEY, S. & RACINE, I. (2015). Does perception precede production in the initial stage of French nasal vowel quality acquisition by Japanese learners? A corpus-based discrimination experiment. *18th International Congress of Phonetic Sciences, ICPHS 2015*, paper n°894.
- DI CRISTO, A. (2016). *Les musiques du français parlé : essais sur l'accentuation, la métrique, le rythme, le phrasé prosodique et l'intonation du français contemporain*. De Gruyter.
- DÍAZ-CAMPOS, M. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in second language acquisition*, 26(2), 249-273.
- D'IMPERIO, M. & MICHELAS, A. (2009). Interface entre structure syntaxique et structure prosodique : le syntagme intermédiaire en français. *Actes de la Conférence Interface Discours & Prosodie (IDP)*, 145-156.
- D'IMPERIO, M., MICHELAS, A. & PORTES, C. (2016). École d'Aix et approches tonales de l'intonation du français : un aperçu des recherches depuis les années 1970. *Langue française*, (3), 31-46.
- DUBĚDA, T. (2009). L'accent initial en tchèque et en français : un cas de transfert positif? *Acquisition et interaction en langue étrangère*, (2/2009), 43-66. <https://doi.org/10.4000/aile.4536>
- DULAY, H. C. & BURT, M. K. (1974). Errors and Strategies in Child Second Language Acquisition. *TESOL Quarterly*, 8(2), 129. <https://doi.org/10.2307/3585536>

- DUPOUX, E., SEBASTIÁN-GALLÉS, N., NAVARRETE, E. & PEPPERKAMP, S. (2008). Persistent stress 'deafness' : The case of French learners of Spanish. *Cognition*, 106(2), 682-706. <https://doi.org/10.1016/j.cognition.2007.04.001>
- DURAND, J., LAKS, B. & LYCHE, C. (2002). La Phonologie du français contemporain (PFC) : usages, variétés et structure. *Romanistische Korpuslinguistik romance corpus linguistics*, 93-106.
- ERVIN-TRIPP, S. M. (1974). Is Second Language Learning like the First. *TESOL Quarterly*, 8(2), 111. <https://doi.org/10.2307/3585535>
- ESCUDERO, D., GONZÁLEZ-FERRERAS, C. & CARDEÑOSO-PAYO, V. (2018). Analysis of the efficiency of repeating activities for improving prosody in L2 pronunciation training. *Speech Prosody 2018*, 299-303. <https://doi.org/10.21437/SpeechProsody.2018-61>
- FANT, G., KRUCKENBERG, A. & NORD, L. (1991). Prosodic and segmental speaker variations. *Speech communication*, 10(5-6), 521-531.
- FELDHAUSEN, I. & DELAIS-ROUSSARIE, E. (2012). La structuration prosodique et les relations syntaxe/ prosodie dans le discours politique. *Actes de la conférence conjointe JEP-TALN-RECITAL 2012, volume 1 : JEP*, 9-16.
- FIELD, J. (2003). Promoting perception : Lexical segmentation in L2 listening. *ELT journal*, 57(4), 325-334.
- FLEGE, J. E. (1995). Second language speech learning : Theory, findings, and problems. *Speech perception and linguistic experience : Issues in cross-language research*, 92, 233-277.
- FLEGE, J. E. (2008). Give Input a Chance! In T. PISKE & M. YOUNG-SCHOLTEN (Éd.), *Input Matters in SLA* (p. 175-190). Multilingual Matters. <https://doi.org/10.21832/9781847691118-012>
- FLEGE, J. E. & FLETCHER, K. L. (1992). Talker and listener effects on degree of perceived foreign accent. *The Journal of the Acoustical Society of America*, 91(1), 370-389. <https://doi.org/10.1121/1.402780>
- FLEGE, J. E. & HILLENBRAND, J. (1986). Differential use of temporal cues to the /s-/z/ contrast by native and non-native speakers of English. *The Journal of the Acoustical Society of America*, 79(2), 508-517. <https://doi.org/10.1121/1.393538>

- FLEGE, J. E., MUNRO, M. J. & MACKAY, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America*, 97(5), 3125-3134. <https://doi.org/10.1121/1.413041>
- FLEGE, J. E., YENI-KOMSHIAN, G. H. & LIU, S. (1999). Age Constraints on Second-Language Acquisition. *Journal of Memory and Language*, 41(1), 78-104. <https://doi.org/10.1006/jmla.1999.2638>
- FLYNN, S. (1987). *A parameter-setting model of L2 acquisition : experimental studies in anaphora* (T. 5). Springer, Dordrecht.
- FÓNAGY, I. (1980). L'accent français : accent probabilitaire (dynamique d'un changement prosodique). *Studia Phonetica Montréal*, 15, 123-233.
- FOUGERON, C. é. & JUN, S.-A. (1998). Rate effects on French intonation : prosodic organization and phonetic realization. *Journal of Phonetics*, 26(1), 45-69. <https://doi.org/10.1006/jpho.1997.0062>
- GABRIEL, C., STAHNKE, J. & THULKE, J. (2014). On the acquisition of French speech rhythm in a multilingual classroom : Evidence from linguistic and extra-linguistic data. *SHS Web of Conferences*, 8, 1267-1283.
- GADET, F. (1996). Niveaux de langue et variation intrinsèque. *Palimpsestes. Revue de traduction*, (10), 17-40. <https://doi.org/10.4000/palimpsestes.1504>
- GERMAIN, A. & MARTIN, P. (2000). Présentation d'un logiciel de visualisation pour l'apprentissage de l'oral en langue seconde. *Apprentissage des Langues et Systèmes d'Information et de Communication (Alsic)*, 3(1), 61-76.
- GLUHAREVA, D. & PRIETO, P. (2017). Training with rhythmic beat gestures benefits L2 pronunciation in discourse-demanding situations. *Language Teaching Research*, 21(5), 609-631. <https://doi.org/10.1177/1362168816651463>
- GRANENA, G. & LONG, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, 29(3), 311-343.
- GROVER, C., JAMIESON, D. G. & DOBROVOLSKY, M. B. (1987). Intonation in English, French and German : Perception and Production. *Language and Speech*, 30(3), 277-295. <https://doi.org/10.1177/002383098703000307>

- GUYOT-TALBOT, A., HEIDLMAYR, K. & FERRAGNE, E. (2016). Entraînements à la prosodie des questions ouvertes et fermées de l'anglais chez des apprenants francophones. *Actes de la conférence conjointe JEP-TALN-RECITAL 2016. volume 1 : JEP*, 265-273.
- HAMADA, Y. (2019). Shadowing : What is It ? How to Use It. Where Will It Go ? *RELC Journal*, 50(3), 386-393. <https://doi.org/10.1177/0033688218771380>
- HIRATA, Y. & KELLY, S. D. (2010). Effects of Lips and Hands on Auditory Learning of Second-Language Speech Sounds. *Journal of Speech, Language, and Hearing Research*, 53(2), 298-310. [https://doi.org/10.1044/1092-4388\(2009/08-0243\)](https://doi.org/10.1044/1092-4388(2009/08-0243))
- HIRATA, Y., KELLY, S. D., HUANG, J. & MANANSALA, M. (2014). Effects of Hand Gestures on Auditory Learning of Second-Language Vowel Length Contrasts. *Journal of Speech, Language, and Hearing Research*, 57(6), 2090-2101. [https://doi.org/10.1044/2014\\_JSLHR-S-14-0049](https://doi.org/10.1044/2014_JSLHR-S-14-0049)
- HIRST, D., DI CRISTO, A. & ESPESER, R. (2000). Levels of representation and levels of analysis for the description of intonation systems. *Prosody : Theory and experiment* (p. 51-87). Springer.
- HIRST, D. & ESPESER, R. (1993). Automatic modelling of fundamental frequency using a quadratic spline function. *Travaux de l'Institut de Phonétique d'Aix-en-Provence*, (15), 75-85.
- HORGUES, C. (2010). *Prosodie de l'accent français en anglais et perception par des auditeurs anglophones* (Thèse de doctorat). Université Paris Diderot . Paris 7.
- IGARASHI, Y. (2015). Intonation. In H. KUBOZONO (Éd.), *Handbook of Japanese phonetics and phonology* (p. 525-568). De Gruyter Mouton.
- IGARASHI, Y. & KOISO, H. (2012). Kumatsu kyōkai onchō no pitchirenji ni ataeru yōin "nihongo hanashi kotoba kōpasu" no bunseki [Factors Affecting Pitch Ranges of Boundary Pitch Movements : An Analysis of the Corpus of Spontaneous Japanese]. *2ème édition du Workshop sur les corpus linguistiques du japonais [notre traduction]*.
- INOUE, F. (1997). Intonēshon no shakaisei [notre traduction : approche sociolinguistique de l'intonation]. *Akusento, intonēshon, rizumu to pōzu [Accent, intonation, rythme et pause]* (p. 143-168). Sanseido.
- JAMES, E. F. (1976). The acquisition of prosodic features of speech using a speech visualizer. *14(3)*, 227-244. <https://doi.org/doi:10.1515/iral.1976.14.3.227>

- JUN, S.-A. & FOUGERON, C. (2000). A Phonological Model of French Intonation. In N. IDE, J. VÉRONIS & A. BOTINIS (Éd.), *Intonation* (p. 209-242). Springer Netherlands. [https://doi.org/10.1007/978-94-011-4317-2\\_10](https://doi.org/10.1007/978-94-011-4317-2_10)
- KAGLIK, A. & BOULA DE MAREÛIL, P. (2010). Polish-accented French prosody in perception and production : transfer or universal acquisition process? *Speech Prosody 2010-Fifth International Conference*.
- KAMINSKAÏA, S. (2016). Syntax-Prosody interface : a case study of enumerations. *41st Forum of the Linguistic Association of Canada and the United States (LACUS)*.
- KAMIYAMA, T. (2005). L1 nihongo washa no furansugo hatsuon ni okeru purosodjī : sono keikō no shikaku-ka to sūchika no kokoromi [La prosodie de phrases lues en français par des apprenants japonophones : essai de représentation graphique et de quantification de quelques tendances observées]. *Toukyou daigaku gaikokugo kyōuikugaku kenkyūkai 2004 nendo kiyō* [Bulletin de la FLTA (Foreign Language Teaching Association) de l'Université de Tokyo 2004-2005, 103-110.
- KAMIYAMA, T. & SAKAMOTO, M. (2010). Intonation des questions totales en français langue étrangère : suffit-il d'enseigner et apprendre la montée finale? *XXVIIIèmes Journées d'étude sur la Parole*, 393-396.
- KAMIYAMA, T. & YAMAMOTO, Y. (2007). Visual representation of prosody for tactful communication skills—the case of request in Japanese as a Foreign Language taught to French university students. *Phonetics Teaching and Learning Conference*, Paper ID 37.
- KARTUSHINA, N., HERVAIS-ADELMAN, A., FRAUENFELDER, U. H. & GOLESTANI, N. (2015). The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds. *The Journal of the Acoustical Society of America*, 138(2), 817-832. <https://doi.org/10.1121/1.4926561>
- KAWAHARA, S. (2016). Japanese has syllables : a reply to Labrune. *Phonology*, 33(1), 169-194.
- KAWAHARA, S. & SHINYA, T. (2008). The Intonation of Gapping and Coordination in Japanese : Evidence for Intonational Phrase and Utterance. *Phonetica*, 65(1-2), 62-105. <https://doi.org/10.1159/000130016>
- KAWAKAMI, S. (1963). Bunmatsu nado no jōshōchō ni tsuite [On sentence-final pitch rises]. *Kokugo Kenkyū*, (16), 25-46.

- KENDRICK, H. (1997). Keep them talking ! A project for improving students' L2 pronunciation. *System*, 25(4), 545-560. [https://doi.org/10.1016/S0346-251X\(97\)00043-2](https://doi.org/10.1016/S0346-251X(97)00043-2)
- KOISO, H. (2014). Nihongo jihatsuon ni okeru fukugō kyōkaionchō to tōgokōzō to no kankei [The relationship between BPM and syntactic structures in spontaneous Japanese]. *Onsei kenkyu*, 18(1), 57-69.
- KOISO, H., HORIUCHI, Y., TUTIYA, S., ICHIKAWA, A. & DEN, Y. (1998). An analysis of turn-taking and backchannels based on prosodic and syntactic features in Japanese map task dialogs. *Language and speech*, 41(3-4), 295-321.
- KOJIMA, K. (1978). Nihonjin gakusei no furansugohatsuonhō [notre traduction : La prononciation française des apprenants japonophones]. *La Societe Japonaise de Didactique du Français*, 6, 29-47. [https://doi.org/10.24495/efj.6.0\\_29](https://doi.org/10.24495/efj.6.0_29)
- KRASHEN, S. D. (1985). *The input hypothesis : Issues and implications*. Addison-Wesley Longman Limited.
- KUBOZONO, H. (2014). Japanese pitch accent in a typological perspective. *4th International Symposium on Tonal Aspects of Languages (TAL 2014)*, 195-198.
- KUSUMOTO, Y. (2015). Exploring the effects of shadowing on prosody. *EPIP4 4th International Conference on English Pronunciation : Issues & Practices*, 82-85.
- LABRUNE, L. (2006). *La phonologie du japonais*. Peeters Publishers.
- LADD, D. R. (2008). *Intonational Phonology* (2<sup>e</sup> éd.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511808814>
- LADEFOGED, P. (1975). *A Course in Phonetics*. New York : Harcourt Brace Jovanovich.
- LADO, R. (1957). *Linguistics across cultures : Applied linguistics for language teachers*. University of Michigan press.
- LEE, S. (2014). The Realization of French Rising Intonation by Native Speakers of American English. *Speech Prosody 2014*, 757-761. <https://doi.org/10.21437/SpeechProsody.2014-140>
- LENNEBERG, E. H. (1967). The biological foundations of language. *Hospital Practice*, 2(12), 59-67.
- LÉON, P. (2009). Nouveau regard sur la phonostylistique. *La linguistique*, 45(1), 159-170.
- LEPETIT, D. (1989). Cross-linguistic influence in intonation : French/Japanese and French/English. *Language Learning*, 39(3), 397-413.

- LI, P., BAILLS, F. & PRIETO, P. (2020). Observing and producing durational hand gestures facilitates the pronunciation of novel vowel-length contrasts. *Studies in Second Language Acquisition*, 42(5), 1015-1039.
- LONG, M. H. (1990). Maturational Constraints on Language Development. *Studies in Second Language Acquisition*, 12(3), 251-285. <https://doi.org/10.1017/S0272263100009165>
- LONG, M. H. (2003). Stabilization and Fossilization in Interlanguage Development. *The Handbook of Second Language Acquisition* (p. 487-535). John Wiley & Sons, Ltd. <https://doi.org/https://doi.org/10.1002/9780470756492.ch16>
- MAEKAWA, K. (2006). Analysis of language variation using a large-scale corpus of spontaneous speech. *International Symposium on Linguistic Patterns in Spontaneous Speech (LPSS)*, 27-50.
- MAEKAWA, K., KIKUCHI, H., IGARASHI, Y. & VENDITTI, J. (2002). X-JToBI : An eXtended J-ToBI for spontaneous speech, 1545-1548.
- MARCHAND, A. & LE GAC, D. (2020). Acquisition de la prosodie en langue additionnelle : l'accentuation en français par des locuteurs adultes plurilingues de langue initiale turque. *SHS Web of Conferences*, 78, 09007. <https://doi.org/10.1051/shsconf/20207809007>
- MARTIN, P. (1975). Analyse phonologique de la phrase française. *Linguistics*, 13(146). <https://doi.org/10.1515/ling.1975.13.146.35>
- MARTIN, P. (2006). Intonation du français : parole spontanée et parole lue. *Estudios de Fonetica experimental*, 134-162.
- MARTIN, P. (2008). L'intonation du français : le vilain petit canard parmi les langues romanes? *Language design : journal of theoretical and experimental linguistics*, (Special Issue), 1-13.
- MARTIN, P. (2018). *Intonation, structure prosodique et ondes cérébrales : introduction à l'analyse prosodique*.
- MARTIN, P. (2019). Enseignement de l'intonation en FLE aujourd'hui. *Recherches en didactique des langues et des cultures*, 16(1). <https://doi.org/10.4000/rdlc.4431>
- MARTINET, A. (1968). *La linguistique synchronique*. Presses universitaires de France (PUF).

- McKINNON, S. (2017). TBLT instructional effects on tonal alignment and pitch range in L2 Spanish imperatives versus declaratives. *Studies in Second Language Acquisition*, 39(2), 287-317. <https://doi.org/10.1017/S0272263116000267>
- MERTENS, P. (1993). Accentuation, intonation et morphosyntaxe. *Travaux de linguistique*, 26, 21-69.
- MERTENS, P. (2005). Le prosogramme. *Cahiers de l'Institut de Linguistique de Louvain*, 30(1), 7-25.
- MERTENS, P. (2008). Syntaxe, prosodie et structure informationnelle : une approche prédictive pour l'analyse de l'intonation dans le discours. *Travaux de linguistique*, (1), 97-124. <https://doi.org/10.3917/tl.056.0097>
- MERTENS, P. (2009). Prosodie, syntaxe, discours : autour d'une approche prédictive. *Actes du congrès Interfaces Discours Prosodie (IDP)*, 19-32.
- MICHELAS, A. (2011). *Caractérisation phonétique et phonologique du syntagme intermédiaire en français : de la production à la perception*. (Thèse de doctorat). Université de Provence - Aix-Marseille I.
- MORI, Y. (2011). Shadowing with oral reading : Effects of combined training on the improvement of Japanese EFL learners' prosody. *Language Education & Technology*, 48, 1-22.
- MOYER, A. (2004). *Age, accent, and experience in second language acquisition : an integrated approach to critical period inquiry*. Multilingual matters.
- MUNRO, M. J. & DERWING, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech the role of speaking rate. *Studies in second language acquisition*, 23(4), 451-468.
- NEMSER, W. (1971). Approximative systems of foreign language learners. *IRAL*, 9(2), 115-124. <https://doi.org/doi:10.1515/iral.1971.9.2.115>
- NIEBUHR, O., ALM, M., SCHÜMCHEN, N. & FISCHER, K. (2017). Comparing visualization techniques for learning second language prosody : First results. *International Journal of Learner Corpus Research*, 3(2), 250-277.
- OKUNO, T. & HARDISON, D. M. (2016). Perception-production link in L2 Japanese vowel duration : Training with technology. *Language Learning & Technology*, 20(2), 61-80.

- OTAKE, T., HATANO, G., CUTLER, A. & MEHLER, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of memory and language*, 32(2), 258-278.
- OYAMA, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, 5(3), 261-283. <https://doi.org/10.1007/BF01067377>
- PATKOWSKI, M. (1994). The critical age hypothesis and interlanguage phonology. *First and second language phonology*, 205-221.
- PAUKER, E., ITZHAK, I., BAUM, S. R. & STEINHAEUER, K. (2011). Effects of cooperating and conflicting prosody in spoken English garden path sentences : ERP evidence for the boundary deletion hypothesis. *Journal of Cognitive Neuroscience*, 23(10), 2731-2751.
- PEIRCE, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of neuroscience methods*, 162(1-2), 8-13.
- PÉLISSIER, M. (2018). *Effets d'entraînements explicites et implicites sur l'acquisition de la syntaxe de l'anglais par des apprenants francophones : étude en potentiels évoqués* (Thèse de doctorat). Université Sorbonne Paris Cité.
- PELLEGRINO, E. & VIGLIANO, D. (2015). Self-imitation in prosody training : a study on Japanese learners of Italian. *Workshop on Speech and Language Technology in Education (SLaTE)*, 53-57.
- PEŠKOVÁ, A. (2019). L2 Italian and L2 Spanish vocatives produced by L1 Czech learners : Transfer and prosodic overgeneralization. *Proceedings of the International Congress of Phonetic Sciences (ICPhS)*, 1932-1936.
- PIERREHUMBERT, J. & HIRSCHBERG, J. B. (1990). The Meaning of Intonational Contours in the Interpretation of Discourse. In P. R. COHEN, J. MORGAN & M. E. POLLACK (Éd.), *Intentions in communication* (p. 271-311). MIT Press.
- PIKE, K. L. (1945). *The Intonation of American English*. Ann Arbor : University of Michigan Press.
- PILLOT-LOISEAU, C. & XIE, H. (2018). Transfert rythmique du chinois mandarin au français dans l'apprentissage du Français Langue Étrangère : acoustique et perception. In F. NEVEU, B. HARMEGNIES, L. HRIBA & S. PRÉVOST (Éd.), *SHS Web of Conferences* (p. 09001). <https://doi.org/10.1051/shsconf/20184609001>

- PORTES, C. & BERTRAND, R. (2005). Some cues about the interactional value of the «continuation» contour in French. *Actes, Discours et Prosodie comme Interface Complexe (IDP)*, 1-14.
- PORTES, C., BERTRAND, R. & ESPESSER, R. (2007). Contribution to a grammar of intonation in French. Form and function of three rising patterns. *Nouveaux cahiers de linguistique française*, 28, 155-162.
- PUGA, K., FUCHS, R., SETTER, J. & MOK, P. (2017). The Perception of English Intonation Patterns by German L2 Speakers of English. *Interspeech 2017*, 3241-3245. <https://doi.org/10.21437/Interspeech.2017-1279>
- PURCELL, E. T. & SUTER, R. W. (1980). Predictors of pronunciation accuracy : A reexamination. *Language Learning*, 30(2), 271-287. <https://doi.org/10.1111/j.1467-1770.1980.tb00319.x>
- QUENÉ, H. & VAN DELFT, L. (2010). Non-native durational patterns decrease speech intelligibility. *Speech Communication*, 52(11-12), 911-918.
- R CORE TEAM. (2018). *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
- RAMÍREZ VERDUGO, D. (2006). A Study of Intonation Awareness and Learning in Non-native Speakers of English. *Language Awareness*, 15(3), 141-159. <https://doi.org/10.2167/la404.0>
- RAMUS, F., NESPOR, M. & MEHLER, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265-292.
- RASIER, L. & HILIGSMANN, P. (2007). Prosodic transfer from L1 to L2. Theoretical and methodological issues. *Nouveaux cahiers de linguistique française*, 28(2007), 41-66.
- RICHARDS, J. C. (1971). A Non-Contrastive Approach to Error Analysis. *ELT Journal*, 25(3), 204-219. <https://doi.org/10.1093/elt/XXV.3.204>
- RICHARDS, J. C. (2015). *Error analysis : Perspectives on second language acquisition*. Routledge.
- RONGNA, A. & HAYASHI, R. (2012). Accuracy of Japanese pitch accent rises during and after shadowing training. *Speech Prosody 2012*, 214-217.
- ROSSI, M. & DI CRISTO, A. (1980). Un modèle de détection automatique des frontières intonatives et syntaxiques. *Journées d'Etude sur la Parole (JEP)*, 213-238.

- ROSSI, M. (1980). Le français, langue sans accent ? *L'Accent en français contemporain (Studia Phonetica)*, 15, 13-51.
- ROSSI, M. (1999). *L'intonation : le système du français : description et modélisation*. Editions Ophrys.
- RUDE, M. (2012). Native-like duration ratio of stressed vs. unstressed syllables through visualizing prosody. *Speech Prosody 2012*, 254-257.
- SAITO, K. (2012). Effects of Instruction on L2 Pronunciation Development : A Synthesis of 15 Quasi-Experimental Intervention Studies. *TESOL Quarterly*, 46(4), 842-854. <https://doi.org/10.1002/tesq.67>
- SANDRYHAILA, D. (2019). *Visualisation de la mélodie de la parole dans l'enseignement du français langue étrangère* (Thèse de doctorat). Université de Paris.
- SANTIAGO, F. (2014). *Systèmes prosodiques et acquisition d'une L2 : production et perception des mouvements mélodiques en français et en espagnol* (Thèse de doctorat). Université Paris 7.
- SANTIAGO, F. & DELAIS-ROUSSARIE, E. (2012). La prosodie des énoncés interrogatifs en français L2, 265.
- SANTIAGO, F. & DELAIS-ROUSSARIE, E. (2015). The Acquisition of Question Intonation by Mexican Spanish Learners of French. In E. DELAIS-ROUSSARIE, M. AVANZI & S. HERMENT (Éd.), *Prosody and Language in Contact* (p. 243-270). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-662-45168-7\\_12](https://doi.org/10.1007/978-3-662-45168-7_12)
- SASAKI-HARA, K. (2004). *Nihongo onsei danwa no inritsukōzō [Prosodic structure of Japanese discourse]* (Thèse de doctorat). Tokyo University of Foreign Studies (T.U.F.S.)
- SAUVAGE, J. & BILLIÈRES, M. (2019). Enseigner la phonétique d'une langue étrangère : Bilan et perspectives. *Recherches en didactique des langues et des cultures*, 16(1). <https://doi.org/10.4000/rdlc.4234>
- SCHACHTER, J. (1974). An error in error analysis 1. *Language learning*, 24(2), 205-214.
- SCHEGLOFF, E. A. (1982). Discourse as an interactional achievement : some uses of "uh huh" and other things that come between sentences. In D. TANNEN (Éd.), *Analyzing Discourse : Text and Talk* (p. 71-93). Georgetown University Press.
- SCHWAB, S. (2012). Do native Spanish speakers transfer accentual acoustic properties from Spanish to French L2? *Research on Prosody in Switzerland*.

- SCHWAB, S. & LLISTERRI, J. (2012). The role of the acoustic correlates of stress in the perception of Spanish accentual contrasts by French speakers. *Speech Prosody 2012*, 350-353.
- SELINKER, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10(3), 209-232.
- SELTING, M. (2000). The construction of units in conversational talk. *Language in Society*, 29(4), 477-517. <https://doi.org/10.1017/S0047404500004012>
- SHELDON, A. & STRANGE, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English : Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, 3(3), 243-261. <https://doi.org/10.1017/S0142716400001417>
- SHIKI, O., MORI, Y., KADOTA, S. & YOSHIDA, S. (2010). Exploring differences between shadowing and repeating practices : An analysis of reproduction rate and types of reproduced words. *ARELE : annual review of English language education in Japan*, 21, 81-90.
- SHIZUKA, T. (1992). Task variation and accuracy predictor in interlanguage phonology production. *The bulletin of the Kanto-koshin-etsu English Language Education Society*, 7, 63-79.
- SO, C. K. & BEST, C. T. (2014). Phonetic Influences on English and French Listeners' Assimilation of Mandarin Tones to Native Prosodic Categories. *Studies in Second Language Acquisition*, 36(2), 195-221.
- TAHTA, S., WOOD, M. & LOEWENTHAL, K. (1981). Foreign Accents : Factors Relating to Transfer of Accent from the First Language to a Second Language. *Language and Speech*, 24(3), 265-272. <https://doi.org/10.1177/002383098102400306>
- TARONE, E. (1980). Some influences on the syllable structure of interlanguage phonology. *International review of applied linguistics*, 18(2), 139-152.
- TARONE, E. (1982). Systematicity and attention in interlanguage. *Language Learning*, 32(1), 69-84. <https://doi.org/10.1111/j.1467-1770.1982.tb00519.x>
- TATEISHI, M. & WINTERS, S. (2013). Does ultrasound training lead to improved perception of a non-native sound contrast? Evidence from Japanese learners of English. *Proceedings of the 2013 annual conference of the Canadian Linguistic Association*, 1-15.

- TROUBETZKOY, N. S. (1949). *Principes de phonologie* (J. CANTINEAU, Trad.). Paris : C. Klincksieck.
- VAISSIÈRE, J. (2010). Le français, langue à frontières par excellence. In D. DELOMIER & M.-A. MOREL (Éd.), *Frontières, du linguistique au sémiotique* (p. 10-20). Lambert Lucas.
- van MAASTRICHT, L., KRAHMER, E. & SWERTS, M. (2016). Native speaker perceptions of (non-)native prominence patterns : Effects of deviance in pitch accent distributions on accentedness, comprehensibility, intelligibility, and nativeness. *Speech Communication*, 83, 21-33. <https://doi.org/10.1016/j.specom.2016.07.008>
- VENDITTI, J. J. (2005). The J\_ToBi Model of Japanese Intonation. *Prosodic Typology* (p. 172-200). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199249633.003.0007>
- VENDITTI, J. J., MAEDA, K. & SANTEN, J. P. v. (1998). Modeling Japanese boundary pitch movements for speech synthesis. *The third ESCA/COCOSDA Workshop (ETRW) on Speech Synthesis*, 317-322.
- VENDITTI, J. J., MAEKAWA, K. & BECKMAN, M. E. (2008). Prominence marking in the Japanese intonation system. In S. MIYAGAWA & S. SAITO (Éd.), *Handbook of Japanese linguistics* (p. 456-512). Oxford University Press.
- VERDUGO, D. R. & TRILLO, J. R. (2005). The pragmatic function of intonation in L2 discourse : English tag questions used by Spanish speakers. *Intercultural Pragmatics*, 2(2). <https://doi.org/10.1515/iprg.2005.2.2.151>
- WENK, B. J. & WIOLAND, F. (1982). Is French really syllable-timed? *Journal of phonetics*, 10(2), 193-216.
- WHITE, L. (1985). The “pro-drop” parameter in adult second language acquisition. *Language learning*, 35(1), 47-61.
- WODE, H. (1994). Nature, nurture, and age in language acquisition : The case of speech perception. *Studies in second language acquisition*, 16(3), 325-345.
- WU, M. (2017). *Perception and production of Cantonese tones by speakers with different linguistic experiences* (Thèse de doctorat). University of Melbourne.
- XI, L., WACHS, S. & RIDOUANE, R. (2020). Production of French final stressed syllables in Accentual Phrase by Chinese learners : A pilot study. *Speech Prosody 2020*, 895-899. <https://doi.org/10.21437/SpeechProsody.2020-183>

- YONEYAMA, K., FON, J. & KOISO, H. (2003). Durational and prosodic patterning at discourse boundaries in Japanese spontaneous monologs. *Proc. 15th International Congress of Phonetic Sciences*, 2637-2640.
- YOSHIMURA, Y. & MACWHINNEY, B. (2007). The effect of oral repetition on L2 speech fluency : An experimental tool and language tutor. *Workshop on Speech and Language Technology in Education*.
- YOUNG-SCHOLTEN, M. (1993). *The Acquisition of Prosodic Structure in a Second Language*. Max Niemeyer Verlag. <https://doi.org/doi:10.1515/9783111715254>
- YUAN, C., GONZÁLEZ-FUENTE, S., BAILLS, F. & PRIETO, P. (2019). Observing pitch gestures favors the learning of Spanish intonation by Mandarin speakers. *Studies in Second Language Acquisition*, 41(1), 5-32. <https://doi.org/10.1017/S0272263117000316>